HISTORICAL SOCIOLINGUISTICS

PETER LANG

Double Standards

Codified Norms and Norms of Usage in European Languages (1600–2020)

Stephan Elspaß, Imke Mendoza,
Bernhard Pöll and Erik Schleef (eds)

HISTORICAL SOCIOLINGUISTICS

Standardisation is not a natural form of language change. The aim of such efforts is usually to create a largely uniform standard, often backed by a strong standard language ideology. However, standard languages are sometimes Janus-faced. Additional standard varieties emerge, leading to situations with 'double standards' or even 'multiple standards'.

This volume brings together fifteen contributions on 'double standards' in twelve languages – and various standard varieties thereof – from three language families, thus providing a novel comparative view on standardisation. Three main aspects of the 'double standards' concept are identified: the descriptive vs prescriptive notion of standards, the coexistence of standard varieties in pluricentric or pluriareal settings, and the impact of the cohabitation of several standards in native or foreign language teaching.

Stephan Elspaß is Professor of German Linguistics at the University of Salzburg (Austria).

Imke Mendoza is Professor of Slavic Linguistics at the University of Salzburg (Austria).

Bernhard Pöll is Professor of Romance Linguistics at the University of Salzburg (Austria).

Erik Schleef is Professor of English Linguistics at the University of Salzburg (Austria).

ISBN 978-1-80374-940-2

9 781803 749402

www.peterlang.com

Double Standards

HISTORICAL SOCIOLINGUISTICS

Series editors

Nils Langer
Stephan Elspaß
Joseph Salmons
Wim Vandenbussche

Vol. 9

PETER LANG

Oxford · Berlin · Bruxelles · Chennai · Lausanne · New York

Double Standards

Codified Norms and Norms of Usage in European Languages (1600–2020)

Stephan Elspaß, Imke Mendoza
Bernhard Pöll and Erik Schleef (eds)

PETER LANG
Oxford · Berlin · Bruxelles · Chennai · Lausanne · New York

Bibliographic information published by the Deutsche Nationalbibliothek.
The German National Library lists this publication in the German National Bibliography; detailed bibliographic data is available on the Internet at http://dnb.d-nb.de.

A catalogue record for this book is available from the British Library.

Library of Congress Control Number: 2025008217

Open access publication supported by the Paris Lodron University of Salzburg Publication Fund.

Cover image: Double herm. Marble, Roman copy after a Greek original. Vatican museums. Photographer Marie-Lan Nguyen (2009). Available under a CC-BY 3.0 license, Wikimedia Commons.

Cover design by Peter Lang Group AG

ISSN 2296-1909
ISBN 978-1-80374-940-2 (Print)
ISBN 978-1-80374-944-0 (E-PDF)
ISBN 978-1-80374-945-7 (E-PUB)
DOI 10.3726/b22873

Published by Peter Lang Ltd, Oxford, United Kingdom
info@peterlang.com - www.peterlang.com

This publication has been peer reviewed.

Contents

Figures

Tables

Acknowledgements

The chapters in this volume are based on a selection of papers presented at the conference *Double standards: Codified norms and norms of usage in European languages (1600–2020)* held at the University of Salzburg in April 2022. Many individuals have made important contributions to this volume and have shaped the final outcome. First of all, we would like to thank the authors of the individual chapters for their contributions, as well as for their cooperation and patience in the editing process. We would also like to express our gratitude to several colleagues who acted as reviewers of the individual chapters and to two anonymous reviewers who provided very constructive advice on the draft volume. Many thanks go to Michael Ryan for proofreading the entire volume and for his useful advice on stylistic issues and to Lara Wlcek, Clara Buchegger and Olha Erdyv for assisting with the editing process of this book. We would also like to thank the staff at Peter Lang for their invaluable support, in particular Laurel Plapp, who has offered expert professional guidance and advice that was extremely helpful to us. Finally, we are grateful to our departments at the University of Salzburg – the Department of German Language and Literatures, the Slavic Department, the Department of Romance Studies and the Department of English & American Studies – for their financial support in publishing this volume.

Stephan Elspaß, Imke Mendoza, Bernhard Pöll and Erik Schleef
Salzburg, December 2024

STEPHAN ELSPAß, IMKE MENDOZA,
BERNHARD PÖLL AND ERIK SCHLEEF

1 Double Standards: Codified norms and norms of usage in European languages (1600–2020)

ABSTRACT
This chapter introduces the concept of 'double standards', which provides the title for the volume. Based on a critical reflection of Haugen's (1966) classic matrix model, the chapter focuses on socio-historical issues of standardisation in European languages over the last four centuries. Three main aspects of the 'double standards' concept are identified: the descriptive vs. prescriptive notion of standard, the coexistence of standard varieties in pluricentric or pluriareal settings, and the impact of the cohabitation of multiple standards in native or foreign language teaching. This is followed by a brief presentation of the contributions in this volume and their respective positioning in relation to these three aspects. The chapter concludes by arguing that a future 'comparative standardology' (Joseph 1987) can greatly benefit from a comparison of the standardisation histories of different languages (from different language families) if it systematically considers these three aspects as well as the role of language ideologies – not least with regard to current countervailing or corrective trends towards standardisation such as destandardisation, restandardisation or demotisation.

1. The controversial nature of standardisation

The histories of most European languages in the last two hundred to four hundred years (or even beyond) have been marked by the development of standard languages and the effects of standardisation on other language varieties. Standardisation, however, is not a natural form of language change. One characteristic of natural languages is their variability and changeability. In contrast, standardisation, especially when imposed 'from above' (Rutten & Vosters 2021), involves 'the suppression of optional variability' (Milroy & Milroy 2012: 6) by imposing 'uniformity upon a class of objects' (Milroy 2001: 31) and the ensuing retardation of change. This kind of standardisation is strongly linked to standard language ideology (Milroy 2001; Milroy & Milroy 2012; Lippi-Green 2012). Among the few attempts to describe the mechanisms of standardisation across languages, Haugen's (1966) matrix model of the formation of a standard variety within a historical language is the most widely received and influential. It identifies four aspects of standardisation: 'selection of norm', 'codification of form', 'elaboration of function' and 'acceptance by the community'. Selection and codification are the two processes that inherently work against variation and change, and as they can be linked to the work of language authorities who favour certain varieties and variants over others, standardisation can be a very conflict-laden process. Further controversies may arise from the notion of 'elaboration', which in the conventional understanding of Haugen's model is largely seen as being controlled by authorities in the education and cultural sectors but may just as well be applied to the functional spread of non-standard varieties, for example, as the 'elaboration of vernacular writing' in today's digital media (Androutsopoulos 2011: 153). Similarly, the aspect of 'acceptance' is by no means limited to standard languages – and even there, controversies can arise because 'the community' may not always agree on what should be accepted as the standard.

In recent years, as part of a revision of Haugen's work, scholars have increasingly raised the question of the extent to which his model is relevant and still up to date, e.g. from a historical-sociolinguistic perspective (e.g. Elspaß 2020; Rutten & Vosters 2020; Ayres-Bennett 2021: 51–6). The model has often been (mis)interpreted as a teleological phase model, and it has been criticised for disregarding the embedding of standardisation in Europe in national language discourses. Accordingly, the call for more advanced standardisation models which are able to adequately capture processes of diversification of standard languages (Hundt 2013), such as those discussed in pluricentricity and pluriareality research (cf. Meer & Durgasingh 2025a; Part III of this volume), standardisation in multilingual contexts (such as in Luxembourg, cf. Vári & Tamburelli in this volume), standardisation of languages without a nation-state (such as Catalan, cf. Calafat Vila in this volume), cases of failed standardisation (such as for Low German, cf. Langer 2001) or processes such as destandardisation (cf. Kristiansen 2021) has become more urgent.

Thus, the focus of standardisation studies in the last twenty-five years or so has been not only on linguistic but increasingly on socio-historical, political and socio-cultural circumstances that have (or have *not*) led to or are presently leading to the standardisation of varieties and variants that were previously only used regionally or in certain social groups. Standardisation processes of individual languages can differ considerably, ranging from processes of supra-localisation of formerly diatopic or diastratic spoken varieties, eventually resulting in written standard varieties (such as in the case of French, cf. Lodge 1999, 2004), to levelling processes of written varieties, forming the basis of spoken standards (such as in the case of German, cf. Mattheier 2003), or the merger of two (or more) varieties (as in Russian, cf. Keipert 1999). Moreover, there can be different opinions among speakers of languages – sometimes even within an individual language – as to what exactly should be recognised as standard language(s) – see the aspect of acceptance mentioned above. As a consequence of the diverse nature of standardisation histories and their contextualisation in the respective national language discourses, present-day language communities can also differ considerably with respect to speakers' attitudes towards standard varieties (cf. Lenz, Soukup & Koppensteiner 2022) and the underlying standard language ideologies (cf. Kristiansen & Coupland 2011; Milroy & Milroy 2012; Lippi-Green 2012).

The present volume is dedicated to the question of standardisation, primarily from a sociolinguistic – and in some contributions also explicitly historical-sociolinguistic – perspective, partly also taking into account discourse-linguistic and cognitive-linguistic aspects. The different understandings of what is and what can be considered 'standard' in languages are captured here with the concept of 'double standards'.

2. Double standards

The term 'standard' is truly Janus-faced in many respects, and it is worth taking a closer look at it from different angles. There are three aspects of the double – or even multiple – implications of 'standard' that the contributions in this volume address: firstly, the descriptive vs. prescriptive notion of standard, secondly, the coexistence of standard varieties in pluricentric or pluriareal settings, and, thirdly, the impact of the cohabitation of one or more standards for native or foreign language teaching.

2.1 Standards between codified norms and norms of usage

Joseph (1987) introduced the important distinction between the emergence of 'language standards' as a sociolinguistic universal and 'standard languages', i.e. 'a specifically Western concept that has been spread by cultural tradition' (Joseph 1987: 7). The emergence of language standards may be at the heart of standardisations 'from below' (Elspaß 2021), whereas 'standard languages' are the product of standardisation efforts 'from above' (Rutten & Vosters 2021). Related to this differentiation is the distinction between a notion of 'standard' which draws on descriptive norms of usage – and as with 'language standards' basically applies to and is inherent in any language variety – and a more prescriptive concept of 'standard' that is mainly based on man-made codified norms, which are defined within specific ideological frameworks of one nation-one language cultures. 'Standard languages' in the latter sense often originate in literary standards (cf. terms like 'Literatursprache', 'langue littéraire', 'литературный язык / literaturnyj jazyk'). Traditional conceptualisations of the standard languages are mostly based on written varieties, less often on spoken varieties (cf. Ayres-Bennett & Bellamy 2021: 4). The term 'double standards' can therefore refer to different notions of standard in spoken and written language, which not only apply to media-related differences (e.g. pronunciation norms vs. spelling norms), but also to different register-related norms of usage in grammar and lexis (between more informal and more formal registers, or between 'language of immediacy' and 'language of distance' according to Koch & Oesterreicher 2012). Due to a growing awareness of such disparities, efforts have been intensified in recent years, especially in historical sociolinguistics, to compare norms and usage(s) in the standardisation histories of European languages (cf. Rutten, Vosters & Vandenbussche 2014). The contributions in Part I of this volume are dedicated to the relationship between codified norms and norms of usage in the histories of French (Ayres-Bennett), German (Hübner), Slovenian (Jelovšek) and (American) English (Kostadinova) as well as such dual norms in present-day Catalan (Calafat Vila).

Wendy Ayres-Bennet's (University of Cambridge) contribution *The French Tradition: Double Standards?* is devoted to the normative situation in seventeenth-century France, a particularly fertile ground for investigating the complex relations between overt normative regulations and actual usage. Not only did this century see the foundation of the French Academy, but also the publication of influential prescriptive works. While there is often a clear contradiction between prescriptivism and linguistic reality, cases can be observed where seventeenth-century usage coincides with normative stipulations. This shows that prescriptive works also recorded contemporary usage, contrary to what a simplistic *a posteriori* analysis might suggest. Another important focus of this chapter is the question of how language change is reflected in different genres (letters, memoirs, novels, essays, theatre, poetry etc.). Which of these genres or text types is more representative of norms of usage than others remains an open question.

In her paper on *The German dative marker* -e *in the early modern period: Double standards between codified norms and norms of usage*, Julia Hübner (University of Hamburg) focuses on norms and usage(s) of a particular marker in the German nominal inflection system, i.e. the *-e*-suffix in the dative singular of masculine and neuter nouns. In Early New High German, the use and the omission of the dative *-e* coexisted, and the variation was subject to factors such as region and register. Previous research has often maintained that – parallel to the general levelling of cases – there was a tendency in Early New High German to drop the dative *-e*, while prescriptive grammars, which were oriented to the 'ideal' of classical languages, argued in favour of retaining inflectional marking. Hübner's analyses are based on a corpus of early modern textbooks for the teaching of foreign language which contain prescriptive norms in their grammatical part and examples of (albeit fictitious) language use in model dialogues. Hübner is able to demonstrate that the textbook authors applied 'double standards' by teaching both codified norms in the grammars and norms of usage in the model dialogues.

In *Which standard for Modern Czech? Three proposals from the turn of the eighteenth and nineteenth centuries*, Alena A. Fidlerová (Charles University Prague) discusses three approaches to a possible standardisation of Czech. She compares the model of Josef Dobrovský and his followers, which ultimately became successful and is now the basis of Modern (written) Standard Czech, to the proposals by Jan Václav Pohl and František Jan Tomsa with respect to linguistic features and the attitude of the authors towards older language stages and contemporary colloquial speech. Fidlerová shows that, contrary to the common narrative about the history of the Czech standard language, some of Tomsa's and Pohl's ideas have lasted until today. She also discusses the positions and the roles of the authors within the contemporary scholarly community and argues that Dobrovský's success was mainly due to his intellectual authority and, finally, the promotion of his concept by his student Václav Hanka.

Alenka Jelovšek (Research Centre of the Slovenian Academy of Sciences and Arts) looks at *Slovenian norms of address in the first half of the nineteenth century: Between purism and politeness*. She focuses on a transitional period in the history of address forms when two forms of polite address were competing. In the old system, the second plural was used in rural areas and the third plural in urban contexts. At the same time, the latter was considered to be more polite than the former. This system was challenged by purist grammarians, who perceived the third plural as a Germanism and suggested that it be replaced by the second plural. In her analysis of various types of documents authored by linguistic experts of the nineteenth century, Jelovšek shows that this typologically unusual shift from a more polite to a less polite form was difficult to implement and that even the grammarians themselves employed the third plural when addressing persons of higher rank or German-speaking individuals.

Three contributions in this part of the book present case studies of 'double standards' in present-day languages.

In their contribution *Norms of usage and codified norms manifesting in explicit and implicit attitudes? An evaluation of standard varieties in Luxembourg and Belgium*, Judit Vári (Ludwig Maximilian University Munich) & Marco Tamburelli (Bangor University) analyse how double-standard norms manifest themselves at the level of perception of the individual in a comparative study. One of the most interesting points about the comparison is that the norms of usage, in both cases based on Moselle-Franconian dialects, are relatively similar in the two neighbouring language communities that were studied, but that they are associated with the norms of different standards, namely Standard Luxembourgish in Luxembourg and Standard German in East Belgium. Standard Luxembourgish not only developed later than Standard German but is also linguistically closer to the vernacular in Luxembourg than Standard German is to non-standard varieties in East Belgium. The authors hypothesised that the differences between the language communities will result more prominently in speakers' explicit attitudes towards the standards than in their implicit attitudes towards the usage norms of the vernaculars. However, the results show that this was not the case. The authors conclude that the differences in the acceptance of usage norms were greater than assumed and need to be better controlled in future studies.

Viktorija Kostadinova's (University of Amsterdam) chapter contrasts *'Double Standards' in American English*, i.e. standards of American English prescribed in usage guides and actual language use. She focuses on ten variable features in three corpora, including recent online speech. Kostadinova investigates an aspect of standardisation that is not explicitly mentioned in any of Haugen's models, namely that of prescription, which other standardisation models, such as those of Milroy & Milroy (2012) and Nevalainen & Tieken-Boon van Ostade (2006) specifically include. However, she not only focusses on prescription alone, but also explores how prescription aligns with actual usage, thereby placing her research well into the scope of the current volume. Her results demonstrate a variety of scenarios that may occur in the tension between prescription and usage. Kostadinova sheds light on the factors that may result in prescription and usage aligning (e.g. *dived/dove*) or not aligning (e.g. *whom/who*). She argues that in these kinds of 'double-standard' situations, a comparison of edited written texts with spoken or online language can be particularly fruitful as different norms appear to be tied to specific contexts of use.

Rosa M. Calafat Vila's (University of the Balearic Islands) paper on *Lexicon and language ideologies: Double solutions for Catalan language use in the media* deals with the tensions between the traditional Catalan standard (purged of Spanish influence on the lexical and semantic levels) and a new media standard which is open to a form of spoken Catalan highly influenced by the dominant contact language. Calafat Vila first critically discusses examples of the latter model, in particular cases of lexical substitution and loss, semantic impoverishment and over-differentiation due to intensive language contact. She then analyses the way style guides deal with this situation and the public debate around the two models. Criticising the laxism of the media (reflected in the style guides), language professionals (university lecturers, linguists, editors) use social media as a forum to defend the traditional normative model and challenge the inclusion of hispanisms into the authoritative Catalan dictionary published under the auspices of the *Institut d'Estudis Catalans*.

The last contribution in this part of the volume is concerned with the consequences of standardisation in German. Simon Pröll (University of Freiburg i. Br.) looks at *The two faces of Standard German: Structural consequences of nativisation*. Pröll claims that recent decades have brought about a critical number of native speakers of Standard German for the first time in its history. This nativisation of the spoken standard is resulting in a situation in which the written and spoken standards are increasingly drifting apart, as the latter is now subject to variation and trajectories of change like any other natural language. Evidence comes from an ongoing change in the suprasegmental system of southern German speakers and from case syncretisms in the formal speech of TV newscasters in Germany.

2.2 Standards in 'pluricentric' settings

Another manifestation of double standards arises from the situation of pluricentric and pluriareal standard languages (for more recent conceptualisations cf. Pöll 2005; Auer 2014, 2021; Wide, Norrby & Oakes 2021; Elspaß 2025; Niehaus 2025). According to Clyne, '[t]he term *pluricentric(ity)* indicates that a language has more than one centre, i.e. several centres, each providing a national variety with its own norms' (Clyne 1989: 358). Critics have questioned the unclear status of centres within this notion of pluricentricity (cf. Auer 2021: 30–5), as it seems to equate centres with nations and it 'is not defined on the level of language use or the language representations of the lay speakers' (Auer 2013: 19). In contrast, the term '*pluriareal(ity)* indicates that a language has more than one standard variety, each with its own norms of usage. There may be more than one standard variety within a nation, and, in contiguous language areas, standard varieties may also transcend national borders'. (Elspaß 2025). A prominent example is the coexistence of *Nynorsk* and *Bokmål* as equally taught standard varieties in Norway, which are 'linguistically very close, but sociolinguistically clearly different' due to 'the two opposed social bases' on which they were constructed in the nineteenth century (Jahr 2003: 351). The two concepts do not seem to be mutually exclusive but are actually compatible to an extent (Durgasingh & Meer 2025; Meer & Durgasingh 2025b; Niehaus 2025). It follows from both concepts that within languages there can be two (e.g. in Dutch, Norwegian) or more (e.g. English, Spanish, German) standard varieties. A problem of 'double standards' arises here from the question of whether there can also be several standard varieties or more than one centre within a country. Apart from the question of the spatial distribution of standard varieties, there is the problem of hierarchisation and the recognition – or 'acceptance', to take up the fourth aspect in Haugen's model – of different standard varieties and standard variants in a particular context, e.g. whether a dominant or non-dominant standard variety and its variants are preferred in a particular situation (cf. Clyne 1992: 459). Many of the world's standard languages are asymmetrically pluricentric, i.e. acceptance is not equally pronounced for the various standard varieties. While this is certainly the case for Spanish or German, a similar point can be made for English (for non-dominant varieties), Portuguese, French and, to a certain degree, also Russian. These issues undoubtedly have an important impact on both mother tongue and foreign language teaching (see Section 2.3). Finally, the question arises as to what extent the modelling of standardisation or the revision of existing models takes into account the *historical* pluricentricity and/or pluriareality and the pluricentric and/or pluriareal development of today's standard languages.

Three papers – on Spanish (Tacke), German (Niehaus) and Russian (Wingender) – address the problem of double (or more) standards arising from present-day pluricentric or pluriareal settings.

Felix Tacke (University of Marburg) presents *A corpus-driven approach to emerging national standard varieties in Spanish: The Corpus Radiofónico Panhispánico (CO.RA.PAN).* Tacke presents the conception and the goals of a new, ongoing project that aims at (partly) overcoming the lack of empirical studies on the pluricentric nature of Spanish. His *Corpus Radiofónico Panhispánico* (CO.RA.PAN) consists of recordings of 'model speakers' (newscasters of major radio stations in Spanish-speaking America) and provides authentic oral speech data in formal situations, viz. where the respective standard or near-standard variety is realised. The value of the project lies, among other things, in the fact that it allows the different status of widespread variants to be determined with regard to their degree of 'standardness'. The aspiration of syllable-final /s/ is a good example. It is attested in several countries (e.g. Argentina, Chile, Cuba, Uruguay), but while it is part of the (broadcasting) standard in Argentina, this is not the case in Uruguay, as the data show.

Konstantin Niehaus (University of Salzburg) turns our attention to the question of *Double standards in Austria? Enregistering 'Austrian German' and 'German German' in the late twentieth century*. 'Double standards' manifest themselves here in at least two ways: firstly, in different regional standard varieties (or 'standards of usage'), which are due not least to the linguistic differences in usage norms in the west and east of the country. Secondly, they arise from the (apparent) competition between two national standard varieties of German in Austria.

Unlike Tacke, Niehaus is not so much interested in *describing* varieties on the basis of measurable linguistic differences, but rather in the discursive construction of national varieties. He uses the framework of 'enregisterment' 'to analyse how "Austrian German" has come to be recognised as a socially meaningful semiotic register which indexes national identity'.

MONIKA WINGENDER (University of Gießen) discusses types of *Slavic standard languages – with a focus on the Russian standard language between mono- and pluricentricity*. Given the high number of speakers of Russian and its large geographic extension, the question whether Russian is a pluricentric language arises naturally. While the idea of pluricentricity runs counter to the original, strictly monocentric conception of the Russian standard language, the collapse of the Soviet Union gave rise to regional variants of Russian in the post-Soviet states that have a sizeable Russian-speaking community or use Russian as an official language, such as Belarus, Kazakhstan, Kyrgyzstan and Ukraine. Wingender explores the possibilities of the emergence of national varieties of Russian in these countries and finds that to date there have been no serious attempts at norm setting or codification. The author pays special attention to Ukraine, where the Russian invasion in February 2022 ended all debates about a Ukrainian variant of Russian and Russian became the language of the enemy.

2.3 (Competing) standards in formal education

The contributions to Part III of this volume are concerned with approaches to and the management of different standard varieties in formal education. We can distinguish between the coexistence (or competition) of different standards at the diatopic level, which we discussed in the previous section, and competing standards on the level of (in)formality, with informal, vernacular registers at one end point and/or formal registers on the other end point of a 'standard' scale.

2.3.1 *Informal vs. formal standards*

A classic example of competing standards is Greek, with *Demotic* Greek, which emerged from vernacular registers, and *Katharevousa*, the cultivated literary language based on Ancient Greek, as opponents. *Demotic* Greek was only established as the sole language of the education system in 1976 after decades of public debate and a long phase of diglossia (Frangoudaki 2008). Recent cases of competing standards, which arise from the fact that spoken standard varieties are moving further and further away from traditional written standard varieties (cf. Haas 1986; Elspaß & Pöll 2024), harbour particular potential for conflict in L1 teaching.

In her chapter on *Competing standards in university students' formal production: A corpus-based study*, Silvia Ballarè (University of Bologna) presents such a case study from Italian. She contrasts students' written productions in formal genres with existing standard language corpora. Her analyses highlight the tensions between two normative models that determine the shape of present-day formal written Italian. The presence of and competition between the old literary standard, which still exerts some influence, and the so-called neo-standard, which incorporates features traditionally considered to be non-standard, gives rise to a 'heterogeneous variety' characterised by hypercorrections and a mixture of features pertaining to either the older or the new standard language model. It is unclear if this 'third way' will eventually turn into a more coherent scenario, viz. a stable variety in between the old literary standard and the neo-standard.

Heimir F. Viðarsson's (University of Iceland) contribution, *Doubling down on the scope and shape of standard norms: The role of compulsory education (1907–) in changing the tone of Icelandic language policy*, explores a proposal by Ottósson (1990), who has argued that Iceland's introduction of compulsory schooling in 1907 resulted in a more direct and stringent tone of Icelandic grammar books as they addressed a new, young audience. While the standardisation of Icelandic is rooted in the nineteenth century, Viðarsson investigates the implementation of Standard Icelandic in the twentieth-century education system and

Icelandic language policy. He analyses twelve grammar books published before and after the Education Act of 1907 and finds evidence in the grammars of increasingly explicit prescriptive remarks. He is also able to document a broadening in focus towards spoken Icelandic. Yet, as many other studies have previously shown, there is also evidence that the effect of Icelandic standardisation and prescriptive grammars has been overestimated.

2.3.2 Pluricentric and pluriareal standard languages in foreign language teaching

Further problems and/or controversies on competing standard varieties in the education sector arise from pluricentric or pluriareal situations on the diatopic level and the sometimes thin political line between standard varieties of one pluricentric language and the emergence of new standard varieties (and the ensuing discussions about their recognition, e.g. in the case of Bosnian, Croatian, Montenegrin and Serbian, Catalan and Valencian, etc.). The 'asymmetric pluricentricity' mentioned in the previous section is mirrored, for instance, by textbooks: they usually come from, or are based upon, the dominant variety of a language, e.g. in the case of Spanish it is 'Peninsular Spanish' (cf. Corti & Pöll 2017, 2023) and in the case of German it is '(Northern) German German' (cf. Ender & Shafer accepted). A systematic consideration and reflection of the impact of the pluricentricity/pluriareality of standard languages on teaching and learning these languages is still rather the exception than the rule (cf. Ruck & Shafer 2020). This situation is potentially conflictual in at least two respects – and can be demonstrated again using the examples of Spanish and German. First, German and Spanish are not only taught as (second) languages in Germany and Spain (and other Spanish- and German-speaking countries), but also in countries where the target language is, for the most part, devoid of authentic communicative functions (German as a Foreign Language/*DaF*; Spanish as a Foreign Language/*ELE*) and in which a monolithic view on standard languages hinders learners from establishing an adequate understanding of language variation in real life; second, teachers of German and Spanish may be native speakers of non-dominant varieties, e.g. speakers of a variety of 'Austrian German' or 'Cuban Spanish', or non-native speakers of the target languages.

Research into foreign language acquisition – that is, studying a pluricentric language in a country where the target language is *not* the dominant language – has revealed that monocentric norms continue to prevail (e.g. Bieswanger 2008; Hägi 2006; Ransmayr 2006, 2012; Schmidlin 2019). In contrast, in the second language context, local standards are inherently present. This immediate access to locally shaped standard norms may have a strong impact on increased awareness of pluricentricity and language variation. While some small-scale studies have explored what teachers or students know about language variation in these contexts (e.g. Clark & Schleef 2010 for English; Ortiz-Jiménez 2019 for Spanish; Racine, Schwab & Detey 2013 for French), there is still a remarkable lack of systematic study across the different (dominating and non-dominating) centres of these language areas and a much needed reconsideration of language standards, e.g. when teaching German in Germany or Switzerland, English in England or South Africa, French in France or Quebec, Portuguese in Portugal or Brazil, Spanish in Spain or Argentina, etc.

In her contribution *A prescriptive source with a spoken language imprint? Tense and aspect in Argentinean language textbooks*, Guro Nore Fløgstad (University of Southeast Norway) examines Argentinean textbooks for the teaching of Spanish as a first language (in elementary schools) with respect to a feature that singles out Spanish in Argentina, viz. the marked preference for the synthetic preterite and its heavy use to the detriment of the compound perfect, especially, but not exclusively, in children and adolescents. She can show that, while textbooks in general are conservative and forbid – or simply ignore – regionalisms, those in use in Argentina surprisingly reflect the actual usage of this specific feature. There is no adherence to a possible supranational Latin American norm (where this feature exists too, but is less pervasive than in Argentina) or to a Peninsular normative model (according to which the Argentinean usage would appear as deviant).

3. Comparative standardology

Research on standardisation has often been conducted within individual philologies. However, this view 'from within', especially when overlaid with national historical narratives, obscures a more detached view of standardisation. A comparative perspective is therefore necessary in order to contextualise the socio-historical narratives of standardisation. This volume therefore takes a further step towards a 'comparative standardology' (Joseph 1987: 13) by bringing together studies on standardisation and 'double standards' in twelve languages – and various standard varieties thereof – from three language families. So far, comparative work has been conducted mostly within these large language families, i.e. the Germanic (e.g. Deumert & Vandenbussche 2003), Romance (e.g. Laferl & Pöll 2007; Lebsanft & Tacke 2020) and Slavic languages (e.g. Müller & Wingender 2013). One aim of this volume, however, is to allow for a comparative perspective *across* language families (cf. Elspaß & Pöll 2024). The main chapters of the volume are therefore not organised according to languages or language families, but according to the thematic focal points presented in Part 2. The first insight from this overview is that there are topics that are not specific to the standardisation processes and their sociolinguistic implications in individual languages or language families. The second general conclusion that can be drawn is that studies on standardisation can always benefit from looking beyond the confines of individual languages. This is why we call for more direct comparisons of standardisation processes in two or more varieties. This volume can serve as a springboard for such comparisons as we have identified themes and tendencies that permeate standardisation processes discussed in this edition.[1] Particularly, the three aspects of 'double standards' discussed

1 The volume is limited to European languages (and some of their North American and South American standard varieties); however, this is not to say that standardisation is a purely European phenomenon (see works on non-European and post-colonial settings, e.g. Calabrese, Chambers & Leitner 2015).

here highlight three central themes that appear to be worth addressing and exploring in greater depth in future studies in comparative standardology:

- the oscillation of standard languages between codified norms and usage norms,
- the coexistence of standard varieties in pluricentric or pluriareal constellations,
- the effects of the coexistence of one or more standards in L1 or L2 teaching.

Further topics for future research include the following:

- the role of standard language ideology and other language ideologies such as homogenism and purism in standardisation processes (e.g. Thomas 1991; Cameron 2012; Langer & Davies 2005),
- destandardisation, demotisation, or restandardisation processes as counter-movements or corrective forces to standardisation (e.g. Coupland & Kristiansen 2011; Kristiansen 2021),
- the extent to which the aforementioned aspects and processes may have different effects on 'old' vs. 'new', on 'big' vs. 'small' standard languages or on dominant vs. non-dominant standard varieties.

In line with the three aspects of 'double standards' addressed in this book, the chapters show the wealth of different methods of data collection and data analysis in current standardisation research, ranging from quantitative to qualitative and also mixed approaches. The volume focusses primarily on linguistic aspects of standardisation processes. For future research, it would be beneficial to incorporate approaches from anthropological linguistics and to employ ethnographic methods to a greater extent in research on 'double standards', e.g. Woolards's work on 'Double talk' and the role of language ideologies in the context of the status of Catalan and Castilian in twenty-first-century Spain (Woolard 1989, 2016).[2] In

2 We would like to thank an anonymous reviewer for drawing our attention to the contribution of anthropological linguistics to and ethnographic work on standardisation.

particular, it can be instructive to explore the representations of different standards in the minds of the speakers. For obvious reasons, such studies could of course only be related to current standardisation processes, not to historical ones.

In conclusion, a comparative approach is essential in order to understand more general principles of standardisation that go beyond the particular cases of individual languages in their specific socio-historical contexts. Principles identified in this way can refine standardisation models and, in theoretical terms, enhance their generalisability and predictability. Such refined models can then inform research into current and future standardisation processes and their consequences for language teaching.

Bibliography

Auer, Peter. 2014. Enregistering pluricentric German. In Augusto Soares da Silva (ed.), *Pluricentricity: Language Variation and Sociocognitive Dimensions*. Berlin, Boston: de Gruyter Mouton. 19–48.

Auer, Peter. 2021. Reflections on Linguistic Pluricentricity. *Sociolinguistica. European Journal of Sociolinguistics* 35, 29–47.

Ayres-Bennett, Wendy. 2021. Introduction. In Wendy Ayres-Bennett & John Bellamy (eds), *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 1–24.

Ayres-Bennett, Wendy. 2021. Modelling language standardization. In Wendy Ayres-Bennett & John Bellamy (eds), *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 27–64.

Ayres-Bennett, Wendy & John Bellamy (eds). 2021. *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press.

Bieswanger, Markus. 2008. Varieties of English in current English language teaching. *Stellenbosch Papers in Linguistics* 38, 27–47.

Calabrese, Rita, J. K. Chambers & Gerhard Leitner (eds). 2015. *Variation and Change in Postcolonial Contexts*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Cameron, Deborah. 1995. *Verbal Hygiene*. London New York: Routledge.

Clark, Lynn & Erik Schleef. 2010. The acquisition of sociolinguistic evaluations among Polish-born adolescents learning English: Evidence from perception. *Language Awareness* 19(4), 299–322.

Clyne, Michael. 1989. Pluricentricity: National variety. In Ulrich Ammon (ed.), *Status and Function of Languages and Language Varieties*. Berlin, New York: de Gruyter. 357–71.

Clyne, Michael (ed.). 1992. *Pluricentric Languages: Differing Norms in Different Nations*. Berlin, New York: Mouton de Gruyter.

Corti, Agustín & Bernhard Pöll. 2017. Pluralidad lingüística y cultural en la formación de profesores de E/LE: estado de la cuestión y desiderata. In Eva Leitzke-Ungerer & Claudia Polzin-Haumann (eds), *Varietäten des Spanischen im Fremdsprachenunterricht. Ihre Rolle in Schule, Hochschule, Lehrerbildung und Sprachzertifikaten*. Stuttgart: ibidem. 225–42.

Corti, Agustín & Bernhard Pöll. 2023. Moving from awareness and understanding to tolerance and promotion? On Spanish teacher training students' knowledge of and attitudes towards language variation. In Marcus Callies & Stefanie Hehner (eds), *Pluricentric Languages and Language Education: Pedagogical Implications and Innovative Approaches to Language Teaching*. London: Routledge. 57–72.

Coupland, Nikolas & Tore Kristiansen. 2011. SLICE: Critical perspectives on language (de)standardization. In Tore Kristiansen & Nikolas Coupland (eds), *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus. 11–35

Deumert, Ana & Wim Vandenbussche (eds). 2003. *Germanic Standardizations. Past to Present*. Amsterdam, Philadelphia: Benjamins.

Durgasingh, Ryan & Philipp Meer. 2025. Pluricentricity AND pluriareality: Building the case for complementarity. In Philipp Meer & Ryan Durgasingh (eds), *Pluricentricity and Pluriareality: Dialects, Variation, and Standards*. Amsterdam, Philadelphia: Benjamins. 187–95.

Elspaß, Stephan. 2020. Alternative sources of data for alternative histories of standardisation. *Language Policy* 19(2), 281–99.

Elspaß, Stephan. 2021. Language Standardization in a View 'from Below'. In Wendy Ayres-Bennett & John Bellamy (eds), *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 93–114.

Elspaß, Stephan. 2025. Pluriareal languages and the case of German. In Philipp Meer & Ryan Durgasingh (eds), *Pluricentricity and Pluriareality: Dialects, Variation, and Standards*. Amsterdam, Philadelphia: Benjamins, 15–44.

Elspaß, Stephan & Bernhard Pöll. 2024. Between formal and informal standards. 'Double standards' in French, German and (Brazilian) Portuguese compared. *Energeia. Online Journal for Linguistics, Language Philosophy and History* *of Linguistics* 9, 166–205. (Special Issue: *Informal Standardization and the Regionalization of Language Norms*, ed. Felix Tacke.)

Ender, Andrea & Naomi Shafer. Accepted. German as a second/foreign language. In Joshua Bousquette & Simon Pickl (eds), *The Oxford Handbook of the German Language*. Oxford: OUP.

Frangoudaki, Anna. 2008. Diglossia and the present language situation in Greece: A sociological approach to the interpretation of diglossia and some hypotheses on today's linguistic reality. *Language in Society* 21(3), 365–81.

Haas, William H. (ed.). 1986. *Standard Languages: Spoken and Written*. Manchester: Manchester University Press.

Hägi, Sara. 2006. *Nationale Varietäten im Unterricht Deutsch als Fremdsprache*. Frankfurt: Lang.

Haugen, Einar. 1966. Dialect, Language, Nation. *American Anthropologist* 68(4), 922–35.

Hundt, Marianne. 2013. The diversification of English: Old, new and emerging epicentres. In Daniel Schreier & Marianne Hundt (eds), *English as a Contact Language*. Cambridge: Cambridge University Press. 182–203.

Jahr, Ernst Håkon. 2003. Norwegian. In Ana Deumert & Wim Vandenbussche (eds), *Germanic Standardizations. Past to Present*. Amsterdam, Philadelphia: Benjamins. 331–53.

Joseph, John E. 1987. *Eloquence and Power. The Rise of Language Standards and Standard Languages*. London: Pinter.

Keipert, Helmut. 1999. Geschichte der russischen Literatursprache. In Helmut Jachnow (ed.), *Handbuch der sprachwissenschaftlichen Russistik und ihrer Grenzdisziplinen*. Wiesbaden: Harrasowitz. 726–79.

Koch, Peter & Wulf Oesterreicher. 2012. Language of immediacy – language of distance: Orality and literacy from the perspective of language theory and linguistic history. In Claudia Lange, Beatrix Weber & Göran Wolf (eds), *Communicative Spaces: Variation, Contact, and Change. Papers in Honour of Ursula Schaefer*. Frankfurt: Lang. 441–73.

Kristiansen, Tore. 2021. Destandardization. In Wendy Ayres-Bennett & John Bellamy (eds), *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 667–90.

Kristiansen, Tore & Nikolas Coupland (eds). 2011. *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus Press.

Kunnan, Antony John & Nick Saville. 2021. Setting standards for language learning and assessment in educational contexts: A multilingual perspective. In Wendy Ayres-Bennett & John Bellamy (eds), *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 496–516.

Laferl, Christopher F. & Bernhard Pöll (eds). 2007. *Amerika und die Norm: Literatursprache als Modell?* Tübingen: Niemeyer.

Langer, Nils. 2003. Low German. In Ana Deumert & Wim Vandenbussche (eds.). *Germanic Standardizations. Past to Present*. Amsterdam: Benjamins. 282–302.

Langer, Nils & Winifred V. Davies (eds). 2005. *Linguistic Purism in the Germanic Languages*. New York: de Gruyter.

Lebsanft, Franz & Felix Tacke (eds). 2020. *Manual of Standardization in the Romance Languages*. Berlin, Boston: de Gruyter.

Lippi-Green, Rosina. 2012. *English with an Accent: Language, Ideology, and Discrimination in the United States*. 2nd ed. London, New York: Routledge.

Lodge, R. Anthony. 1993. *French: From Dialect to Standard*. London: Routledge.

Lodge, R. Anthony. 2004. *A Sociolinguistic History of Parisian French*. Cambridge: Cambridge University Press.

Mattheier, Klaus J. 2003. German. In Ana Deumert & Wim Vandenbussche (eds), *Germanic Standardizations. Past to Present*. Amsterdam, Philadelphia: Benjamins. 212–44.

Meer, Philipp & Ryan Durgasingh (eds). 2025a. *Pluricentricity and Pluriareality: Dialects, Variation, and Standards*. Amsterdam, Philadelphia: Benjamins.

Meer, Philipp & Ryan Durgasingh. 2025b. Modeling variation: Pluricentricity and pluriareality, the debate surrounding both models, and potentials for their complementarity. In Philipp Meer & Ryan Durgasingh (eds), *Pluricentricity and Pluriareality: Dialects, Variation, and Standards*. Amsterdam, Philadelphia: Benjamins, 1–14.

Milroy, James. 2001. Language ideologies and the consequences of standardization. *Journal of Sociolinguistics* 5(4), 530–55.

Milroy, James & Lesley Milroy. 2012. *Authority in Language. Investigating Language Prescription and Standardisation* (4th ed.). London: Routledge.

Nevalainen, Terttu & Ingrid Tieken-Boon van Ostade. 2006. Standardisation. In Richard Hogg & David Denison (eds), *A History of the English Language*. Cambridge: Cambridge University Press. 271–311.

Niehaus, Konstantin. 2025. Regiocentric use and national indexicality: Enregisterment as a theoretical integration for standard German. In Philipp Meer & Ryan Durgasingh (eds), *Pluricentricity and Pluriareality: Dialects, Variation, and Standards*. Amsterdam, Philadelphia: Benjamins, 66–89.

Ortiz-Jiménez, Macarena. 2019. Actitudes lingüísticas de los profesores de español en España y Australia hacia las variedades dialectales. *Journal of Spanish Language Teaching* 6(2), 182–98.

Ottósson, Kjartan G. 1990. *Íslensk málhreinsun: Sögulegt yfirlit*. Reykjavík: Íslensk málnefnd.

Pöll, Bernhard. 2005. *Le français langue pluricentrique? Études sur la variation diatopique d'une langue standard*. Frankfurt: Lang.

Racine, Isabelle, Sandra Schwab & Sylvain Detey. 2013. Accent(s) suisses(s) ou standard(s)? Approche perceptive dans quatre régions de Suisse romande. In Anika Falkert (ed.), *La perception des accents du français hors de France*. Mons: Éd. CIPA. 41–59.

Ransmayr, Jutta. 2006. *Der Status des österreichischen Deutsch an nichtdeutschsprachigen Universitäten. Eine empirische Untersuchung*. Frankfurt, Bern: Lang.

Ruck, Julia & Naomi Shafer (eds). 2020. *National Standards – Local Varieties: A Cross-linguistic Discussion on Regional Variation in Foreign Language Studies* [Special Issue]. *Critical Multilingualism Studies* 8(1).

Rutten, Gijsbert & Rik Vosters. 2020. *Revisiting Haugen: Historical-Sociolinguistic Perspectives on Standardization* [Special Issue]. *Language Policy* 19(2).

Rutten, Gijsbert & Rik Vosters. 2021. Language standardization 'from above'. In Wendy Ayres-Bennett & John Bellamy (eds*), The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 65–92.

Rutten, Gijsbert, Rik Vosters & Wim Vandenbussche (eds). 2014. *Norms and Usage in Language History, 1600–1900. A sociolinguistic and comparative perspective.* Amsterdam, Philadelphia: Benjamins.

Scharloth, Joachim. 2005. Asymmetrische Plurizentrizität und Sprachbewusstsein. Einstellungen der Deutschschweizer zum Standarddeutschen. *Zeitschrift für Germanistische Linguistik* 33(2), 236–67.

Schmidlin, Regula. 2019. Stand und Perspektiven linguistischer Plurizentrik und mögliche Konsequenzen für Deutsch als Fremdsprache. In Elisabeth Peyer, Thomas Studer & Ingo Thonhauser (eds), *IDT 2017*. Vol. 1: *Hauptvorträge*. Berlin: Schmidt. 35–47.

Thomas, George. 1991. *Linguistic Purism*. London, New York: Longman.

Wide, Camilla, Catrin Norrby & Leigh Oakes (eds). 2021. *New Perspectives on Pluricentricity*. [Special Issue]. *Sociolinguistica* 35.

Woolard, Kathryn. 1989. *Double Talk: Bilingualism and the Politics of Ethnicity in Catalonia*. Stanford: Stanford University Press.

Woolard, Kathryn. 2016. *Singular and Plural: Ideologies of Linguistic Authority in 21st Century Catalonia*. Oxford: Oxford University Press.

PART I

Standards between codified norms and norms of usage

Wendy Ayres-Bennett

2 The French tradition: Double standards?

ABSTRACT
The French grammatical tradition has long been considered the most prescriptive in Western Europe, making it particularly fertile ground for exploring the relationship between codified norms and norms of usage. My examples come from seventeenth-century France, the period when the French Academy, symbol of linguistic prescriptivism, was founded. As background, I discuss the relevant terminology and suggest some possible refinements to the way prescriptive texts are analysed. In the main part of the article I demonstrate how, when so-called prescriptive texts are examined in detail, the relationship between prescriptive norms and norms of usage is more complex than might be expected. While prescriptive norms do at times contradict norms of usage, they also frequently appear to reflect them, even when the metalanguage used is prescriptive. Finally, I consider how linguistic change spreads through different genres in a large multi-genre corpus and how far this informs us about which text types constitute the best source for historical sociolinguists.

1. Introduction

The French grammatical tradition has long been considered the most prescriptive in Western Europe (Brunot 1909: 4; Lodge 1993: 153, 159). Indeed, the leading contemporary linguist Bernard Cerquiglini (2012: 5) has gone so far as to argue that language in France has taken on the role of state religion. The origins of this prescriptivism are often said to lie in the foundation of the French Academy in 1634, followed in the following decades of the seventeenth century by the publication of some of the foundational metalinguistic texts by Vaugelas (1647), Bouhours (1674, 1675, 1692) and Andry de Boisregard (1689, 1693) amongst others. As a result, Sampson (1993: 7) claims not only that few languages have been exposed in such a sustained way to prescriptive influences as French, but also that for the past four centuries many of the various bodies and individuals seeking to 'direct' the language 'have commanded, and continue to command, very considerable attention and favour among the French'. Similarly, Nigel Armstrong and Ian Mackenzie (2013: 6) speak about 'the internalization by [French] speakers of an unusually normative reflex of the ideology of the standard', so that, in their view, the French situation 'illustrates the standardizing principle in a particularly vivid form'. This contrasts with the findings of Olivia Walsh (2016): only 19% of the French respondents to her survey were aware of French linguistic legislation today (compared with 100% of Québécois), and only 8% could name the *Loi Toubon*, which is a major piece of legislation that makes the use of French mandatory in education, commerce, audiovisual media, the workplace and public meetings, and which states that all public service contracts may not contain any foreign terms or expressions if an equivalent French term exists.[1] Moreover, while the legislation affects usage in the public sphere, it appears to have little or no impact on private and especially spoken usage.

In this article, I examine the validity of the claims about the strength of the normative tradition in France, and in particular consider the relationship between codified norms and norms of usage, using examples from the French seventeenth century to illustrate my arguments. In relation to the notion of 'double standards' which are the subject of this volume, I ask whether the codified norms in France (i) reflect norms of usage, (ii) drive norms of usage, or whether (iii) the codified norms and norms of usage are simply double tracks which do not really impinge on each other.

1 <https://www.legifrance.gouv.fr/loda/id/LEGITEXT000005616341>, accessed 8 June 2024.

As background to the current study, in Section 2 I start by presenting some of the terminology concerning 'double standards' and my understanding of it, as set out in previous publications (Ayres-Bennett 2016, 2020, 2021). In the main part of the article I explore the validity of the descriptive – prescriptive dichotomy through a number of seventeenth-century French case studies, looking in particular at the extent to which codified norms and norms of usage are in sync or not. I further consider the extent to which we can use corpus data to track how changing usage spreads through different genres or texts types, in particular considering whether this allows us to determine whether the change – or the emerging norms of usage – comes 'from above' (Rutten & Vosters 2021) or 'from below' (Elspaß 2021), and the extent to which multi-genre corpora can be useful in addressing such questions.

2. Terminology and the analysis of prescriptive texts

The terminology to describe different types of linguistic norms – notably codification, prescription and purism – has been subject to considerable variation between different scholars. In Haugen's classic model of standardisation in its first iteration (Haugen 1966), of the three terms mentioned above only codification is included. The same is true of his revised 1987 version (Table 2.1) although more detail is given about how codification – glossed as 'standardization procedures' and seen as part of corpus planning – is conceived, with a distinction made between the reduction of language to writing (graphization), the formulation of grammatical rules (grammatication), and the selection of the appropriate lexicon (lexication).

James and Lesley Milroy, on the other hand, include both codification and prescription in their model (Milroy & Milroy 1991: 27), a distinction which I also adopt. According to my definition, codification involves the production of grammars, dictionaries and other metalinguistic works. Metalinguistic texts can then be characterised as being either descriptive or prescriptive in nature. Descriptive texts are based on

Table 2.1 Haugen's 1987 model of language planning.

	Form (policy planning)	Function (cultivation)
Society (status planning)	(1) Selection (decision procedures) (a) Identification of problem (b) Allocation of norms	(3) Implementation (educational spread) (a) Correction procedures (b) Evaluation
Language (corpus planning)	(2) Codification (standardization procedures) (a) Graphization (b) Grammatication (c) Lexication	(4) Elaboration (functional development) (a) Terminological modernization (b) Stylistic development

the descriptive norm: this starts from usage and describes what is 'normal', 'regular', or 'frequent' in language usage without making a value judgement about it. The prescriptive norm, conversely, tends to be more subjective: it prescribes what should be said – or more usually written – based on value judgements. The prescriptive norm may begin with the observation of usage, but then a notion of what is correct or incorrect is overlaid.

As a number of linguists have noted, the opposition between descriptivism and prescriptivism is not clear-cut, and it is therefore better to think of it as a continuum (Joseph 1987: 18). A schematic representation of the relationship between codification and the descriptive and prescriptive norms is presented in Figure 2.1.

This then gives us our 'double standards': norms of usage (the descriptive norm) and the prescriptive norm, which we will explore in more detail in the case studies below (Section 2).

It may also be worth mentioning here that, like Walsh (2016), I distinguish prescription and purism – two concepts which are also sometimes not clearly differentiated (see, for example, Brousseau 2011: 4; Curzan 2014: 24). While it is true that purism shares common elements with prescriptivism, notably that, in general, only one form of the language is the correct one, purism introduces two additional ideas: (i) this language is pure and, therefore, to change it equals contamination, corruption, or decline of some sort; (ii) the language must be protected from

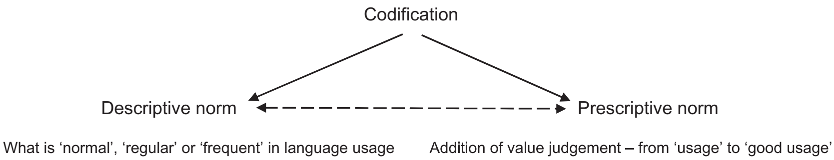


Figure 2.1 Codification, the descriptive and prescriptive norms.

this contamination and preserved in its current state (or, alternatively, if the language has already begun to be corrupted, the corrupted part must be removed) (see Figure 2.2). The addition of this additional layer of value judgement is, naturally, optional.

In his discussion of purism, George Thomas introduces a measure of intensity for purism, distinguishing between mild, moderate and extreme purism on the basis of the extent to which non-puristic factors are taken into consideration, and assigning a numerical score to the different types (Thomas 1991: 170). I have argued that, in evaluating how prescriptive a particular text is, it might be useful to adopt the same kind of scale of intensity, or at least to consider the relative weight of descriptive and prescriptive elements (Ayres-Bennett 2020). In the 2020 article I demonstrate how you might use such a scale of intensity to evaluate how authors and texts are more or less prescriptive.

Furthermore, I have argued that it is important to distinguish and refine consideration of what aspect of metalinguistic texts or approaches we are considering when we are evaluating them as to whether they are prescriptive or not (Ayres-Bennett 2020, 2021). As a minimum, we should differentiate in my view whether a text or author is (i) prescriptive in intention/purpose, (ii) prescriptive in expression, or (iii) prescriptive in

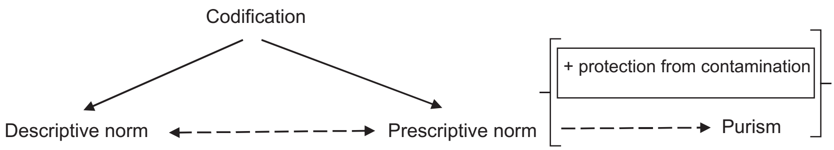


Figure 2.2 Codification, prescription and purism.

effect.[2] If we separate out and analyse in detail these factors, many of the so-called icons of prescriptivism are more complex than they would first appear, as my case studies will show.

3. Case studies

3.1 Methodology

Now that I have established my use of terminology, I want to present a number of case studies to explore the extent to which prescriptive norms do – or do not – conform with norms of usage. My examples are all taken from the work of Claude Favre de Vaugelas since he has been constantly characterised in the literature on the French grammatical tradition as being highly prescriptive and, indeed, sometimes purist (Cerquiglini 2008: 7, 10; Paveau & Rosier 2008: 17; Coski 2011: 77).[3] During his lifetime he published a single volume of remarks on the French language (Vaugelas 1647) which was an immense publishing success to judge by the number of editions published throughout the seventeenth century (Ayres-Bennett 2018a: 37–42). In 1690 Louis-Augustin Alemand announced that he had discovered an unpublished volume of observations by Vaugelas (Vaugelas 1690). In fact, these were the observations which Vaugelas had, for various reasons, chosen not to publish in 1647 (Ayres-Bennett 2018a: 26–8).

2 This embraces questions around the reception of the text. As Joseph (1987: 18) points out, 'unfortunately, even if one takes great pains to write a descriptive grammar, readers may impose a prescriptive interpretation on it'.

3 I and others have also conducted similar studies for Lowth for the English tradition and for Schottelius for the German tradition, since they have likewise been characterised as icons of prescriptivism (McLelland 2011; Tieken-Boon van Ostade 2011; Ayres-Bennett 2016).

I want to concentrate my analysis on the second type of prescription I identified, that is, the use of metalanguage and the attitude – whether descriptive or prescriptive – that this represents. Many critics have based their judgement about whether a text is prescriptive or not mainly, if not solely, on the type of metalanguage employed. This is perhaps therefore where we can best focus on the question of double standards and consider the extent to which texts that are prescriptive in expression follow norms of usage or not. If the text employs prescriptive language we need to determine whether the judgement captures 'normal usage', albeit couched in prescriptive language, or whether it is prescribing and thereby valorising a minor or even a non-attested form. In the former case, we need to establish whether what is offered is a description of the dominant usage; or guidance about changing usage (for instance, which form or construction is taking over as the dominant usage); or information about different social or stylistic values associated with a particular usage. Conversely, it is important to ascertain whether descriptive metalanguage indeed reflects a descriptive attitude and therefore provides norms of usage or whether it might hide what is actually a prescriptive attitude and behaviour.

In order to assess the extent to which norms of usage are being followed – at least in written texts – I selected a number of apparently prescriptive observations and then traced usage of that linguistic feature during the seventeenth century. For this I used the Frantext database which is a multi-genre corpus of different texts; for the seventeenth century it currently holds 636 texts, comprised of over 21.7 million words.[4] My case studies present examples of pronunciation, verb morphology and syntax, and add to the studies of *print/prindrent/prinrent*, *print/prit*, *commencer à/commencer de* and clitic climbing discussed in previous publications (Ayres-Bennett 2004: 208–19, 2016, 113–14).

4 <www.frantext.fr> The studies were conducted in February 2023, when the database for the seventeenth century comprised 21,761,026 words.

3.2 Case study 1: [o] vs [u]

My first case study deals with the variation in pronunciation between [o] and [u] in tonic or pretonic position. Usage was very divided between [o] and [u] in the sixteenth century in a large number of words (Thurot 1881–1883: I, 240), including in common words like *chose/chouse* ('thing'), *arroser/arouser* ('to water'), *portrait/pourtrait* ('portrait'), and *fossé/foussé* ('ditch'). The sixteenth-century grammarian Meigret (1980 [1550]: 6) and Bèze (1584: 17), the author of a work on French pronunciation, already comment, on the one hand, on the use of [u] for [o], indicated in the spelling by *ou* instead of *u*, in words such as *trop* (pronounced as *troupe*), *nostre* (pronounced as *noustre*) and *vostre* (pronounced as *voustre*); and on the other hand on [o] for [u], notated in the spelling as *u* for *ou*, in words such as *course* (pronounced as *corse*), *trousse* (pronounced as *trosse*) and *coup* (pronounced as *cop*). In the early seventeenth century, the grammarians Maupas (1618: 9v-10r) and Oudin (1632: 9) condemn the pronunciation with [u] in *chose* (*chouse*) and *composa* (*compousa*), Maupas commenting that the variant *chouse* for *chose* is 'une niaise et vicieuse prononciation' ('a silly and erroneous pronunciation'). This is what Vaugelas has to say on the question:[5]

> Portrait, pourtraict
>
> Il faut dire *portrait*, et non pas *pourtrait* avec un *u*, comme la plus part ont accoustumé de le prononcer, et de l'escrire. Il est vray qu'on a fort long-temps prononcé en France l'*o* simple comme s'il y eust eu un *u* apres, et que c'eust esté la diphthongue *ou*, comme *chouse*, pour *chose*, *foussé*, pour *fossé*, *arrouser*, pour *arroser*, et ainsi plusieurs autres. Mais depuis dix ou douze ans, ceux qui parlent bien disent *arroser, fossé, chose*, sans *u*, et ces deux particulierement, *foussé,* et *chouse*, sont devenus insupportables aux oreilles delicates. (Vaugelas 1647: 340)
>
> 'You must say *portrait*, and not *pourtrait* with a *u*, as most people used to pronounce and write it. It is true that for a very long time simple *o* was pronounced as if there were a following *u*, and as if it were the diphthong *ou*, as in *chouse* for *chose*,

5 I have kept the original spelling except that the use of *u/v* and *i/j* has been modernised.

> *foussé* for *fossé*, *arrouser* for *arroser*, and many others similarly. But in the last ten or twelve years, those who speak well say *arroser*, *fossé*, *chose*, without a *u*, and in particular *foussé* and *chouse* have become intolerable to sensitive ears'.[6]

The metalanguage used by Vaugelas in this example is clearly prescriptive: 'you must say … and not … as most people used to pronounce and write it'; 'those who speak well say …'; 'have become intolerable to sensitive ears'. The results of looking at usage of this variant in Frantext are presented in Table 2.2; Figure 2.3 shows the percentage of occurrences of each of the forms over the century. Following the consistent usage in the metalinguistic texts, we associate the spelling *pourtrait* with the pronunciation

Table 2.2 Variation between *pourtrait* and *portrait* 1600–1699.

	1600–1609	1610–1619	1620–1629	1630–1639	1640–1649	1650–1659	1660–1669	1670–1679	1680–1689	1690–1699
pourtrai(c)t	20	37	39	38	13	0	3	1	0	0
portrai(c)t	18	8	113	149	97	149	110	175	112	193

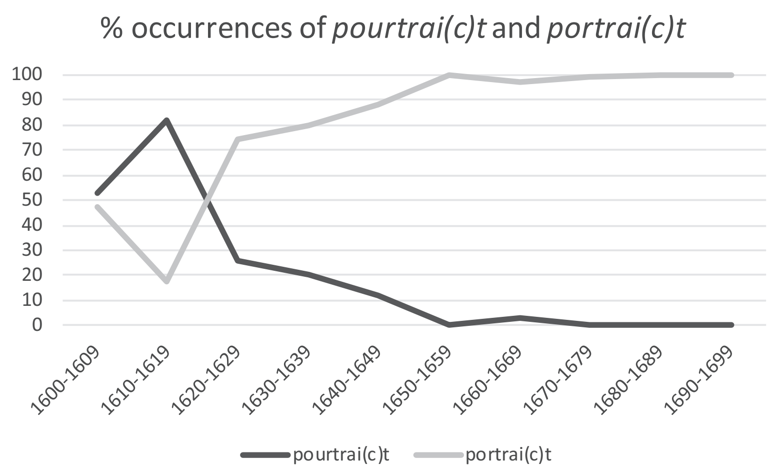


Figure 2.3 Variation between *pourtrait* and *portrait* 1600–1699.

6 All translations are my own.

with [u], noting that Vaugelas considers that this variation affects both the spoken and written language.[7]

The data show clearly that while the spelling with *ou* dominated in the first two decades of the century, by the time of the publication of the *Remarques* in the 1640s, the variant spelled with *o* had become dominant. Usage then seems to reflect Vaugelas's judgement regarding the emerging prevailing usage, despite the prescriptive language of the observation. Given that written usage is likely to lag somewhat behind pronunciation, the accuracy of the chronology for the change offered in the remark is also striking, particularly since Vaugelas did not have the benefit, as we do, of analysing twenty-one million words of texts.

3.3 Case study 2: Favorit *vs* favori

My second case study comes from Vaugelas's *Nouvelles Remarques* and concerns the variation between *favorit* and *favori* for the masculine form of the noun 'favourite'. Vaugelas's remark (1690: 98) is again expressed in prescriptive terms:

> *Favori* ou *favorit*.
>
> Il faut dire *favori*, non *favorit*, comme écrivent plusieurs bons Ecrivains: et toutefois on dit *favorite*, et non *favorie*. Ce mot, à mon avis, vient de l'Italien *favorito*. C'est ce qui fait que l'on retient le *t* au féminin: mais au masculin, parce qu'il finit le mot, il seroit trop rude: et en effet on l'a dit au commencement, et puis peu-à-peu on l'a ôté pour adoucir et polir le mot. Car comme une piéce d'argent ou quelque autre monnoye vient à se polir par l'usage à force de passer par les mains des personnes; aussi les mots s'adoucissent à force de passer par la langue et la bouche de ceux qui parlent.
>
> 'You must say *favori* and not *favorit*, as many good authors write: and nevertheless you say *favorite* and not *favorie* [in the feminine singular]. In my opinion, this word comes from the Italian *favorito*. That is why the *t* is retained in the feminine: but in the masculine, because it comes at the end of the word, it would

7 Théodore de Bèze (1584: 17) also uses the formulation 'scribunt et legunt' ('write and read').

> be too harsh; and in fact it was initially pronounced and then little by little the *t* was removed to soften and polish the word. For just as a silver coin or some other money gradually gets polished through usage by being passed through the hands of people, so words come to sound sweeter by being passed through the tongue and the mouth of those who speak'.

This is an interesting remark, since the wording seems to suggest that he is not following his own definition of where to find good usage, namely the *la plus saine partie de la cour* ('the "healthiest" part of the court') for spoken language, and *la plus saine partie des auteurs du temps* ('the "healthiest" part of contemporary writers') for written usage (Vaugelas 1647: *Préface* II.3). When we look at usage as represented in Frantext, there are no examples of *favorit*, whatever the genre, after the 1620s, and only *favori* is found. Since the number of occurrences of this example are low, Table 2.3 and Figure 2.4 present the raw figures rather than the percentages of each variant.

Table 2.3 Variation between *favorit* and *favori* 1600–1699.

	1600–1609	1610–1619	1620–1629	1630–1639	1640–1649	1650–1659	1660–1669	1670–1679	1680–1689	1690–1699
favorit	7	0	1	0	0	0	0	0	0	0
favori	1	0	0	7	4	55	19	57	35	41

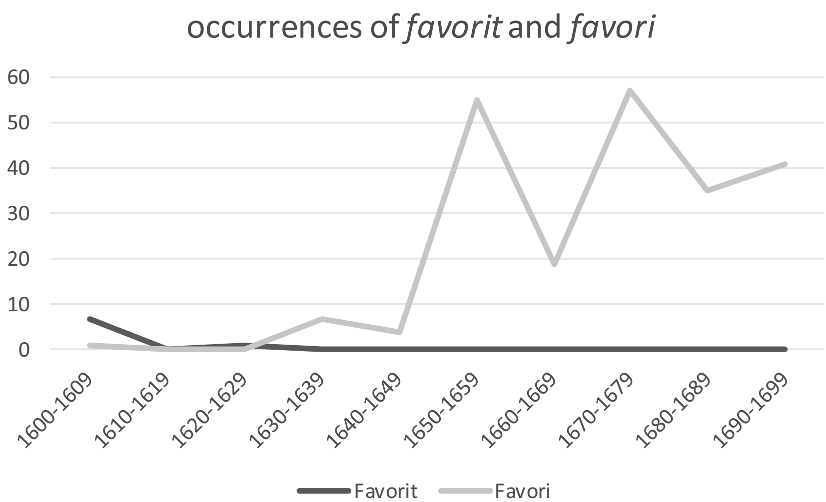


Figure 2.4 Variation between *favorit* and *favori* 1600–1699.

These data may explain why Vaugelas chose not to publish this remark in 1647. By the time of publication, the question no longer fell into what he terms *usage douteux* ('doubtful usage'), and the remark had therefore become redundant. Whatever the case, the observation seems to capture, rather than prescribe, what was 'normal' at the time.

3.4 *Case study 3: Verbal syntax*

In a series of observations on the question of verbal syntax, Vaugelas discusses whether various French verbs take a direct object (DO) or should be followed by the preposition *à*: this gives oppositions such as *survivre/survivre à* ('to survive'), *éviter/éviter à* ('to avoid'), *prier/prier à* ('to pray') and *ressembler/ressembler à* ('to resemble'). Not all of them restrict usage to one form: concerning *survivre*, for instance, Vaugelas (1647: 162) accepts both forms, and leaves it up to the reader to judge which form sounds better in which context. In the case of *ressembler/ressembler à* ('to resemble'), however, Vaugelas (1647: *480–*481) judges that it requires what he calls the dative (i.e. *ressembler à*) and that the other construction is old:

> *Ressembler.*
>
> On demande si *ressembler*, regit aussi bien l'accusatif, que le datif; car personne ne doute qu'il ne regisse le datif. M. de Malherbe a escrit en un certain lieu, *gardons nous de le ressembler*, et en un autre, *avecque ce langage et autres qui le ressemblent*, et M. Bertaut lui a fait aussi regir l'accusatif en cette fameuse stance [quotes the stance]. Il y a beaucoup d'autres Autheurs qui luy donnent l'accusatif, mais ce sont les vieux, et non pas les modernes; Ce qui fait voir que c'estoit la vieille façon de parler, que de luy faire regir l'accusatif, et qu'aujourd'huy il demande tousjours le datif. Il est vray qu'en faveur de la poësie j'ay oüy dire à plusieurs personnes tres-sçavantes en nostre langue, qu'en vers ils le souffriroient à l'accusatif, aussi bien qu'au datif, mais qu'en prose ils le condamneroient absolument.
>
> 'People ask whether *ressembler* can take the accusative as well as the dative; for nobody doubts that it can take the dative. M. de Malherbe [one of his authorities] wrote in a certain place, *we must be careful not to resemble him* [DO], and in another, *with this language and others which resemble it* [DO], and M. Bertaut

made it govern the accusative in this famous stance [...]. There are many authors who make it take the accusative, but they are old and not modern ones. This shows that it was the old way of speaking to make it govern the accusative and that today it always takes the dative. It is true that regarding poetry I have heard many people who are very knowledgeable about our language say that in poetry they would tolerate it in the accusative as well as the dative, but that in prose they would condemn this completely'.

Of the 168 occurrences of *ressembler* in the infinitive in Frantext for the period 1600–1699, some have to be discarded, for example where there is

Table 2.4 Variation between *ressembler à* and *ressembler* + DO 1600–1699.

	1600–1609	1610–1619	1620–1629	1630–1639	1640–1649	1650–1659	1660–1669	1670–1679	1680–1689	1690–1699
ressembler à	8	8	17	8	2	12	19	20	10	13
ressembler + DO	2	4	4	7	0	0	2	0	0	0

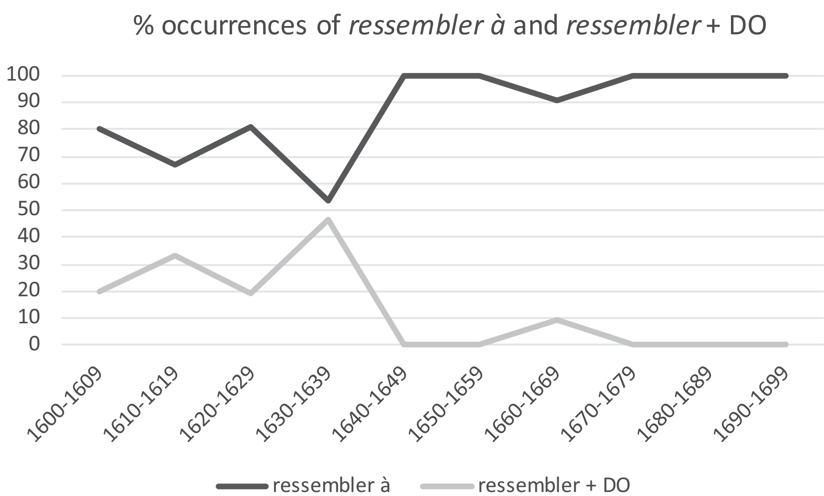


Figure 2.5 Variation between *ressembler à* and *ressembler + DO* 1600–1699.

a pronoun which is ambiguous[8] as to case or where the verb is used reflexively. Of the remaining 136 occurrences, there are 117 of *ressembler à* and nineteen of *ressembler* + DO. The findings are set out in Table 2.4 and Figure 2.5. From the 1640s on, *ressembler à* appears virtually categorical, the only exceptions being two examples in the 1660s. Vaugelas, publishing in the 1640s appears to have captured the majority usage.

In the case of *éviter* (Vaugelas 1647: 248–9), there is a clear value judgement ('very bad', 'fault') expressed:

> *Eviter.*
>
> Plusieurs luy font regir le datif, & disent *eviter aux inconveniens,* mais tres-mal, et ce qui a donné lieu à cette faute, c'est que l'on dit ordinairement, *pour obvier aux inconveniens*; mais *eviter*, regit l'accusatif, et *obvier* le datif.
>
> 'Many people make it govern the dative and say *eviter aux inconveniens* ['to avoid drawbacks'], but this is very bad, and what has generated this fault is the fact that ordinarily you say *pour obvier aux inconveniens* ['to obviate drawbacks'], but *eviter* governs the accusative and *obvier* the dative'.

Searching just the infinitival forms of *éviter*, we find that there are 1,753 occurrences of the verb in Frantext for the period 1600–1699. Of these only twenty-seven, or approximately 1.5%, are examples of *éviter à/au/aux* + nominal direct object; twenty of these occur pre-1630, and there are only two examples for the period 1630–1649. Here again, the majority usage seems to have been captured by the *remarqueur*. The case of *prier* in the example *prier Dieu/prier à Dieu* ('to pray to God') is even more extreme, since the form judged 'old' by Vaugelas (1647: 479) *prier à Dieu* is not attested in Frantext throughout the century, whereas there are 186 occurrences of *prier Dieu* in the infinitive.

8 Where there is a first- or second-person pronominal object such as *je vous évite* ('I avoid you'), the form for the direct and indirect objects is identical, so it is impossible to judge the case of the pronoun.

3.5 *Case study 4: Third-person singular present subjunctive of* avoir

My next example is a case of verb morphology: the variation between *il ayt/il ait* and *il aye/il aie* for the third-person singular present subjunctive of the verb *avoir* ('to have'). In his grammar, Maupas (1618: 92v) offers both forms without comment, but Oudin (1632: 145), while including both forms in his verb tables, nevertheless observes that *ait* is better than *aye*. In his *Remarques*, Vaugelas (1647: 90) is more dogmatic in the way he describes usage, and this might be thought to be a covertly prescriptive comment dressed up as a descriptive observation:

> *Aye*, ou *ayt*.
>
> Le verbe *avoir*, en l'optatif et au subjonctif, ne dit jamais, *aye*, en la troisiesme personne, mais tousjours *ayt*, soit en vers, ou en prose. Ce n'est pas qu'autrefois on n'ayt escrit, *aye*, mais on ne l'escrit plus qu'en la premiere personne: comme, *je prie Dieu que j'aye bon succés de, etc.* et *qu'il ayt bon succés, afin que j'aye, et afin qu'il ayt.*
>
> 'The verb *avoir*, in the optative[9] or subjunctive is never *aye* in the third person, but always *ayt*, whether in verse or in prose. It is not that *aye* was not used in the past, but it is no longer used except in the first person, as in I pray to God that I have success in, etc. (*j'aye*) and that he has success in (*il ayt*), so that I have (*j'aye*), and so that he has (*il ayt*)'.

If we take 'never' literally, then the observation is not entirely descriptive as Table 2.5 and Figure 2.6 show, since in the 1640s there are still seven attestations of the older form. However, all of these appear in correspondence, and six of them were published in 1649, two years after the publication of the observation. After something of a peak in the 1630s, when use of the old form reached 18.2% of occurrences, in the 1640s there was a sharp decline in frequency of the older form to 9.1%. This looks like the end stages of the change, and therefore Vaugelas's comment might be felt to be justified.

9 Vaugelas still refers to the Latin category of optative to refer to the semantic function of the subjunctive to express a wish.

Table 2.5 Variation between *il aye* and *il ayt* 1600–1699.

	1600–1609	1610–1619	1620–1629	1630–1639	1640–1649	1650–1659	1660–1669	1670–1679	1680–1689	1690–1699
il aye/ il aie	16	4	9	24	7	1	5	1	5	2
il ayt/ il ait	33	44	110	108	70	123	116	91	157	125

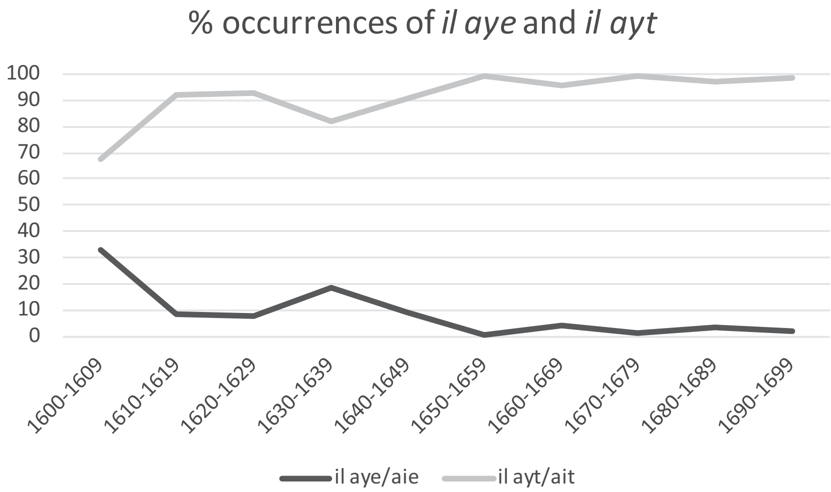


Figure 2.6 Variation between *il aye* and *il ayt* 1600–1699.

3.6 Summary

In all of the examples discussed above, as in the ones analysed in previous publications (*print/prindrent/prinrent*; *print/prit*; *commencer à/commencer de*; clitic climbing), Vaugelas's observations record contemporary usage with a good degree of accuracy. Moreover, the *remarqueur* is clearly reflecting 'normal usage' rather than driving it. The observations show how it is misleading simply to judge an author or a text by looking at the metalanguage. In all of the observations we have analysed the norms of usage and the prescriptive norms seem broadly to overlap rather than being discrete double tracks.

4. Tracking change by genre or text type

There has been a great deal of discussion as to which text types might best provide evidence for language change (Ayres-Bennett 2014), in particular when trying to track changes 'from below' (Elspaß 2020). In my final examples, I present two cases where I have undertaken a finer analysis of certain corpus attestations according to the genre in which they are found. In so doing, I acknowledge the limitations of dealing with a multi-genre corpus of texts, many of which are literary. We are not dealing here with the invaluable source of informal personal letters ideally written ideally by those with low literacy skills, which have been described by Nevalainen & Raumolin-Brunberg (2012: 32) as the 'next best thing' to authentic spoken language, and which have transformed the work, for instance, on the history of Dutch (Rutten & van der Wal 2014). Unfortunately, for seventeenth-century metropolitan French, these have been to date extremely scarce, although the transcription of French letters from the Prize Papers as part of the MACINTOSH project, for instance, promises to open up exciting new sources for research on this period.[10] Moreover, I am not denying that many, or indeed the majority of, changes arise 'from below'. My point is rather that, as a result of what we see from the analysis of large-scale corpora, we need to be careful about making generalisations about the relative merits of different text types for research in historical sociolinguistics and how well they may, or may not, show early signs of language change. My previous analyses of clitic climbing (Ayres-Bennett 2004: 208–19) and of

10 <https://www.prizepapers.de/>, accessed 14 October 2024. The MACINTOSH project (<https://lettresoutremer.huma-num.fr/>) undertakes to provide diplomatic editions of a selection of 1,000 letters for the period 1652–1760 by 2026. However, these are very much concentrated in the eighteenth century.

the variation in adverbial contexts between *après/en après/par après* ('after, afterwards') (Ayres-Bennett 2018b) seem to refute the simple hypothesis that, since the majority of changes occur in speech, we might expect them to be attested first in textual genres considered closest to speech, such as correspondence and direct speech in theatre and other genres, and to show up latest in more literary and 'artificial' genres such as poetry. In the case of clitic climbing I found that, while theatre was indeed more progressive than oratory or fables written in verse, Mme de Sévigné's letters, the majority of which were written to her daughter, were slower to show the predominance of the new construction than theatre. Moreover, within theatre it was the verse comedies and tragedies of Pierre Corneille which led the way in attestations of the new construction, rather than Molière's comedies that have often been deemed a good source of 'more spoken' usage or of the vernacular (Lodge 1991). The case of *après/en après/par après* (Ayres-Bennett 2018b) was somewhat different: here it did not make much sense to think of this as a change 'from above' or 'from below'. In fact, the genres where the forms going out of usage were most retained were a mixture of formal genres (treatises and essays) and those which are often considered more informal because of their personal nature (memoirs, journals, travelogues). It seems that these expressions were associated rather with genres where the sequence of events or the sequence of arguments is vital. Unsurprisingly then, the forms survive longest in their core genre of essay/treatise.

In what follows, I will briefly outline my methodology before introducing two new case studies to add to the existing evidence.

4.1 Methodology

For this study, each of the attested forms under consideration was classified according to its genre. Many of the texts in Frantext are already classified by genre, so it was decided to retain the categories employed there, and to use these to categorise those texts not already assigned a genre. Seven broad categories were employed, which in some cases merged separate poetry essays novels theatre sermons memoirs correspondence

Figure 2.7 Continuum of genres used in the analysis.

but related categories used in Frantext: poetry; essays/treatises [essays]; novels/novellas/tales (*contes*) [novels]; theatre; sermons; memoirs/travelogues [memoirs]; correspondence. The designation in square brackets is the short form which is used in the results for convenience. As Figure 2.7 shows, these genres might be broadly placed on a continuum ranging from more formal literary genres to more personal ego-documents.[11] If changes occur from below, then it might be predicted that they would be attested first in correspondence, memoirs and travelogues, and more slowly in more formal genres such as poetry and essays/treatises.

4.2 Case study 5: Intrigue

My first example considers in which genres a new word is first attested and how it subsequently spreads from there to different genres. There is a significant number of comments on new words in the *Remarques* (1647). In the case of the word *intrigue* ('intrigue, plot [of a story, play, etc.]'[12]), Vaugelas (1647: 126) makes the following observation:

11 On ego-documents, see, for instance van der Wal & Rutten (2013). The term is used to refer to writings in which the writer or 'I' is present in the text as the writing and describing subject. It thus comprises letters and autobiographical works such as memoirs, journals and travelogues.

12 There is not space in this article to explore in depth the semantics of this term or the contexts in which it occurs, but it is worth noting that in the period 1620–1649 both the sense of 'intrigue', often associated with the court, and the sense of 'plot' are found, the first being the more common usage as attested in Frantext.

Intrigue

La pluspart font ce mot feminin, je dis *la pluspart*, parce qu'il y en a qui le font de l'autre genre; il faut dire *intrigue*, avec un *g*, et non pas *intrique*, avec un *q*, comme force gens le disent et l'escrivent. C'est un nouveau mot pris de l'Italien, qui neantmoins est fort bon, et fort en usage.

'Most people make this word feminine, I say most people because there are some who give it the other gender; you must say *intrigue* with a *g* and not *intrique* with a *q*, as many people say and write it. It is a new word taken from Italian, which nevertheless is very good and very much in use'.

In Frantext for the period 1600–1699 there are in total 354 occurrences of the lemma *intrigue*, and twenty-three of *intrique* dating from 1615 to 1682, of which twenty are by Pierre Corneille, suggesting it was perhaps something of an idiosyncrasy in his usage. The comment recommending the form *intrigue* rather than *intrique* therefore seems to follow norms of usage. The same is true of the observation that most people make the word feminine: while there are many cases where the gender of the word is not marked because it begins with a vowel, of the 354 occurrences only four are clear examples of the masculine (1624–1646). Figure 2.8 illustrates the number of occurrences of the noun over the century. While there are relatively few attestations in the first half of the century, the first

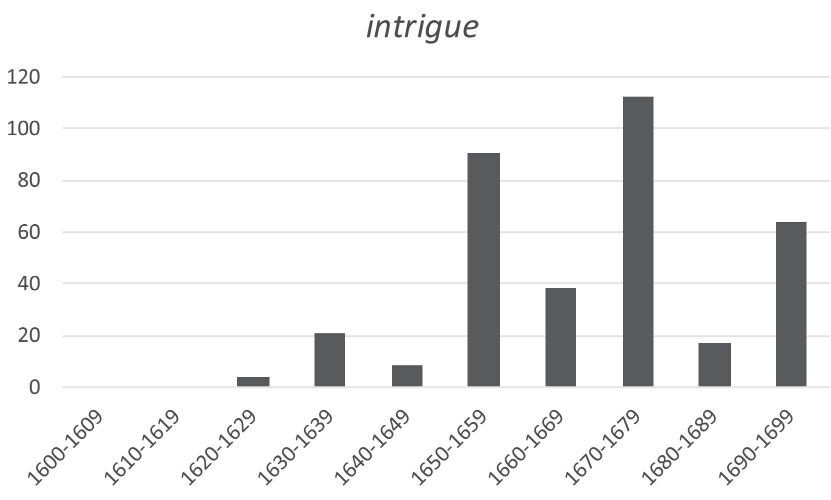


Figure 2.8 Occurrences of the lemma *intrigue* 1600–1699.

occurring in 1623, by the second half the term seems to have become well established in usage.

Turning now to the genres in which it is found, the data are provided in Table 2.6 and Figure 2.9. Looking first at the total figures, we can see

Table 2.6 Occurrences of *intrigue* by genre 1600–1699.

	Poetry	Essays	Novels	Theatre	Sermons	Memoirs	Correspondence
1600–1609	0	0	0	0	0	0	0
1610–1619	0	0	0	0	0	0	0
1620–1629	0	2	0	0	0	0	2
1630–1639	0	7	0	12	0	0	2
1640–1649	1	0	0	5	0	0	2
1650–1659	5	71	12	1	0	0	1
1660–1669	1	6	7	13	11	0	0
1670–1679	2	29	41	5	1	30	4
1680–1689	0	4	1	2	2	0	8
1690–1699	0	19	17	5	3	14	6
TOTAL	9	138	78	43	17	44	25

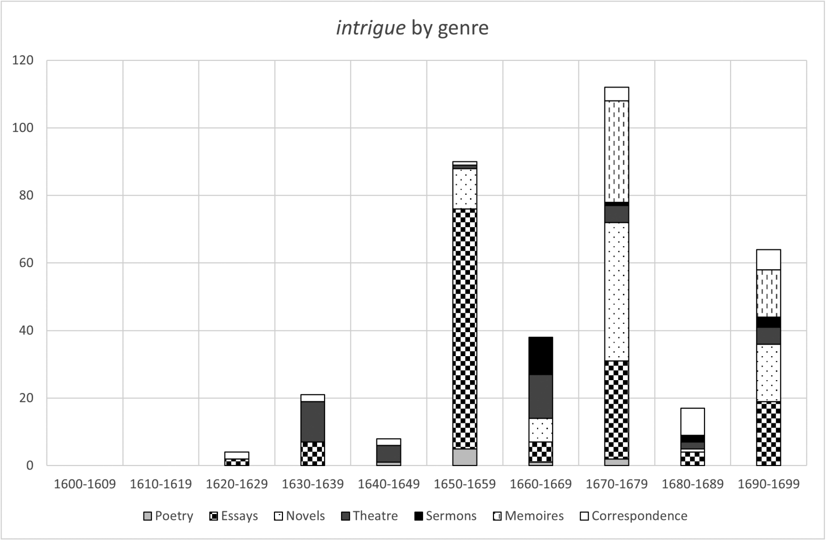


Figure 2.9 Occurrences of *intrigue* by genre 1600–1699.

that all genres are represented. The form is favoured by the more formal, less personal genres (poetry, essays, novels), but there are still more than a third of the examples (36.4%) in what we might think of as the more personal genres. Moreover, when we look at where the form is first attested and how it spreads through the different genres, it is difficult to discern any clear patterns. In the 1620s there are two examples in essays and two in correspondence, which are at different ends of our continuum of genres. In the 1630s, these two genres are still represented, but there are more examples in theatre, a fact which can in part at least be explained by one of the meanings of the term being the plot of a play. In the 1640s, theatre still remains the most common genre, but the term is also attested in poetry and correspondence. It is therefore difficult to think of this being clearly a change 'from above' or 'from below' or to argue – within the limited parameters of the texts included in Frantext – that one text type is a better source of incoming usages than another. Moreover, even in a large database such as Frantext, results can be skewed by the use of an individual author as we have seen above: for the 1650s, sixty-six of the examples come from the abbé d'Aubignac's 1657 essay, *La Pratique du théâtre*, in which the author seeks to set out the profession of playwright.

4.3 *Case study 6:* ains

My final example is of the conjunction *ains* which has the sense of 'but' after a negative, equivalent to *sondern* in German, for instance. This time we are looking at an example of a form going out of usage. On the surface, it is once again not clear whether we are dealing with a descriptive or prescriptive comment (Vaugelas 1690: 284–5):

> *Ains.*
>
> *Ains* n'est plus en usage parmi les bons Auteurs: aussi ne le dit-on jamais à la Cour, si ce n'est en raillant avec cette queuë, *ains au contraire*. J'étois present quand M. de

> Malherbe en avertit M. Coëffeteau qui en usoit au commencement de ses Oeuvres: mais à la vie de Tibére, si je ne me trompe, ou environ, il commence à ne s'en plus servir. Je sçay combien l'usage en est nécessaire, et le besoin qu'on en a à tous propos, pour n'être pas obligé de répéter toujours *mais* dont il faut se servir si souvent. Je sçay aussi que *mais* n'exprime pas toûjours bien la signification d'*ains*, qui a toute autre force à dénoter les choses opposées, en quoy *mais* se trouve foible. Mais il n'y a reméde, l'usage l'a banni, on ne le dit jamais à la Cour, et la régle est générale et sans exception, *que ce qui ne se dit jamais en parlant, ne se dit jamais en écrivant*.
>
> '*Ains* is no longer used by good authors, and it is never used at court unless it is used in jest with this tail, *ains au contraire* [but on the contrary]. I was present when Malherbe informed Coëffeteau of this who used it at the beginning of his works: but from the life of Tiberius [of his *Roman History*], if I'm not mistaken, or there abouts, he begins to no longer use it. I am fully aware that its usage is necessary and that it is constantly needed so that one is not forced always to repeat *mais* [but] which has to be used so often. I know moreover that *mais* does not always have the same meaning as *ains*, which has a quite different force for denoting opposing things, and for which *mais* seems weak. But there is nothing to be done about it, usage has banished it, and it is never used at court, and the rule is general and without exception that *what is never used when speaking is never used in writing*'.

Table 2.7 and Figure 2.10 show clearly the demise in usage of *ains* over the century: of the 1,517 examples of the form in the seventeenth century, 1,503 (99.1%) occur in the first half of the century, and there are only twenty (1.3%) attestations in the 1640s, the decade when the *Remarques* were published.

For this study I have analysed the genres of where the occurrences occur for the period 1600–1649, since thereafter the number of attestations is too low to be informative. These results are presented in Table 2.8 and Figure 2.11. While the genre of essays clearly dominates overall, the second most common genre is correspondence, again at the other end of my continuum, followed by poetry and theatre. In the 1610s, when attestations have already reduced by more than two-thirds, essays and correspondence are still the genres most commonly represented. While it is true that in the 1630s correspondence has the most occurrences, it is notable that the word is also still featuring in poetry. In other words, it is once again difficult to see a clear pattern of loss being associated with either the more formal or the more personal genres.

Table 2.7 Occurrences of *ains* 1600–1699.

1600–1609	1610–1619	1620–1629	1630–1639	1640–1649	1650–1659	1660–1669	1670–1679	1680–1689	1690–1699	TOTAL
934	285	192	72	20	7	1	3	0	3	1,517

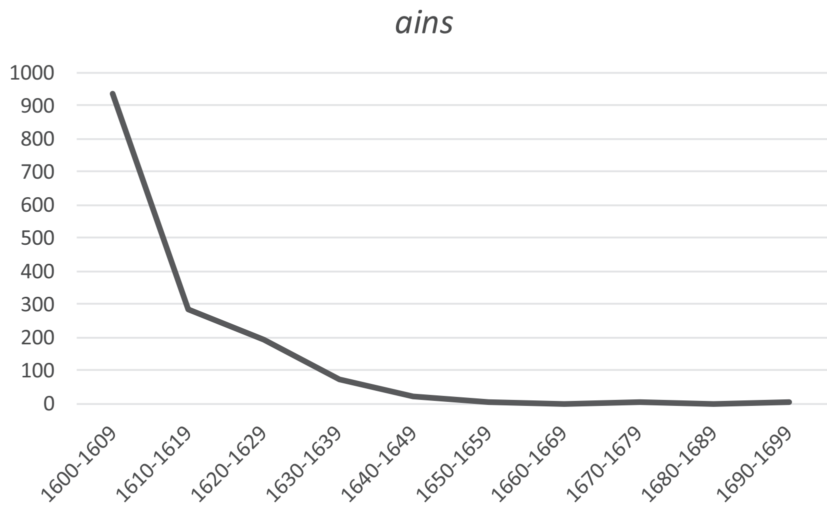


Figure 2.10 Occurrences of *ains* 1600–1699.

Table 2.8 Occurrences of *ains* by genre 1600–1649.

	Poetry	Essays	Novels	Theatre	Sermons	Memoirs	Correspondence
1600–1609	10	864	0	51	5	4	0
1610–1619	6	134	14	0	13	1	117
1620–1629	59	50	9	27	8	0	39
1630–1639	8	1	0	0	5	0	58
1640–1649	12	1	0	0	5	0	2
TOTAL	95	1,050	23	78	36	5	216

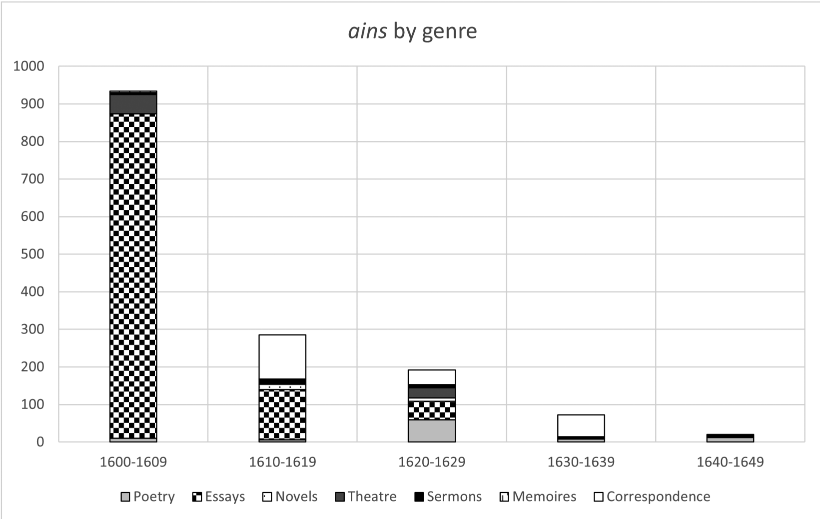


Figure 2.11 Occurrences of *ains* by genre 1600–1649 (not showing occurrences fewer than five).

5. Conclusion

I began by clarifying my usage of some of the terminology that is frequently used in discussing and modelling standardisation and my understanding of the double standards of norms of usage and prescriptive norms. Building on previous work, I argued that, if we are to analyse prescriptive authors and texts properly, we need to be clear which aspect of the work we are addressing. This involves differentiating, for instance, the author's purpose or intention (whether explicit, implicit or indeed covert) and the subsequent impact of the text (e.g. whether the prescriptive influence comes from the original text or a later version or perceptions of that text). Another facet which needs to be analysed carefully is the question of the metalanguage used in metalinguistic texts of all kinds, since this has often been a primary factor in evaluating whether a particular text is prescriptive or not. In so doing, we need to distinguish whether prescriptive metalanguage reflects a prescriptive stance or rather attempts to record a dominant usage or the rise of a form rapidly spreading through usage (thereby aligning prescriptive norms and norms of usage).

The new examples I presented in Section 2 provided further evidence that, when so-called prescriptive texts are examined in detail, the relationship between prescriptive norms and norms of usage is more complex than might have been expected. Vaugelas's reputation as an icon of prescriptivism has not been borne out by my analysis. While there are clearly cases where he is prescriptive, very often his observations do seem to reflect norms of usage, despite the metalanguage used. This suggests that we should consider the relationship between description and prescription not as a question of 'double standards', but as a continuum or a cline, allowing texts and their authors to be more or less prescriptive.

While research on Lowth for England and Schottelius for Germany has thrown up similar patterns, these are all examples of the standardisation of modern European languages in the early modern period. It would therefore be important to multiply this kind of study for other languages and other periods to ascertain whether what I and others have found is just a feature of highly standardised Western European languages or of the early modern period. In conducting these studies of what Joseph (1987: 13) has termed 'comparative standardology', care is needed in applying simple oppositions to different periods and different places where the linguistic, social and cultural contexts may be very different.

The final part of my study offered new examples of tracing the spread of change through a large multi-genre database. While we need to be careful in interpreting the figures, since Frantext is neither a perfectly stratified corpus nor does it contain, for instance, letters by those with low-levels of literacy (for metropolitan French these do not appear to survive in any great numbers until much later), it nevertheless contains over 21.7 million words of seventeenth-century usage, which is a substantial body of written evidence to analyse. In the examples presented in this article – as with previous studies – it was difficult to see any clear patterns of change spreading through different genres or to argue that one type of text was a more 'authentic' representation of norms of usage. In this study at least, it is unclear that the ego-documents or the more personal genres should be given absolute priority by historical sociolinguists.

Bibliography

Andry de Boisregard, Nicolas. 1689. *Réflexions sur l'usage présent de la langue françoise, ou Remarques nouvelles et critiques touchant la politesse du langage*. Paris: Laurent d'Houry.

Andry de Boisregard, Nicolas. 1693. *Suite des Réflexions critiques sur l'usage présent de la langue françoise*. Paris: Laurent d'Houry.

Armstrong, Nigel & Ian E. Mackenzie. 2013. *Standardization, Ideology and Linguistics*. Basingstoke: Palgrave Macmillan.

Ayres-Bennett, Wendy. 2004. *Sociolinguistic Variation in Seventeenth-Century France: Methodology and Case Studies*. Cambridge: Cambridge University Press.

Ayres-Bennett, Wendy. 2014. From *l'usage* to *le bon usage* and back: Norms and usage in seventeenth-century France. In Gijsbert Rutten, Rik Vosters & Wim Vandenbussche (eds), *Norms and Usage in Language History: A Sociolinguistic and Comparative Perspective*. Amsterdam, Philadelphia: Benjamins. 173–200.

Ayres-Bennett, Wendy. 2016. Codification and prescription in linguistic standardisation: Myths and models. In Francesc Feliu & Josep Maria Nadal (eds), *Constructing Languages: Norms, Myths and Emotions*. Amsterdam, Philadelphia: Benjamins. 99–129.

Ayres-Bennett, Wendy (ed.). 2018a. *Claude Favre de Vaugelas, Remarques sur la langue françoise*. Paris: Classiques Garnier.

Ayres-Bennett, Wendy. 2018b. Historical sociolinguistics and tracking language change: Sources, text types and genres. In Wendy Ayres-Bennett & Janice Carruthers (eds), *Manual of Romance Sociolinguistics*. Berlin, Boston: de Gruyter. 253–79.

Ayres-Bennett, Wendy. 2020. From Haugen's codification to Thomas's purism: Assessing the role of description and prescription, prescriptivism and purism in linguistic standardization. *Language Policy* 19(2), 183–213.

Ayres-Bennett, Wendy. 2021. Modelling language standardization. In Wendy Ayres-Bennett & John Bellamy (eds), *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 27–64.

Bèze, Théodore de. 1584. *De Francicae linguae recta pronuntiatione*. Geneva: E. Vignon.

Bouhours, Dominique. 1674. *Doutes sur la langue françoise proposez à Messieurs de l'Académie françoise par un gentilhomme de province*. Paris: Sebastien Mabre-Cramoisy.

Bouhours, Dominique. 1675. *Remarques nouvelles sur la langue françoise*. Paris: Sebastien Mabre-Cramoisy.

Bouhours, Dominique. 1692. *Suite des Remarques nouvelles sur la langue françoise*. Paris: Georges & Louis Josse.

Brousseau, Anne-Marie. 2011. Identités linguistiques, langues identitaires: Synthèse. *Arborescences: Revue d'études françaises*, 1, <http://id.erudit.org/iderudit/1001938ar> (accessed 8 June 2024).

Brunot, Ferdinand. 1909. *Histoire de la langue française des origines à 1900*. Vol. 3: *La Formation de la langue classique (1600–1660)*. Paris: A. Colin.

Cerquiglini, Bernard. 2008. *Merci professeur! Chroniques savoureuses sur la langue française*. Paris: Bayard.

Cerquiglini, Bernard. 2012. *Petites chroniques du français comme on l'aime!* Paris: Larousse.

Coski, Christopher. 2011. *From Barbarism to Universality: Language and Identity in Early Modern France*. Columbia: University of South Carolina Press.

Curzan, Anne. 2014. *Fixing English: Prescriptivism and Language History*. Cambridge: Cambridge University Press.

Elspaß, Stephan. 2020. Alternative sources of data for standardisation histories. *Language Policy* 19(2), 281–99.

Elspaß, Stephan. 2021. Language standardization in a view 'from below'. In Wendy Ayres-Bennett & John Bellamy (eds), *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 93–114.

Haugen, Einar. 1966. Dialect, language, nation. *American Anthropologist* 68(4), 922–35.

Haugen, Einar. 1987. *Blessings of Babel: Bilingualism and Language Planning. Problems and Pleasures*. Berlin, New York, Amsterdam: Mouton de Gruyter.

Joseph, John E. 1987. *Eloquence and Power: The Rise of Language Standards and Standard Languages*. London: Frances Pinter.

Kristiansen, Tore & Nikolas Coupland (eds). 2011. *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus Press.

Le Page, Robert B. 1988. Some premises concerning the standardization of languages, with special reference to Caribbean Creole English. *International Journal of the Sociology of Language* 71, 25–36.

Le Page, Robert B. & Andrée Tabouret-Keller. 1985. *Acts of Identity: Creole-Based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.

Lodge, R. Anthony. 1991. Molière's peasants and the norms of spoken French. *Neuphilologische Mitteilungen* 42, 485–99.

Lodge, R. Anthony. 1993. *French: From Dialect to Standard*. London, New York: Routledge.

Maupas, Charles. 1618 [1607]. *Grammaire et syntaxe françoise*. Orléans: Olivier Boynard & Jean Nyon.

McLelland, Nicola. 2011. *J. G. Schottelius's* Ausführliche Arbeit von der Teutschen HaubtSprache *(1663) and its Place in Early Modern European Vernacular Language Study*. Chichester: Wiley-Blackwell.

Meigret, Louis. 1980 [1550]. *Le Traité de la grammaire française*, ed. Franz Josef Hausmann. Tübingen: Narr.

Milroy, James & Lesley Milroy. 1991. *Authority in Language: Investigating Language Prescription and Standardisation*, 2nd edn. London, New York: Routledge.

Nevalainen, Terttu & Helena Raumolin-Brunberg. 2012. Historical sociolinguistics: Origins, motivations, and paradigms. In Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre (eds), *The Handbook of Historical Sociolinguistics*. Chichester, Malden: Wiley-Blackwell. 22–40.

Oudin, Antoine. 1640 [1632]. *Grammaire françoise rapportée au langage du temps*. Paris: A. de Sommaville.

Paveau, Marie-Anne & Laurence Rosier. 2008. *La Langue française: Passions et polémiques*. Paris: Vuibert.

Rutten, Gijsbert & Marijke J. van der Wal. 2014. *Letters as Loot: A Sociolinguistic Approach to Seventeenth- and Eighteenth-Century Dutch*. Amsterdam, Philadelphia: Benjamins.

Rutten, Gijsbert & Rik Vosters. 2021. Language standardization from above. In Wendy Ayres-Bennett & John Bellamy (eds), *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 65–92.

Sampson, Rodney (ed.). 1993. *Authority and the French Language*. Münster: Nodus.

Thomas, George. 1991. *Linguistic Purism*. London, New York: Longman.

Thurot, Charles. 1881–1883. *De la prononciation française depuis le commencement du XVIe siècle, d'après les témoignages des grammairiens*, 2 vols. Paris: Imprimerie nationale.

Tieken-Boon van Ostade, Ingrid. 2011. *The Bishop's Grammar: Robert Lowth and the Rise of Prescriptivism in England*. Oxford: Oxford University Press.

Vaugelas, Claude Favre de. 1647. *Remarques sur la langue françoise utiles à ceux qui veulent bien parler et bien escrire*. Paris: Veuve Camusat & Pierre le Petit.

Vaugelas, Claude Favre de. 1690. *Nouvelles remarques de M. de Vaugelas sur la langue françoise. Ouvrage posthume. Avec des observations de M.* ***** [Louis-Augustin Alemand]. Paris: G. Desprez.

Wal, Marijke J. van der & Gijsbert Rutten (eds). 2013. *Touching the Past: Studies in the Historical Sociolinguistics of Ego-Documents*. Amsterdam, Philadelphia: Benjamins.

Walsh, Olivia. 2016. *Linguistic Purism: Language Attitudes in France and Quebec*. Amsterdam, Philadelphia: Benjamins.

JULIA HÜBNER

3 The German dative marker *-e* in the early modern period: Double standards between codified norms and norms of usage

ABSTRACT
When studying historical grammar in general as well as particular phenomena such as the German dative *-e*, the task of finding appropriate and reliable sources presents itself as a major challenge. A promising, yet so far rather neglected type of research material is early modern textbooks for the teaching of foreign languages. This study is based on a corpus of fifty-six textbooks from the early modern period, usually containing a grammar and a dictionary as well as model dialogues. The grammar shows the explicit grammatical rules, and the model dialogues provide a deeper insight into the application of these rules. This will be illustrated by the concrete example of the German dative marker *-e*. Since there are significant discrepancies compared to previous studies analysing the development of the dative *-e*, this points out the value as well as the necessity of using different sources to obtain a picture as complete as possible.

1. Introduction[1]

The German language uses grammatical markers to indicate various grammatical relationships between words in a sentence. One such marker

1 I would like to express my gratitude to the organisers as well as the audience of the conference 'Double Standards: Codified norms and norms of usage in European languages (1600–2020)' in Salzburg for their valuable feedback and insightful questions during my presentation, which have contributed to the development of my research. Furthermore, I want to thank the anonymous reviewers of my paper for their helpful comments.

is the dative marker *-e*, which could be attached to certain masculine and neuter nouns in the dative case (cf. Example 1).

(1) im Wald $_{\text{M,Sg,Dat}}$ – im Wald-e $_{\text{M,Sg,Dat}}$ ('in the forest')

In the early modern period, the suffix *-e* was a possible, though not obligatory, marker to indicate the dative. Today, the dative *-e* is barely used and sounds quite archaic. Only in the context of formulaic expressions like *am Fuß-e des Berges* ('at the foot $_{\text{M,Sg,Dat}}$ of the mountain') is the dative *-e* is still regularly used (cf. Elspaß 2012; Radden 2021).

In this paper, I argue that our knowledge of the use and the distribution of the German dative marker *-e* in the early modern period is still fragmentary. Although we already know a great deal about the Early New High German period, many studies have narrowly focused on certain aspects of language use. For example, most early studies in the field of historical linguistics only focused on literary or formal texts written by scholars and elites. Thus, they are only able to describe language history 'from above', excluding most members of early modern society and their language use. Studies by Elspaß (2005, 2015, 2020) and Elspaß et al. (2007) show the importance of including alternative sources and establishing a language history 'from below' to portray early modern language use as accurately as possible.

This exploratory study of the use of dative *-e* broadens our understanding of everyday language in the early modern period by drawing on multilingual textbooks from the seventeenth century.

These textbooks presented relatively authentic everyday language use, which was most notably conveyed in the form of model dialogues written by non-scholars. This study aims to show that we can gain new perspectives on the use of dative *-e* by considering these dialogues as a source, since they are modelled on authentic spoken language. The textbooks allow us to reveal underlying double standards in the use of the dative *-e*. Whereas the grammar conveys a normative standard, the model dialogues teach a standard of usage including variation.

In addition, it sheds light on the normative concept inherent in these textbooks. This raises the question of whether the growing ideology of normativity and uniformity may have impacted actual language use and how textbook authors handled this tension. With this in mind, I want to demonstrate the value and relevance of early modern textbooks as a source for studies in historical sociolinguistics.

The paper begins by giving a brief overview of the current state of research on the development of dative *-e*. It then reviews early modern textbooks as a source for historical linguistics including not only the characteristics, but also the challenges of dealing with this kind of material. This is followed by a short overview of the methodology and the analysis of the use of the dative *-e* in early modern textbooks. Finally, I will discuss the results and draw a conclusion.

2. What do we already know about dative *-e*?

There is a relatively large amount of research on the German dative marker *-e*. However, the majority of these studies take grammars or literary texts as a source to investigate the dative. Thus, the issue of the extent to which dative *-e* was used in everyday communication has been left largely unaddressed.

In this section, I provide a brief overview of the most important research findings. Data from several studies suggest that there are two main factors influencing the use of dative *-e*: space and time. Beginning in the Old High German period (750–1050) and continuing to New High German, the inflection system was subject to processes of change, which consisted particularly in the weakening or omission of endings in noun inflection and a change from case inflection to number inflection (cf. Wegera & Solms 2000: 1542; von Polenz 2013: 277). In the seventeenth and especially in the eighteenth century, this process resulted in a controversy over unstressed *-e* in final syllables. Proponents wanted to retain the unstressed *-e* in its function to distinguish, inter alia, case and to preserve the inflection system. Opponents argued for its deletion, justifying this decision with their orientation to the actual language use. Although these studies differ slightly in their details, their tendency is similar: they all agree on the fact that the use of dative *-e* has declined since Middle High German (1100–1350) and was only interrupted by a phase of restitution in some regions in the seventeenth and eighteenth centuries (cf. Habermann 1997; Hartweg & Wegera 2005; Takada 1998).

These studies have also pointed out the influence of the regional distribution on the use of the dative *-e*. Based on data from the Bonn Corpus of Early New High German (covering the period from 1350 to 1700), von Polenz (2013: 277) notes that in the period from 1450 to 1600, dative *-e* is missing in approximately 50% of the Upper Saxon texts. In West Central German, dative *-e* had already largely disappeared by this time, and in Upper German it had disappeared almost completely (cf. Habermann 1997: 438–39; Hartweg & Wegera 2005: 154). Wegera & Solms (2000: 1543) also note that the apocope of *-e* remains predominant in most regions, but that since the sixteenth century, dative *-e* slightly increases in its use again, especially in East Central German – i.e. Thuringian and Upper Saxon – texts.

Until the seventeenth century, dative *-e* is still apocopated in almost all regions except Thuringian and Upper Saxon (cf. von Polenz 2013: 277). Studies based on the Bonn Corpus of Early New High German show that the *-e* is successively restored to almost 40% in Thuringian and well over 90% in Upper Saxon texts from 1650 to 1700. Around 1700, dative *-e* is described as almost completely restored in Upper Saxon texts, remaining an exception only in Upper German (cf. Wegera 1987: 117).

To sum up, starting from dative *-e* as an unmarked form, the *-e* becomes a marked form before becoming a question of style in the eighteenth and nineteenth century (cf. Habermann 1997: 472). Only in phraseological word combinations has the *-e* been preserved for a very long time, in some cases up until today (cf. Elspaß 2012: 212). A short overview of the development of dative *-e* can be found in Table 3.1.

The early modern grammarians also dealt with the use of the dative *-e* (Table 3.2). Whereas Albertus (1573) and Kromayer (1618) argue for a dative without *-e*, Gueintz (1641) and Schottelius (1663) use dative *-e* consistently. We may observe here how the ongoing standardisation of German influences the authors' prescription of how to construct the dative (cf. McLelland 2014: 256). There is only one grammarian who tolerated variation and declared both forms as fit for usage: Clajus (1578:

Table 3.1 Percentages of *e*-apocope depending on region (cf. Wegera 1987).

	Upper German	Central German		
		Moselle-Franconian	Rhine Franconian-Hessian	East Central German
1550–1600	90–100 %	90–100 %	90–100 %	50–60 %
1650–1700	90–100 %	90–100 %	70–80 %	20–30 %

Table 3.2 The use of dative *-e* in early modern grammar.

Albertus (1573)	-Ø
Clajus (1578)	both forms
Kromayer (1618)	-Ø
Gueintz (1641)	*-e*
Schottelius (1663)	*-e*

45–6) explains that nouns ending in gemination (like *Bett* 'bed') can be used with or without *-e* to mark the dative.

A potential limitation to these approaches is that they deal with similar source material such as grammars, literary or institutional texts. Thus, they focus on language from the writing elites who typically perceived language from a normative point of view in order to refine it. This highlights the need for a new perspective on dative *-e* based on language use in everyday communication.

3. Early modern foreign language textbooks as a source for historical linguistics

When studying historical grammar in general, as well as special phenomena such as the German dative marker *-e*, the task of finding appropriate and reliable sources is a major challenge. This task is particularly difficult if we seek to examine historical texts by authors beyond the writing elite. A promising, yet so far rather neglected, type of research material is early modern textbooks for the teaching of foreign languages. In these textbooks, the authors apply, explain and transfer linguistic knowledge to enable learners to communicate in a foreign language. These textbooks, which gained wide distribution in early modern Europe, were multilingual and contained up to eight languages such as German, French, Italian, Spanish, English and Dutch but also Czech, Polish and Swedish in some cases. They were mainly addressed to tradespeople, craftsmen and travellers of all sorts. Thus, they were addressed to an at least literate but non-scholarly audience. Furthermore, there are also some books specifically addressed to soldiers and to aristocrats, as well as textbooks for women.

Another characteristic of these textbooks is that they were revised and reissued frequently – sometimes over a period of more than 100 years. In some cases, we can even observe aspects of language change in one revised and republished text written by one author. This allows us to keep track of the phenomena of language change such as sentence internal capitalisation or the development of German future constructions (cf. Hübner 2023).

The structure of the textbooks is often very similar. They typically contain a preface, a grammar, model dialogues and wordlists. Furthermore, there can be some additional elements such as model letters, prayers, or instructions for appropriate behaviour. In the preface, the authors justify their decision to write the book. They explain its advantages and its structure and give an explanation of how to use it. In this regard, the authors may differ in their concepts of language learning: some argue for the importance of memorising rules and then practising dialogues afterwards, while others emphasise the importance of using a foreign language in order to learn it.

As an illustration, the language teacher Sumaran (1617: 17) argues for rule-based acquisition to master a foreign language (ex. 2):

(2) Derwegen vmb zuerlangen die volkommenheit diſer Sprachen/ so in diſer Grammatica begriffen/raht ich dir ehe du in diſer Lånder/da man diſe ſprachen redt/verraiſeſt/ daß du zu einen Sprachmaiſter geheſt / der dich underweiß wol zuleſen vñ außſprechen/ die Sprach die du willst lernen/ dir zaigen die fundamenta derselbigen/ vnd dz er dich außwendig lerne Nomenclaturam ſo ich dir in

> diſem Buch hab beygeſetzt / zugethan mit den Verbis oder Coniugationibus [...] Da du ſolches ein halb jar thun wirſt/ wirſtu sehen/ dz du mehr haſt zugenomen / als wann du 2. jahr in Franckreich oder Welſchland werest geweſen / one die gemeltē fundamentis.
>
> 'To attain perfection in these languages contained in this grammar, I advise you to go to a language teacher before travelling to these countries where they speak this language. He will teach you to read and properly pronounce the language you want to learn and show you the basics of it. And he shall teach you the words that are in this book together with the verbs or conjugations. If you do this for half a year, you will see that you will have learned more than if you had been in France or any other Romance-speaking country [e.g. Italy] for two years without these aforementioned basics'.

However, the anonymous author of the *Tresor en trois langues, Françoise, alemande, latine* (1679: 11) explains that too many rules may hamper understanding. Therefore, he argues for a more usage-based foreign language acquisition including a reduction of rules (ex. 3):

> (3) [...] ist vor unnȍhtig erachtet worden / eine lange und beschwerliche Syntaxin (Wortfůgung) darzu zu setzen/ welche selten grossen Nutzen bringet; weilen die zierliche Wortstellung vielmehr durch die Übung und Lesung guter Autoren als durch so vielfaltige Regeln/ (Lehrsåtze) und Anmerckungen / welche der Gedåchtnuß eine Last sind / und dem Gemůht wenig Lust machen / erlehrnt wird.
>
> 'It is not necessary to use long and burdensome syntax (word order), which is seldom of much use, because appropriate word order is learned much more by practice and reading good authors than by so many rules (doctrines) and remarks which are a burden to the memory and give little pleasure to the mind'.

Furthermore, most textbooks contain a grammar section. The content and the elaborateness vary depending on the textbook, but it mainly contains basic information about the grammatical structure of a language on the level of phonology and morphology. The grammar is usually presented in paradigms complemented by explanations in a running text. In a few (predominantly) early textbooks, the paradigms are missing, or the grammar is presented in the form of a dialogue. In some textbooks, we can also find chapters concerning aspects of syntax, which was quite progressive for that time. The scholarly grammar tradition includes aspects of syntax systematically only from the eighteenth century on (cf. Moulin-Fankhänel 2000).

Every textbook also contains model dialogues. The dialogues are multilingual and mostly arranged in parallel columns. There are up to ten different interlocuters in different social settings. Usually, the dialogues contain everyday communicative situations such as asking for directions, sales talk, collecting debts, table talk, or small talk. These dialogues serve as sequences or chunks that the learner may use in different situations when travelling through Europe. Finally, textbooks often include word-lists that are arranged in alphabetical order or thematically.

Concerning the question of the normativity of dative *-e*, I examine the grammar as well as the model dialogues. The grammar shows the explicit grammatical rules, and the model dialogues give an insight into the application of these rules in 'spoken' everyday language.[2] With these means, we can study both the normativity and the use of dative *-e* in order to identify the underlying double standards.

Another crucial factor that distinguishes the textbooks from many other early modern texts is the authors themselves. Although some of the books are written by scholars, most are written by non-scholars (such as merchants or language teachers) with multilingual backgrounds. Some authors may be experts in domain-specific language use, and they have the cultural capital to write and publish their books, but most authors do not belong to the writing elites.

To sum up, there are two important characteristics of early modern textbooks that make them a useful source on questions of normativity and grammar. First, they are focused on language use and everyday communication. Their goal is to enable learners to communicate in a foreign language. Second, they are written by a fairly heterogeneous group of authors that consisted mostly of non-scholars.

2 The model dialogues contain a spoken language which is relatively authentic since they are addressed to non-scholars to enable them to communicate in everyday situations (cf. Ackermann 2021; Hübner & Gennies 2021).

4. Method

This study is based on the BDaFL (*Berliner Datenbank frühneuzeitlicher Fremdsprachenlehrwerke* 'Berlin data base of early modern foreign language textbooks'; Simon, Gennies & Hübner 2021), a data base of 150 multilingual foreign language textbooks from the early modern period, predominantly published in the seventeenth century and usually containing a grammar and a dictionary as well as model dialogues for various communicative situations. The corpus consists of textbooks containing German as well as other languages. The textbooks are all digitised, but they are not machine-readable, and therefore examples of linguistic items under study are collected and annotated manually. My analysis is based on a subcorpus of fifty-six textbooks from the BDaFL for two reasons. First, the main corpus includes some textbooks which would be inappropriate for analysing grammatical aspects of language use. Some books are reprints or even unmodified copies just published with modified metadata (e.g. author or date). These books are of interest regarding publishing processes but do not meet the authenticity criterion since the author, the year of publication and the corresponding text could be incorrect (cf. Hübner & Gennies 2021). Second, a smaller corpus was necessary as manual annotation is a lengthy process. The subcorpus consists of fifty-six textbooks (wherever possible) balanced in the dimension of mother tongue of author, year of publication as well as included languages.

To analyse the use of dative *-e* in the textbooks (N = 56), I annotated the grammar section as well as the dialogues. However, fifteen textbooks do not contain a grammar section and twenty-seven do not contain any information on the declension of the nouns and thus no information on the dative.

In addition, I annotated the first fifty dative forms from every model dialogue in the textbooks. I took only masculine and neuter dative forms into account where an *-e* could potentially be placed due to the phonological structure. Although the research literature does not mention any rules for an obligatory *-e*, there are certain restrictions for *-e* placement. For example, *-e* is not permitted for nouns ending with *-em, -en, -el, -er* or vowels (cf. Behaghel 1909: 34–8; Konopka 2010: 27).[3] This preselection resulted in a dataset of 2,610 instances of dative forms.

Previous studies have explored the conditions influencing the use of dative *-e*. These studies serve as a starting point for my analysis of the model dialogues. To identify the potential factors influencing the use of dative *-e*, the following parameters were used:

- word formation: while monosyllabic words tend to have the *-e* suffix, words with suffixes or proper names and foreign words tend to omit the *-e* suffix. No clear tendency can be established for compounds (cf. Behaghel 1909: 34–6; Konopka 2010: 27);
- the gender of the noun, its frequency, as well as its occurrence within a prepositional phrase (cf. Konopka 2010: 28);
- the metadata of the textbook and the author.

The data analysis was conducted in three steps. First, I analysed the use of dative *-e* in the textbook dialogues and the potential influencing factors involved. In order to test the impact of these variables, a binary logistic regression was applied. The regression as a multi-factorial procedure was used because the dependent variable is dichotomous (*-e* or without *-e*) and several potential independent variables exist. Second, I focused on the grammar sections in the textbooks to portray the normative ideas that might lay behind the formal grammatical descriptions. Finally, the variables were analysed concerning their respective characteristic values. In each case, the dependent variable is the use of *-e*.

3 These criteria are based on texts from different epochs. To avoid anachronistic conclusions, I annotated 250 datives without any phonological or morphological restriction (e.g. nouns ending in *-em*). By means of this pilot study, these restrictions could also be confirmed for dative *-e* within the textbooks.

5. Results and discussion

Previous research has established that during the early modern period, there is variation concerning the dative *-e*. That means that the marked and the unmarked forms coexisted. We now turn to the textbooks to see if and how the authors encountered this variation.

In order to obtain a holistic picture of the variation, I took two approaches. First, I reconstructed the norms of usage through the model dialogues. Second, I reconstructed the codified norm based on the formal grammar sections. We cannot be completely sure about the intention of the authors when writing the textbooks. However, since they are meant to teach German as a foreign language and to enable learners to communicate successfully, they can be seen as texts with a normative intention, or at least they were perceived as such by learners.

5.1 Norms of usage

This section deals with the question of whether there are any regularities underlying the use of those forms. In total, this study is based on 2,610 dative forms (first fifty datives from each textbook). Within the model dialogues (even of the same textbook), dative forms with and without *-e* coexist. However, in most cases (85%), the authors use the unmarked variant (without *-e*). That means that the use of dative *-e* in the dialogues is neither a matter of free variation nor dependent on generally valid rules. Above all, there are various factors which influence the use of the dative *-e* to a different extent: gender, word formation (e.g. compound nouns), word frequency and year of publication. Beyond that, extra-linguistic factors such as author, mother tongue of the author and the target language of the textbook play an important role.[4] There is only one variable which can be disregarded: the occurrence of the dative in a prepositional phrase has hardly any influence. There is only a slight difference regarding the dative endings within a prepositional phrase (fewer number of *-e*) than in datives without a preposition, this difference is not significant (cf. Hübner 2023).

To illustrate that the authors apply double standards (norms of usage and a codified standard norm) even within one textbook, I now focus on the year of publication and the word frequency as potential factors influencing the use of the variants. Figure 3.1 illustrates the proportion of dative *-e* in the model dialogues of the textbooks develops from 1535 to 1715. Each point represents the proportion of *-e* within one published textbook based on fifty dative forms.

This tendency is confirmed statistically by a binary logistic regression using the lrm- and the glm-function in R. To reduce the number of levels for one of the independent variables,[5] 'year of publication' is broken down to four periods from 1600 to 1750. Due to a lack of textbooks from the sixteenth century that include dialogues, the first period covers 100 years instead of fifty years. The time span from 1500 to 1600 is taken as the reference category, and the model compares the probability of the occurrence of *-e* of all the cases where it could theoretically appear in the other time periods with this reference category (cf. Table 3.3).

Table 3.3 also contains p-values, which indicate whether a factor has a significant influence as well as coefficients. Coefficients indicate the direction and the strength of the influence and are of particular importance. If the coefficient is positive, this characteristic value promotes the probability of *-e*. If the coefficient is negative, the characteristic value lowers the probability of *-e*. The p-value indicates the probability that the observed result or an even more extreme result will occur if the null hypothesis is

4 These results are based on Hübner (2023) and were obtained through a quantitative analysis of fifty-six textbooks. Their correlations were then verified using logistic regression.

5 It is necessary to make such a reduction to guarantee the operability of the model since too many single data points are obstructive.

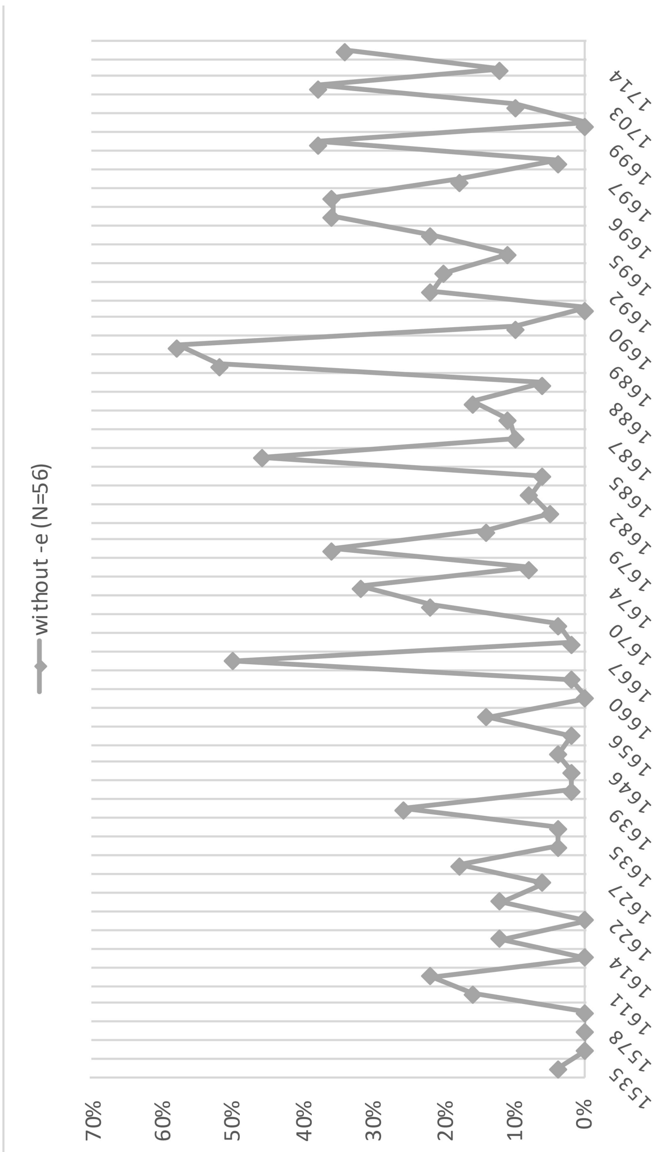


Figure 3.1 Influence of the variable year of publication (N = 56 textbooks).

Table 3.3 Influence of the variable 'year of publication' on the use of dative *-e*.

characteristic values	coefficients	p-value
1500–1600		
1601–1650	2.41774	*
1651–1700	3.05586	**
1701–1750	2.82686	**

Significance: < 0.001 '**' 0.01 '*'
Nagelkerke Pseudo-R^2: 0.232; C = 0,732

true (= no impact of the independent variable). A low p-value means that it is unlikely that there is no impact of the independent variable.

The variable 'frequency' aims to capture the influence of pattern formation on the use of the dative *-e* variant in model dialogues. In this case, the question is whether some lexemes are more likely to be used with *-e* than others. For this purpose, the lexemes were first divided into five frequency classes. Ideally, such a classification would be based on the word frequency in early modern language use. Due to the fact that corpora of Early New High German usually comprise written language texts such as documents, literary works and religious texts and not everyday language use, frequency cannot be reliably reconstructed. However, by comparing the frequency in different sources such as corpora and dictionaries (*Wortschatz Leipzig*, dlexdb.de and deutschestextarchiv.de), and the use of the nouns within the dialogues, four frequency classes were established, ranging from very rare (e.g. *Gespräch* 'talk' and *Schlaf* 'sleep') to very frequent lexemes (e.g. *Haus* 'house' and *Wort* 'word'). Table 3.4 illustrates the results of the logistic regression.

Table 3.4 Influence of the variable 'frequency' on the use of dative *-e*.

characteristic values	coefficients	p-value
very rare		
rare	0.03066	0.894020
Frequent	0.07017	0.733120
very frequent	1.01918	***

Significance: 0 '***'
Nagelkerke Pseudo-R^2: 0.232; C = 0,732

As already explained for the factor year of publication, the first characteristic value (here 'very rare') serves as the reference category. Table 3.4 shows that only the very frequent lexemes have a significant influence on the use of dative *-e*. For example, the proportion of dative *-e* for the very frequent lexemes is significantly higher (31%) than for the less frequent classes where only 12% of the dative forms have an *-e*. These results show that, at least for some authors, there seems to be a tendency towards pattern formation with the help of the dative *-e*.

Unfortunately, the individual lexemes could not be integrated as a factor into the model due to multicollinearity and a large number of different values. That means that individual variables are highly correlated with each other but are not exactly linearly dependent. However, there are some lexemes where the proportion of *-e* is higher than the average of 15% for all nouns or 17.9% for monosyllabic nouns, including *Bett* ('bed'), *Hof* ('farm, yard') and *Land* ('land, country') (see Figure 3.2).

Dative *-e* is not obligatory in any of the lexemes listed in Figure 3.2. Nevertheless, in some cases, the occurrence within a phraseological context seems to have an influence. While for *Bett* and the most frequent

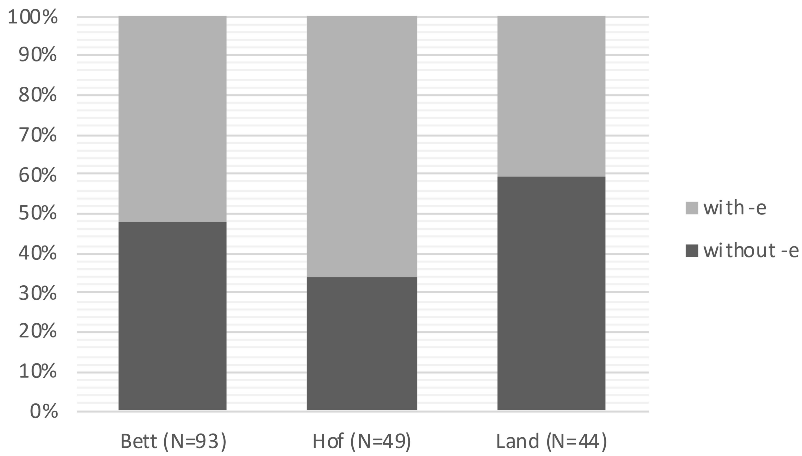


Figure 3.2 Lexeme-specific use of dative *-e*.

collocations *im Bett(e)* ('in bed') and *zu Bett(e)* ('to bed') both forms exist, the lexemes *Hof* and *Land* are predominantly marked with *-e*. Thus, *zu Hofe* ('at the court) (n = 15) is used exclusively with *-e*. And in the case of *auf dem Land(e)* ('in the countryside') (n = 11), except for two unmarked forms, there are only forms with *-e*. However, these findings are only preliminary because of the small number of textbooks involved and need to be validated based on a larger database.

Moreover, there are further indications of a lexeme-specific use of *-e* for individual authors. The French author Duez (1639; 1643; 1656; 1683) uses dative *-e* exclusively in four different nouns: *Bett*, *Feld* ('field'), *Haus* ('house') and *Ernst* ('earnest'). However, Duez also forms the dative of all four nouns without *-e*. Unfortunately, a comparison with the grammar is not possible since none of his grammars addresses the dative.

In conclusion, there is variation within one text written by one author. At the same time, we can observe individual author preferences. Some authors writing in the same period use dative *-e* more often than others (cf. Hübner 2023). Furthermore, there is a higher probability for *-e* in books that were published in the second half of the seventeenth century and when it is a frequent noun.

5.2 Codified norms

In the following, we will have a closer look at the grammar sections of the textbooks. The German language at that time was not a standardised language. However, as already mentioned in the introduction, norms are more or less required when learning a foreign language and especially when writing a textbook. This leads to the assumption that there may be some kind of tension or even contradiction between language reality on the one hand, and the underlying concepts of the textbooks on the other.

Unfortunately, there are only fourteen textbooks containing a grammar as well as an instruction on how to form the dative. Figure 3.3 illustrates that the grammar sections in the textbooks are quite heterogenous.

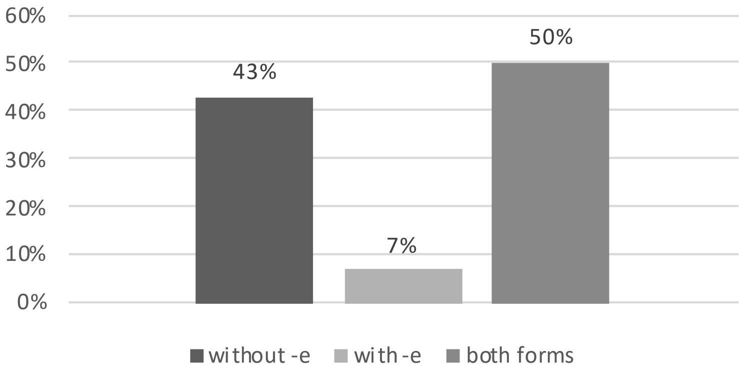


Figure 3.3 Use of dative *-e* in the grammar sections (N=14).

Only one author recommends using the dative *-e* as the appropriate form, whereas 43 % of the authors recommend the unmarked form exclusively to form the dative. Finally, 50 % of the authors use the marked and the unmarked dative as equal options.

Authors such as Martin (1635) explain the formation of the dative within the grammar section mainly with the help of more or less complete paradigms (ex. 4 for *Stab* 'stick').

(4) Der stab
Deß stabes
Dem stabe / ou stab
Den stab
O stab
Vom stabe / ou stab. (Martin 1635: 33)

In some cases, we also find some explicit further explanations on how to handle dative *-e*, but always without normative judgement. Canel (1689), for example, explains in his German grammar that it is possible to complement the dative of certain masculine or neuter nouns with an *-e* (ex. 5).

(5) Lorsque le Genitif singulier d'un nom masculin, ou neutre se termine en es/ ou s/tous les autres cas du singulier sont semblable au Nominatif, excepté les Monosillabes, au Datif ou Ablatif des quels on peut ajouter la voyelle, e/ (Canel 1689: 29)

'When the genitive singular of a masculine or neuter noun ends in *es* or *s*, all the other singular cases are similar to the nominative, except monosyllables to which in the dative or ablative one can add the vowel *e*'.

5.3 Comparison

We have already seen that it would be inaccurate to treat the textbooks as a homogenous text type. It is important to differentiate between grammar and model dialogues of the textbooks since they are heterogeneous with regard to their concepts of norm and variation. First, we can conclude that, compared to the model dialogues, the grammars allow less variation. While nearly every author uses both forms in the model dialogues, about half of the authors decide on one form or the other in their prescription. It can therefore be assumed that the textbooks apply double standards. The model dialogues represent norms of usage which are characterised by a higher tolerance of variation. In contrast, the grammar sections present the codified norm which is characterised by explicit rules and less variation.

If we compare this result to the grammarians' use of the dative *-e*, we can furthermore conclude that the grammar sections of the textbooks are more tolerant towards variation. Whereas all grammarians except Clajus (1578) decide on one form when they explain how to form the dative, 50 % of the textbook authors use both forms in their grammar.

Finally, we compare the results of the textbook analysis to previous research results. Up until now, the research literature has focused primarily on regional distribution concerning the relationship between *-e* and unmarked datives in the early modern period. Only very few studies exist on the further potential restrictions of dative *-e*, mostly concerning synchronic data (cf. Elspaß 2012; Habermann 1997; Radden 2021).

A comparison of the results gained so far reveals that there are some analogies compared to previous research but also some discrepancies. My findings roughly confirm the development of dative *-e* described in the research literature. However, if we take a closer look, we can see that dative *-e* was omitted and restored in a less linear fashion than indicated in previous research. Whereas many studies focused on the influence of the East Central German on the spread of the dative *-e* as well as the diachronic development, the present study includes different factors (word frequency, lexeme-specific use and context of use). My results on the trend mentioned in Section 5.2 confirm the observations from the research literature. Over the course of the seventeenth century, the use of the dative *-e* increased in the dialogues.

However, a look at the regional distribution shows some slight deviations. Whereas the research literature describes dative *-e* as restored almost everywhere except Upper German, the results from the textbooks show that dative *-e* is not an exception in Upper German textbooks.

6. Conclusion

The present research aimed to examine the use of dative *-e* in early modern multilingual textbooks. According to past studies 'from above' based on literary or formal texts, the use of dative *-e* seemed to be teleological and smoothed and mainly influenced by regional factors. Most of these studies focused on the process of standardisation, which gradually led to the omission of dative *-e* in favour of the unmarked form. However, these results demonstrate that actual language use over this time period was characterised by the coexistence of both forms, with a range of factors influencing the use of dative-*e*. Ongoing standardisation is a dynamic process that is by no means straightforward but runs parallel to the development of language use over time.

The textbooks allow us to take a new and more nuanced perspective on language use and also on standardisation. The authors apply 'double standards' even in their own work since most of the textbooks teach two different linguistic norms: a codified norm and a norm of usage, and they can thus be viewed as portraying double standards. The codified norm within the grammar section allows for almost no variation, whereas the norm of usage within the dialogues tolerates variants. The grammars show the explicit grammatical rules, and the model dialogues give an insight into the application of these rules and provide a framework for how language should be used. This also means that the early modern textbooks are characterised by variation. Thus, the textbooks respond not only to the question 'how should the foreign language be used?', but also to 'how is the foreign language actually used?' Therefore, the authors' goal was not about identifying the one correct form, but predominantly about depicting and teaching everyday language use. We can conclude that the textbook authors are participants in normative discourse in such a way that they taught and established norms (of usage) at the same time.

Since there are significant discrepancies compared to previous studies that analyse the development of dative-*e*, this points out the value as well as the necessity of using different sources to obtain as complete a picture as possible. The textbooks allow not only a new perspective on language use 'from below', but also from a multi-perspective view by including grammar sections, dialogues and insights into language awareness (prefaces). These findings raise intriguing questions regarding the nature of language use and language change taking place alongside the well-known history of standardisation. Based on the insights from analyses of such sources as the textbooks, the existing historiography of language should be revised and supplemented. In addition, early modern textbooks can offer insights into the cultural and social context in which language was used. The textbooks include cultural norms for language use, such as rules for politeness or etiquette, or offer information about the social status or prestige associated with different language varieties.

In this way, multilingual textbooks can be a valuable source for historical linguistics, as they provide insight into the way language was taught and understood during a particular period in history. These texts can reveal the norms, standards and expectations for language use at the time, as well as the techniques and methods used to teach language to learners.

Bibliography

Database

Simon, Horst, Linda Gennies & Julia Hübner. 2021. *Berliner Datenbank frühneuzeitlicher Fremdsprachenlehrwerke (BDaFL).* <http://dx.doi.org/10.17169/refubium-28414> accessed 8 June 2024.

Primary sources

Albertus, Laurentius. 1573. *Teutsch Grammatick oder Sprach=Kunst. Certissima Ratio discendae, augendae, ornandae, propagandae, conseruandaeque linguae Alemanorum siue Germanorum, Grammaticis Regvliset exemplis comprehansa & conscripta* (published by Carl Müller-Freureuth. 1895). Straßburg.

Anonym. 1611. *Dictionariolum hexaglosson cum colloquiis aliquot sex linguarum*. Leipzig.Anonym. 1679. *Tresor en trois langues […] Frantzösischer, teutscher und lateinischer Sprach-Schatz*. Basel.

Canel, Pierre. 1689. *Königliche teutsche Grammatic*. Nürnberg.

Clajus, Johannes. 1578. *Die deutsche Grammatik des Johannes Clajus* (published by Friedrich Weidling. 1894). Berlin, New York: de Gruyter Mouton, Ältere deutsche Grammatiken in Neudrucken 2.

Duez, Nathanael. 1683. *Frantzösische Grammatica renovata*. Frankfurt am Main.

Duez, Nathanael. 1639. *Le vray guidon de la langue françoise*. Leiden.

Duez, Nathanael. 1643. *Le vray guidon de la langue françoise*. Leiden.

Duez, Nathanael. 1656. *Kurtze frantzösische Grammatica […] Bref extraict*. Hanau.

Gueintz, Christian. 1641. *Deutscher Sprachlehre Entwurf.* Cöthen. <http://diglib.hab.de/drucke/ko-209-1s/start.htm≥ accessed 8 June 2024.

Kromayer, Johannes. 1618. *Deutsche Grammatica/ Zum newen Methodo/ der Jugend zum besten / zugerichtet. Für die Weymarische Schuel/ Auff sonderbaren Fürstl. Befehl.* Weimar, (Reprint 1986. Hildesheim & New York: Olms).

Martin, Daniel. 1635. *Acheminement a la langue allmande*. Straßburg.

Schottelius, Justus Georg. 1663. *Ausführliche Arbeit von der teutschen HaubtSprache* (published by Wolfgang Hecht. Deutsche Neudrucke. Reihe Barock 11 & 12. Tübingen: Niemeyer).

Sumaran, Juan Angel de. 1617. *Tyrocinium gallicum, italicum et germanicum*. München.

Secondary sources

Ackermann, Tanja. 2021. Bauernkomödien, Fremdsprachenlehrwerke und die Authentizitätsfrage – ein Vergleich anhand von Aufforderungen im 17. Jahrhundert. In Markus Denkler & Michael Elmentaler (eds). *Bauernkomödien des 17. Jahrhunderts als sprachhistorische Quellen*. Wien, Köln & Weimar: Böhlau. 69–103.

Hübner, Julia. 2023. *Norm und Variation. Paradigmenwechsel anhand frühneuzeitlicher Fremdsprachenlehrwerke*. Berlin, New York: de Gruyter.

Hübner, Julia & Linda Gennies. 2021. Zur Authentizität frühneuzeitlicher Sprachlehrwerke. In Julia Hübner & Horst J. Simon (eds). *Fremdsprachenlehrwerke in der Frühen Neuzeit: Perspektiven – Potentiale – Herausforderungen*. Wiesbaden: Harrassowitz. 225–40.

Behaghel, Otto. 1909. Der Dativ der Einzahl männlicher und sächlicher Hauptwörter. *Zeitschrift des allgemeinen deutschen Sprachvereins* 24, 33–9.

Elspaß, Stephan. 2005. *Sprachgeschichte von unten. Untersuchungen zum geschriebenen Alltagsdeutsch im 19. Jahrhundert.* Tübingen: Niemeyer.

Elspaß, Stephan. 2012. Wohin steuern Korpora die Historische Sprachwissenschaft? Überlegungen am Beispiel des *Neuhochdeutschen*. In Péter Maitz (ed.). *Historische Sprachwissenschaft. Erkenntnisinteressen, Grundlagenprobleme, Desiderate*. Berlin, Boston: de Gruyter. 201–25.

Elspaß, Stephan. 2015. Grammatischer Wandel im (Mittel-) Neuhochdeutschen – von oben und von unten. Perspektiven einer Historischen Soziolinguistik des Deutschen. *Zeitschrift für Germanistische Linguistik* 43(3), 387–420.

Elspaß, Stephan. 2020. Alternative sources of data for standardisation histories. *Language Policy* 19(2), 281–99.

Elspaß, Stephan, Nils Langer, Joachim Scharloth & Wim Vandenbussche. 2007. *Germanic Language Histories 'from Below' (1700–2000).* Berlin, New York: de Gruyter.

Glück, Helmut. 2002. *Deutsch als Fremdsprache in Europa vom Mittelalter bis zur Barockzeit*. Berlin, New York: de Gruyter

Habermann, Mechthild. 1997. Das so genannte 'Lutherische e'. Zum Streit um einen armen Buchstaben. *Sprachwissenschaft* 22, 435–77.

Hartweg, Frédéric & Klaus-Peter Wegera. 2005. *Frühneuhochdeutsch. Eine Einführung in die deutsche Sprache des Spätmittelalters und der frühen Neuzeit*. Berlin, New York: de Gruyter.

James, Gareth, Daniela Witten, Trevor Hastie & Robert Tibshirani. 2013. *An Introduction to Statistical Learning: With Applications in R*. New York: Springer.

Konopka, Marek. 2010. Niedrigfrequente grammatische Phänomene als sprachliche Zweifelsfälle. *Korpus – Grammatika – Axiologie* 2, 24–44.

Langer, Nils. 2002. On the Importance of Foreign Language Grammars for a History of Standard German. In Andrew R. Linn & Nicola McLelland (eds). *Standardization: Studies from the Germanic Languages*. Amsterdam, Philadelphia: Benjamins. 67–82.

McLelland, Nicola. 2014. Language description, prescription and usage in seventeenth-century German. In Gijsbert Rutten, Rik Vosters & Wim Vandenbussche (eds). *Norms and Usage in Language History, 1600–1900: A Sociolinguistic and Comparative Perspective*. Amsterdam, Philadelphia: Benjamins. 251–75.

McLelland, Nicola. 2015. *German Through English Eyes: A History of Language Teaching and Learning in Britain, 1500–2000*. Wiesbaden: Harrassowitz.

Moulin-Fankhänel, Claudine. 2000. Deutsche Grammatikschreibung vom 16. bis 18. Jahrhundert. In Werner Besch, Anne Betten, Oskar Reichmann & Stefan Sonderegger (eds). *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. Berlin, New York: de Gruyter. 1903–10.

Polenz, Peter von. 2013. *Deutsche Sprachgeschichte vom Spätmittelalter bis zur Gegenwart. Band II. 17 und 18. Jahrhundert*. 2nd ed. by Claudine Moulin. Berlin, Boston: de Gruyter.

Radden, Günter. 2021. Der verkannte *e*-Dativ im heutigen Deutsch: *im Kreise meiner Lieben* und *am Rande des Abgrunds*. In Anja Binanzer, Jana Gamper & Verena Wecker (eds). *Prototypen – Schemata – Konstruktionen. Untersuchungen zur deutschen Morphologie und Syntax*. Berlin, Boston: de Gruyter. 53–75.

Rutten, Gijsbert, Rik Vosters & Wim Vandenbussche (eds). 2014. *Norms and Usage in Language History, 1600–1900: A Sociolinguistic and Comparative Perspective*. Amsterdam, Philadelphia: Benjamins.

Scharloth, Joachim. 2005. *Sprachnormen und Mentalitäten: Sprachbewusstseinsgeschichte in Deutschland im Zeitraum von 1766–1785*. Tübingen: Niemeyer.

Takada, Hiroyuki. 1998. *Grammatik und Sprachwirklichkeit von 1640–1700. Zur Rolle deutscher Grammatiker im schriftsprachlichen Ausgleichsprozeß*. Tübingen: Niemeyer.

Wegera, Klaus-Peter. 1987. *Grammatik des Frühneuhochdeutschen. Band 3, Flexion der Substantive*. Heidelberg: Winter.

Wegera, Klaus-Peter & Hans-Joachim Solms. 2000. Morphologie des Frühneuhochdeutschen. In Werner Besch, Anne Betten, Oskar Reichmann & Stefan Sonderegger (eds). *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. Berlin, New York: de Gruyter. 1542–54.

ALENA A. FIDLEROVÁ

4 Which standard for Modern Czech? Three proposals from the turn of the eighteenth and nineteenth centuries*

ABSTRACT
This chapter explores three language standardisation proposals of the early phase of the Czech National Revival. It analyses their background, foundational principles and reasons for acceptance or rejection. First, Jan Václav Pohl's proposal is examined, rooted in a Baroque notion of language as a rational system. Pohl advocated elimination of alleged loanwords, and his orthographic reform aimed at greater regularity of spelling. Second, František Jan Tomsa's proposal is discussed, which combined features from classical literary tradition and colloquial spoken usage and brought Czech orthography closer to those of other Slavic languages in order to facilitate communication both between different social classes and different peoples. Finally, the most influential concept, of which Josef Dobrovský was the spiritual father, is introduced, favouring late sixteenth- and early seventeenth-century literary Czech over contemporary usage and comprising an orthographic reform driven by philological precision rather than everyday usability.

1. Introduction

In his article on selected issues in the history of Czech orthography, Berger (1999: 2) wrote that 'nemůžeme popsat cestu, když nevíme nic o slepých uličkách a zacházkách, které se nabízely vedle "správných" cest' ('we cannot describe the path if we know nothing about the dead ends

* This study was supported by the Charles University project *Cooperatio*.

and detours that were offered alongside the "right" paths'). In this chapter I want to take this claim seriously and examine the various proposals that were put forward during the early days of the standardisation of Modern Czech. In addition to the well-known proposal of which the eminent Slavist Josef Dobrovský (1753–1829) is considered the spiritual father, I will also describe two alternative proposals: The first was put forward by Jan Václav Pohl[1] (ca. 1720–1790), who was a teacher of Czech in Vienna, and the second one was produced by the philologist and translator František Jan Tomsa (1751–1814).[2] The inclusion of Pohl among those who intended to participate in the standardisation of Modern Czech may seem surprising from the point of view of the traditional 'school' conception of this period, which tends to see him as a representative of the previous, Baroque epoch (Šlosar, Večerka, Dvořák & Malčík 2009: 100). But already Schamschula (1973: 160–1) thought of him as one of the personalities of the National Revival period and described him 'in mancher Beziehung als der Gesetzgeber der Sprache, dem die weitere Entwicklung des Tschechischen recht gegeben hat' ('in some respects as the legislator of the language, who was proved right by the further development of Czech'). Similarly, Berger (1999: 5) believed that Pohl 'se pokoušel vší silou [...] ovlivňovat standardizaci češtiny' ('tried with all his might [...] to influence the standardisation of Czech'), and according to Newerkla (2016: 111), Pohl was among the first philologists who 'se [...] pokusil ideu obnovení češtiny realizovat' ('attempted to implement the idea of the restoration of Czech'). Pavkovič (2011: 61) called Pohl explicitly 'erste[r] Spracherneuerer eines neuen Typus' ('the first language renewer of a new type').

1 All the authors I deal with used both Czech and German versions of their names, usually in several spelling variations. For the sake of consistency, I decided to stick to the forms most common in Czech contexts today.

2 Pavkovič (2011) also deals in his book with the standardisation proposals of these three philologists, but he does not view them as alternatives, but as representatives of three 'generations'; he places Pohl in the first, Tomsa and František Martin Pelcl (1734–1801) in the second, and Dobrovský in the third. However, I find this division inaccurate, among other things because Pelcl's approach to language is closer to Dobrovský than to Tomsa.

Traditionally, the so-called first phase of the Czech National Revival was associated with the activities of Czech philologists publishing their works at the turn of the eighteenth and nineteenth century, including František Martin Pelcl, František Jan Tomsa and Josef Dobrovský (see e.g. Šlosar, Večerka, Dvořák & Malčík 2009: 105–6). However, Pavkovič (2011: 21, 43–5) moves the beginning of this epoch to the mid-eighteenth century and considers this initial period a 'microepoch', distinct from the following phases of the nation-building process and the standardisation of the national language. It was the time when the basic principles of Modern Standard Czech were negotiated. In retrospect it can be seen as the time of realisation of the first two steps of the standardisation process according to Haugen (1987: 59–61; see also Ayres-Bennett 2021: 31–5), namely selection and codification, and therefore it deserves more scholarly attention than it has received so far. However, in doing so, it is good to keep in mind that the intentions of its protagonists did not always exactly coincide with how their work was later understood and (not) used.

After a brief description of the historical background, I will devote individual sections to each of the standardisation proposals. First, I will discuss the position of each of the three authors within the community of Czech scholars, in particular on what they based their linguistic authority as 'language makers',[3] and what means they used to advance their ideas. I will then very briefly characterise each of the proposals and give selected examples at the level of spelling, morphology and lexis. I will primarily analyse the most advanced versions, that is, works or editions of works that were published last in the lifetime of their authors. However, when justified, I will also refer to earlier works. I will conclude by reflecting on why Dobrovský's proposal was the most successful, at least in the initial period, and by commenting on whether Modern Standard Czech is really based on the approach of Dobrovský and his followers only, or whether it bears any traces of the alternative proposals.

3 On the concept of language making and language makers see Krämer, Vogl & Kolehmainen (2022).

2. Historical background

The beginnings of the standardisation process of Modern Czech occurred in the second half of the eighteenth century, within the period classified by Miroslav Hroch as phase A of the Czech national movement (traditionally called the Czech National Revival), i.e. the period of scholarly interest (Hroch 1985: 23, 44). In that time, the Czech lands were bilingual, with a Czech-speaking majority and a German-speaking minority, and were part of the Habsburg Empire, with the Imperial Court and state administration residing in Vienna and using German as the primary means of communication.

Although the Czech language has had a literary tradition since the late thirteenth century, since the early seventeenth century it has gradually moved into the position of the L(ow) language (Fishman 1967). It was a consequence of the unhappy fight of Czech non-Catholic estates against the ruling Habsburg dynasty, resulting in the catastrophic defeat at the battle of White Mountain in 1620, and the subsequent forced recatholicisation and emigration of non-Catholic wealthy and educated elites. Thus, when searching for the 'Golden Age' of the Czech language, Czech patriots of the Enlightenment era looked back to the period of Renaissance Humanism (sixteenth to early seventeenth century), when Czech was widely used not only in literature, higher education, church and state administration, but partly also in the everyday communication of aristocratic elites. The following Baroque period (1620s–1740s) was perceived as a problematic period witnessing the gradual retreat of the Czech language from the literary and public scene and its replacement by Latin and later by German. At the end of this epoch, Czech was actively used mostly by uneducated inhabitants of rural areas and played only a secondary role in state administration. Nevertheless, Czech books continued to be published for the general public, and the notion of literary Czech as distinct from regional dialects was widespread even among lower-class users.

The enlightened educational and administrative reforms of Habsburg Empress Maria Theresa (1717–1780) and her son Joseph II

(1741–1790), seeking greater efficiency in state administration, were gradually more and more based on German as a universal means of communication. Their emphasis on the promotion of German-language education threatened to further displace Czech not only from public, but also from private communication and made Czech educated elites long for the return of the language of the 'Golden Age'. But besides this, there was another impulse urging Czech patriots to cultivate and standardise their language. During the second half of the eighteenth century, educated people at the Habsburg court felt that the Empire was lagging behind in the fields of scholarship and literature. To remedy this situation, German was to replace Latin as the language of higher education, and French as the language of the court and diplomacy. However, Upper German, the written form of a local spoken variety already in use for several centuries, was considered outdated, uncultivated and ultimately unfit for this purpose. Consequently, it was decided to replace it with the already standardised East Central German variety. The reform[4] was heralded by Johann Balthasar von Antesperg (1685–1765) of Lower Bavarian origin, whose grammatical works of the 1730s and 1740s were still based on the Baroque tradition, especially Justus Georg Schottelius (1612–1676). As a result, they were soon overshadowed by the works of Leipzig-based Johann Christoph Gottsched (1700–1766), who became the German-language authority not just in Saxony and Prussia, but also in Austria, and whose works were soon used to cultivate the written standard of adult users and for teaching at schools. The first decisive step towards the reform was taken in 1750 with the appointment of Johann Heinrich Gottlob (von) Justi (1720–1771) as Professor of German eloquence at the elite Viennese school Collegium Theresianum and the introduction of Gottsched's *Grundlegung einer deutschen Sprach kunst* as a textbook there. Later, Johann Christoph Adelung (1732–1806) and his works became the foremost authority, but the process of adoption of the new standard was not without opposition and lasted almost until the end of the century. In fact, it 'resulted in a process of invisibilisation, i.e. a disappearance of certain features in writing through stigmatisation, which

4 For more detail, see e.g. Rössler (1997); Wiesinger (2014: 313–97); Havinga (2018).

increased the gap between written and spoken language use' (Havinga 2018: 21).

According to Wiesinger (2014: 385), '[d]urch diese Verbindung von Sprache, Denken und Wissenschaft als Fortschrittsstreben zur Mehrung des Gemeinwohls treten in Österreich Aufklärung und Sprachreform in engen Zusammenhang' ('through this connection of language, thought and scholarship as a striving for progress leading to the increase of the common good, Enlightenment and language reform come into close connection in Austria'). This concept also naturally appealed to the intellectual elites of the non-German peoples in the Empire, but similar rights to reform, cultivate and standardise their languages were not readily granted to them, as they conflicted with the desired linguistic centralisation of the Empire. However, Czech patriots could not be discouraged by the disfavour of the authorities.

3. Competing standardisation proposals

3.1 Common features

All authors I am going to discuss were convinced adherents of Josephinism, the Austrian version of Enlightenment from above. Pohl taught at schools founded by Maria Theresa and dedicated his works to Joseph II, Tomsa took an active part in the enlightened school reforms by working as a secretary to Ferdinand Kindermann (1740–1801), a leading school reformer in Bohemia, and Dobrovský served for a certain time (1787–1790) as vice-rector and rector of the General Seminary in Hradisko near Olomouc, a Josephinist institution for the education of priests. Each of them also disposed of some type of authority on the basis of which they made their proposals: Pohl was a teacher of Czech as a foreign language, Tomsa was a primary school textbook translator, and Dobrovský an outstanding scholar. Their most obvious common feature is the fact that they wrote their works predominantly in German. This clearly distinguished them from the previous tradition, which used Latin and/or Czech as a metalanguage. According to Pavkovič (2011: 20), the use of German as a metalanguage marked a turning point in the grammaticographic tradition of the Slavic peoples of the Habsburg Empire. These works represented a new type of grammars, both oriented towards German-speaking scholars, and 'durchdrungen von der Idee, die slavische Sprache im Sinne der Aufklärung systematisch zu normieren oder aus dem vorhandenen Sprachmaterial Rückschlüsse auf vorhandene Regeln zu ziehen' ('in the sense of the Enlightenment imbued with the idea of systematically standardising the Slavic language, or of drawing conclusions about existing rules from the existing language material'). This means that all of these authors tried more or less successfully to provide a relatively comprehensive concept of Modern Standard Czech, comprising orthography, grammar and lexicon. The Czech language represented for all of them one of the Slavic dialects, a constituent part (and the best representative) of the Slavic language. Regarding their approach to spelling, all three of them demanded the use of the same system in manuscript and printed texts, thus rejecting the hitherto common practice of having separate orthographies for printers and scribes.[5]

3.2 Proposal of Jan Václav Pohl

3.2.1 Pohl: Linguistic authority

Little is known about Pohl's origins, but he is believed to have been born in Hradec Králové in northern Bohemia, in a linguistically mixed Czech-German environment (Newerkla 2004a: 43, 138). This, together with his almost life-long stay in Vienna, was often used as a pretext for questioning his competence in Czech. However, as his recently rediscovered

5 For more detail on these two orthographic systems see Fidlerová (2020a).

translation of a religious primer into Czech showed, his Czech was of a good standard (Newerkla 2004a: 44, 139).[6]

For most of his life, Pohl worked as a prominent teacher of Czech as a foreign language in Vienna, which justified his aspirations for the position of an outstanding Czech 'language maker'.[7] From its foundation in 1746, he taught Czech at Collegium Theresianum, which educated the sons of Catholic nobles for high positions in the state administration. Throughout its existence (1755–1769), he also taught young noblemen at the *Adelige Militärakademie* ('Noble Military Academy') in the Laimgrube suburb of Vienna (Newerkla 2000: 76; Newerkla 2004a: 43, 138). Even more importantly, he taught Czech to prominent members of the imperial family, including future Emperors Joseph II (since 1755), Leopold II and Francis I/II (Berger 1999: 5; Berger 2000: 63, 66–8). Consequently, Pohl considered himself 'státní expert pro oblast češtiny' ('a state expert in the field of the Czech language') (Newerkla 2001: 108), and based this claim on two pillars: First, on his profession as a language teacher, i.e. a person who grounds his approach to language on knowledge of its rules, and not on mere tradition (see e.g. Pohl 1786a: 40–2), and second on the official approval of the Emperor Joseph II (see e.g. Pohl 1786a: a3r; Pohl 1783: dedication). Moreover, Pohl maintained professional and friendly contacts with some Hungarian and South Slavic philologists, which also enabled him to (try to) claim the linguistic authority of a scholar (Newerkla 2000: 78; Newerkla 2004a: 44, 139; Newerkla 2016: 112).

3.2.2 Pohl's proposal: General characteristics

Pohl published and continually revised a textbook of Czech originally intended for his students. It was printed thrice under the title *Grammatica linguae Bohemicae oder Die Böhmische Sprach-Kunst [...]*

6 An edition of this translation was published by Newerkla (2004b).

7 See Krämer, Vogl & Kolehmainen (2022: 9, 14); Vogl & de Wilde (2022).

(1756, 1764, 1776 without the Latin part of the title) and twice under the title *Neuverbesserte Böhmische Grammatik [...]* (1773, 1783). It contained not only what we nowadays consider grammar but also examples of conversations on various topics and an ever-growing glossary, comprising ca. 6,000 words in the 1783 edition (Berger 2012: 237). Until the end of its existence in 1769, this textbook was compulsory at the Noble Military Academy in Vienna; at the *k. k. Theresianische Militärakademie* ('Theresian Military Academy')[8] in Wiener Neustadt its use lasted even longer, until 1806 (Newerkla 2000: 81–2; Čenský 1876, 93). However, Pohl did not regard it as a mere school textbook, but as a 'zu allgemeinen Nutzen des Publici gefertigte[s] Werk' ('work produced for the general benefit of the public') (Pohl 1773: 'Extrakt...').

Before composing his textbook, Pohl used extracts from the Latin grammar of Václav Jan Rosa (ca. 1630–89) *Grammatica Linguae Bohemicae / Cžechořečnost* (1672)[9] for his teaching (Newerkla 2000: 77–8; Newerkla 2004: 44, 139). This book decisively shaped Pohl's view of language. Rosa's intention was to write a rational, 'analogical' grammar of Czech;[10] he believed that, as an expert on the internal structure of the Czech language, he was entitled not only to describe, but to co-create the language on the basis of its general principles.[11] Pohl's textbook was so strongly influenced by this Baroque work that in some passages it can be considered its German paraphrase. Overall, Pohl adopted Rosa's rationalist approach, conceptualising language as an ideal regular system free of irregularities and gaps. Already in the preface (Pohl 1783: 'Vorrede') he discussed the decisive role of the 'innerliche

8 Established under the name Cadet House (*Kadettenhaus*) in 1752, under this name since 1768 (Newerkla 2012: 75).

9 See also Koupil (2015: 267–324). Despite originating in the Baroque period, this grammar was valued also by other leading 'Revivalists' (see e.g. Dobrovský 1809: X).

10 'Grammaticam Orationis, ſeu Linguae Boëmicae, Analogicam' (Rosa 1672: A3v).

11 In this, Rosa consciously followed the approach advocated in relation to Latin by another important Czech Baroque scholar, Jan Amos Comenius (1592–1670), who eventually advanced his efforts even further, to attempts to create an artificial universal language (see Koupil 2015: 277–8).

Vollkommenheit' ('inner perfection') of a language. According to him, Czech was 'in ihrem Weſen so vollkommene Sprache' ('in its essence such a perfect language'), the original principles of which had been in previous years violated by neglect and infiltration of foreign words, and Pohl aimed to restore it to its original state. Thus, even in the more recent editions of his work, Pohl was strongly influenced by Rosa, not in specific details, but in his overall approach to language: he deemed himself entitled to remove historically created irregularities and improve (presumed) deficiencies.

However, to describe Pohl's approach simply as a continuation of one branch of the Baroque tradition, the rational grammar, might not fully do justice to him. He was an ardent supporter of Joseph II, and it seems improbable that he would not have any contact with the works of his enlightened contemporaries. One possible source of inspiration could have been the work of Johann Heinrich Gottlob Justi, his already mentioned colleague at Collegium Theresianum. In his inaugural lecture, Justi argued that the happiness ('Glückſeligkeit') of a nation depended on the flourishing of academic studies. They, in turn, depended on the degree of perfection of the state language ('Landesſprache'), since academic studies could not be cultivated successfully in an unfit language. If cultivated in a foreign language, they could not lead to the welfare of the state. Similarly to Pohl, Justi talked about the perfection of a language ('die Vollkommenheit der Sprache'), but unlike Pohl, he saw it in the future, not in the past. As languages are subject to change, they can approach the state of perfection through 'rational improvement' ('vernünftigen Verbeſſerung') by experts such as teachers, who are entitled to subject various historical irregularities of the usage to general rules (Justi 1750: 11–4). Thus, although Pohl's approach to language and its perfection grew out of Baroque concepts and he himself never adapted his German style to the requirements of Gottsched and other German-language reformers (Pavkovič 2011: 132–7), in some important respects his views were close to Justi's, if not co-inspired by him.

3.2.3 Pohl's proposal: Orthography

In his reform of Czech orthography (see Pohl 1783: 4–10, 15),[12] Pohl sought to eliminate (most) historical irregularities and (re)introduce rational rules instead. He wanted the same (type of) phenomena to be written using the same (type of) signs and for one sign not to be used to signify different phenomena. In doing this, he was not entirely original; in fact, he mostly modified and systematised ideas already formulated by his model Rosa (Rosa 1672: 408–17). But even so, he was not always totally consistent. For example, to mark vowel length, which is phonemic in Czech, it was mostly customary to use a diacritic over the vowel; only for original Old Czech /ɪː/ a special letter <j> was used instead of <í>. The diacritic mostly took the form of an acute accent (called 'der gezogene Accent' or 'das lange Tonzeichen' by Pohl), i.e. <á>, <é>, <ý>, <ó>, but for historical reasons, in most positions a circle was used over <u> instead (<ů>). Pohl decided to drop the circle and write <ú> with an acute accent in all positions,[13] but, surprisingly, he kept <j>.

Another change concerned the diacritic called 'háček' ('hook') used traditionally both over consonants (<ž>, <č>, <ř>, <ď>, <ť>, <ň>) and the vowel <ě>. The grapheme <ě> was used either to mark palatalisation of the preceding consonant (in <dě>, <tě>, <ně>, pronounced /ɟɛ/, /cɛ/, /ɲɛ/ respectively), or the position of the Old Czech vowel jat' /ie/ after labials /b/, /p/, /v/, /f/, /m/ (<bě>, <pě>, <vě>, <fě>, <mě>, pronounced /bjɛ/, /pjɛ/, /vjɛ/, /fjɛ/, /mɲɛ/ since the fifteenth century). To remove this ambiguity, Pohl firstly reserved <ě> for the second case only, and for the first case moved the diacritic to the consonant. Secondly, he distinguished the shape of the diacritic over <e> (the circumflex accent: 'der gebogene Accent', 'das eingebogene Tonzeichen') from that over the consonants (the dot: 'das Düpflein', 'der Punkt').

In his effort to make Czech orthography as rational and systematic as possible, Pohl went partly beyond Rosa's suggestions. He proposed to

12 For more details, see Berger (1999, 2008, 2010: 348–50); Newerkla (2010).

13 Rosa (1672: 3, 409), on the other hand, preferred to use the circumflex accent instead of the circle.

use <i> and <y> to distinguish between 'soft' and 'hard' consonants in a more regular manner than was customary. As the syllables /dɪ/, /tɪ/, /nɪ/, /t͡sɪ/, /rɪ/, /zɪ/ were always written <dy>, <ty>, <ny>, <cy>, <ry>, <zy>, and the syllables /ɟɪ/, /cɪ/, /ɲɪ/ were written <di>, <ti>, <ni>, then also, according to him, /t͡ʃɪ/, /r̝ɪ/, /ʒɪ/ should be written <ci>, <ri>, <zi>, instead of traditional <či>, <ři>, <ži>.[14] We do not find this suggestion in Rosa; he (1672: 4, 415–7) required /t͡ʃ/, /r̝/, /ʒ/ to be spelled <č>, <ř>, <ž> in all positions.

3.2.4 Pohl's proposal: Morphology

In morphology, Pohl strove for regularity often regardless of usage and the age and origin of a given ending, and even sometimes regardless of what his model Rosa said. Thus, he did not reject the (late medieval) phonological innovations partly perceived in the Humanist period as low style, but he also did not fully embrace Baroque plurality of forms. In the case of 'hard' adjective endings, he did not follow Rosa in allowing for both older and newer forms as equal, but he used the doublets as a means of paradigm regularisation. Instead of the variant pairs *krásný/krásneg* (Nom. sg. masc.), *krásné/krásný* (Nom. sg. neut.), *krásnýho/krásného* (Gen. sg. masc. and neut.) etc. (Rosa 1672: 106–7), he reserved *krásný*, *krásnýho* for the masculina, and *krásné*, *krásného* for the neuters (Pohl 1783: 73).

Following Rosa, Pohl mostly approved of the adoption of endings from one noun paradigm to another, as it was current in the language of the Baroque period. Thus, he allowed for Instr. pl. endings *-y/-i*, *-ami/-emi* and *-mi* to be used with almost any noun, thus achieving a great degree of symmetry in the system of noun declension. Rosa, too, tended to such symmetry, but unlike Pohl he generally also accepted former dual endings in *-ama/-ema*. Pohl rejected them as rustic and uncultivated, which led to a reduction in the maximum number of permissible forms for each

14 In the 1786 modification of his proposal, Pohl went even further and applied the last two changes also to /ʃ/, hitherto written <ſſ> or <sſ> (Pohl 1786a, 1786b).

declension from five to three (Pohl 1783: 45–65; Rosa 1672: 60–100). This tendency to keep symmetry but reduce variability shows once more that despite being a proponent of Baroque rational grammar, Pohl was also already a participant in Enlightenment standardisation and language cultivation efforts.

Pohl even introduced a completely new rule for writing short <i> and long <j> in the endings, a distinction on the boundary between morphology and spelling, as in pronunciation this vowel tends to be shortened word-finally. He insisted that in the endings of nouns, adjectives and comparatives and superlatives of adjectives it had to be short (although traditionally, all of them had <j>), so that they could be distinguished from the third pers. ind. pres. of verbs, where the ending was long (Pohl 1783: 17, 63, 79–80, 84–7). This suggestion, obviously aimed at better distinguishing the forms of nouns and verbs in writing, may also be understood as a rationalising tendency, facilitating the reception of the written text, especially for less experienced users.

3.2.5 Pohl's proposal: Lexicon

Pohl's model Rosa listed forty-five suffixes in his grammar, granting the reader what Ondřej Koupil (2015: 304–5) calls a 'generative' ('generativní') approach to vocabulary: creating new words using rules and units with more or less well-defined meanings. This approach was adopted by Rosa also in his huge, unpublished, but not forgotten *Thesaurus linguae Boëmicae*, intended to present the system of Czech lexicon not only by collecting and classifying existing words, but also – and actually quite often – by filling the (perceived) 'gaps' with the help of the rules (see Koupil 2015: 305–14).

We have already shown that Pohl actually understood his role as restoring Czech to its original, perfect state, to which loan words were the greatest affliction (Pohl 1783: 'Vorrede'). Therefore, they had to be replaced by words of native origin, newly derived if necessary. Although he is traditionally accused of superfluous arbitrary neologisation (see Cuřín 1985: 64), he simply faithfully followed and (supposedly) improved his model, trying to be as systematic as possible. For example, Rosa (1672: 371, 373) listed among the nouns derived from verbs those ending[15] in *-tedłnjk* and denoting persons affected by a given activity (*Kázatedłnjk* 'listener to a sermon', *vkázatedłnjk* 'recipient of a demonstration', *včitedłnjk* 'pupil'), and those ending in *-njk*, to which he ascribed both passive and active meaning (*včedłnjk* 'pupil', *Poručnjk* 'mandatary', *odkaznjk* 'legatee'; *Miłownjk* 'lover', *dělnjk* 'worker'). Pohl (1783: 263–4) attempted to make this more systematic and distinguished between three endings: *-ted(e)łnik* with passive meaning (*Kázatedłnik* 'listener to a sermon', *Riditedłnik* 'subject to the government', *vcitedłnik* 'pupil'), *-telnik* with active meaning (*kázatełnik* 'preacher', *riditełnik* 'governor', *řecitełnik* 'person capable of oratory') and *-nik* with passive meaning (*poručnik* 'recipient of an order', *Odkáznik* 'legatee'; *Obſtrážnik* 'keeper, sequester').

3.3 Proposal of František Jan Tomsa

3.3.1 Tomsa: Linguistic authority

Tomsa[16] was born as a farmer's son in the Czech-speaking region of north-eastern Bohemia. Most of his active life he spent as the economic director of the Normal School Printing Office in Prague, which published mostly Czech translations of German primary school textbooks,[17] and as a translator and corrector of its production. He belonged to the Czech patriotic milieu and was a regular visitor to the house of the Knight Jan

15 I use this rather general term, as the word groups described by Rosa and Pohl were meant to differ not only in the suffix, but also in the base to which it was attached. As the descriptions of these bases are rather complicated and unprecise in some cases, I decided not to reproduce them here for the sake of brevity.

16 For more details, see Fidlerová (2018a).

17 About the important role of elementary language textbooks 'represent[ing] the target language in education', which in turn 'is a central force in the dissemination of language norms', see Havinga (2018: 97–8).

František z Neuberka, where its members met. In his younger years, he was close to Dobrovský, who wrote an introduction to his dictionary of 1791 and published there the first version of his study on the origin and word formation of Czech (Dobrovský 1791). He also managed to win supporters among the aristocracy; he taught Czech privately to the children of Count Christian Philipp Clam-Gallas and his wife Karolina, and probably also to the children of the widowed Princess Elisabeth zu Fürstenberg. Moreover, he twice (unsuccessfully) applied for the position of Professor of Czech language and literature at Prague University.

3.3.2 Tomsa's proposal: General characteristics

In one of his last works, Tomsa (1812b: 41–2) wrote about his idea of proper knowledge of a living language:

> Man ſoll daher, um genauere Kenntniß einer lebendigen Sprache zu erhalten, ſelbe nicht allein aus den guten Schriften, ſondern auch aus der gemeinen Rede ſtudiren. Es ist doch nicht alles grob in der gemeinen Rede, es giebt darunter viel feines.
>
> 'Thus, to acquire more precise knowledge of a living language, it is important to study not only good written works, but also ordinary speech. Not everything in the ordinary speech is coarse, much of it is pleasing'.

In its mature stage, Tomsa's concept of Modern Standard Czech represented a language free from narrowly regional phenomena, phenomena regarded in the Humanist and Baroque literature as low style,[18] and phenomena resulting from careless pronunciation or lack of knowledge of the written tradition.[19] This language was to be based both on the written monuments of the past and on the spoken usage of native speakers

18 In this respect, Tomsa's attitude changed during his life. In his early trilingual Czech-German-Latin handbook for children *Elementarwerk der böhmisch-deutsch- und lateinischen Sprache* (1784), he followed the spoken usage more faithfully and preserved even the allegedly low style phenomena.

19 See e.g. Tomsa (1812a: 5–6).

especially of lower social classes, because there, according to Tomsa, many old and typical Czech expressions not recorded in literature were preserved. Unlike Dobrovský, Tomsa cherished language largely as a means of communication, and thus aimed at the development of the full linguistic variational space of Czech. In this, he may have been inspired by his everyday tasks as a translator of elementary school textbooks and educational books and journals for villagers, where he had plenty of opportunities to experience the communicative inadequacy of the language modelled upon Baroque written usage. The same consideration for communication efficiency led Tomsa to advocate certain changes in Czech orthography, so that 'alle Slawen ihre čechiſchen Bücher leſen könnten' ('all Slavs could read their Czech books') (Tomsa 1812a: 3).

However, Tomsa's concern for spoken language was not only practically oriented. In his opinion, any living language was 'to be viewed as a mirror of the national character of the people' ('die Sprache eines Volks als einen Spiegel ſeines Nazionalcharakters anſehen'), and thus no written sources could make up for living spoken usage because they could never comprise it in its entirety. Tomsa was aware of the difference between the literary and spoken language but claimed that this difference was smaller in Czech and, consequently, that the Czech spoken usage was often quite correct (Tomsa 1812b: 39–42).

3.3.3 Tomsa's proposal: Orthography

Tomsa's proposal for orthographic reform went through several stages, but its basis seems to have always been the same – the desire to bring the graphic system of Czech closer to that of other Slavic languages. Consequently, he was dissatisfied especially with the way the consonant /j/ was spelled in Czech: mostly <g>, less frequently also <y> or <i> depending on the position in the word and other factors. The usage of <g> for /j/ particularly hindered Tomsa's goal, as the Czech phoneme corresponding to original Slavic /g/ is /h/, not /j/ (Tomsa 1802: 9). Tomsa gradually proposed several alterations to this usage; in the final version of his reform, he recommended that /j/ be spelled <j> before a vowel, <y> in the sequence of /ɛj/ which is the result of diphthongisation of /ɨː/ (written <ý>) and <i> after other vowels.[20]

Between 1793 and 1800, Tomsa made use of his position in the Normal School Printing Office and presented his proposal in small, anonymous Czech booklets containing basic spelling instructions intended for primary school children. This brought him various reproaches from public figures, including Josef Dobrovský (1803: 694–5), and eventually probably even a ban from his superiors (see Nejedlý 1828: 12). Consequently, he further refined his proposal in German linguistic treatises, published at his own expense, especially in the works *Über die čechische Rechtschreibung* (1802) and *Grössere čechische Orthographie* (1812a). However, the motivation to increase the similarity of Czech orthography with the writing systems of other Slavic languages was already formulated in one of the Czech booklets: 'Podobněgi geſt ſlowenſké řeči čeſky pſát ai [...] než ag [...]' ('It is more appropriate for the Slavic language to write *ai* [...] rather than *ag* in Czech [...]') (Tomsa 1800: 9). In the German treatise of 1802, Tomsa specified that he had in mind only the Slavic languages spoken in the Habsburg Empire, since the originally formulated general goal could have been understood as a tendency towards Pan-Slavism and Russophilia and consequently considered highly politically dangerous by the government. As his goal he stated the convergence of writing systems of these Slavic languages, being the first, easiest move to the convergence of the languages themselves, which in turn would facilitate their mutual understanding, and thus improve the functioning of the Habsburg Empire in general (Tomsa 1802: 3–4).

In his last treatise on the subject, Tomsa tried to practically demonstrate his aim in the appendix entitled 'Eine Probe, wie man nach der hier vorgeſchlagenen čechiſchen Rechtſchreibung alle ſlawiſche Dialekte ſchreiben könnte' ('A test of how all Slavic dialects could be written according to the Czech spelling proposed here'). Here he provided suggestions for using his system to write Polish, Croatian and – thus departing from his earlier declared intention to stick exclusively to the Slavic languages used in the Habsburg Empire – also Russian (Tomsa 1812a: 41–8).

20 For more details, see Berger (2010: 350–2), Fidlerová (2018b: 235–9).

The same goal was also pursued by his proposal to replace the black-letter (Schwabacher and Fraktur) used for printing Czech during the whole early modern period with the roman type (Antiqua), which was already in use for some West and South Slavic languages, e.g. Polish or Slovene.

3.3.4 Tomsa's proposal: Morphology

Tomsa's best known grammatical work is his *Böhmische Sprachlehre* of 1782, written partly under the influence of Dobrovský and partly modelled after the *Anleitung zur deutschen Sprachlehre* (published repeatedly from 1775 onwards, from 1779 under the title *Verbesserte Anleitung zur deutschen Sprachlehre*) by Johann Ignaz Felbiger (1724–1788), an educationist entrusted with the implementation of the reform of primary education in the Empire (see Keipert 1991). It was used at the University of Vienna (Vintr 2004: 18, 107; Reichel 2004: 31, 124), and thus competed indirectly with Pohl's grammar, which was used in other Czech-teaching institutions in Vienna.

To exemplify Tomsa's approach to morphology in this book, I will use the same examples as with Pohl. In the case of 'hard' adjective endings, Tomsa (1784: 173–85) provided two paradigms: one used by 'good writers' ('bei guten Schriftftellern'), comprising the older forms only, the other that the reader may encounter 'in everyday life' ('im gemeinen Leben'), showing greater variability and listing the newer endings together with selected older and regional forms. Concerning the Instr. pl. of nouns, he gave for masculina the endings *-y/-i* and *-ami/-emi*, but *-mi* exclusively in the declension in which it occurred in Old Czech, and otherwise only as an exception. The same holds generally for the neuters, with the exception of the 'soft' types, where he gave only *-mi* or only *-emi*. For the feminina (and small groups of masculina, which in Old Czech had the same declension as feminina) he allowed only for the endings *-ami/-emi* or *-mi* in the original Old Czech distribution. The originally dual ending *-ama/ -ema* is described as a matter of commonly spoken language only (Tomsa 1782: 66–172). Thus, in this early work Tomsa left a certain degree of variability but compared to Pohl he reduced it and took more account of the Old Czech distribution.

However, some of Tomsa's later works show a certain hesitation between following the classical norm and a greater acceptance of colloquial forms. As already mentioned, in his compact language textbook *Elementarwerk der böhmisch- deutsch- und lateinischen Sprache* (1784), Tomsa presented short Czech utterances of one or two phrases taken directly from ordinary private conversation and containing also newer (and perceived as stylistically lower) morphological and phonetical forms. Similarly, these forms occur occasionally in the numerous records of authentic spoken utterances found in the books *Über die Veränderungen der čechischen Sprach*e (1805) and *Von den Vorzügen der čechischen Sprache* (1812b). But their belonging to the standard is uncertain, as can be seen from Tomsa's transcription of extensive texts he had obtained from his father Jan Tomsa (1728–1803), a simple peasant from a small village in north-eastern Bohemia. Tomsa published them in two versions: first in faithful transliteration, and second in transcription, in which he modified not only the spelling but also the morphology and phonology towards the standard of classical texts (Tomsa 1805: 54–7, 224–32). However, he did not comment on the changes he made, so we cannot be sure about their status.

Finally, in morphology, too, Tomsa occasionally took into account the mutual comprehensibility of Slavic languages. For example, he recommended the rather colloquial infinitives ending in *-t* instead of *-ti*, as they also occur in other Slavic languages (Tomsa 1812a: 8–9).

3.3.5 Tomsa's proposal: Lexicon

Tomsa's works dealing with lexicon show both his appreciation for classical literary tradition and contemporary colloquial language and his dislike for neologisms. In his *Vollständiges Wörterbuch der böhmisch- deutsch- und lateinischen Sprache* (1791), Tomsa included both words he found in older Czech texts and colloquial expressions, but he omitted dialect and regional expressions. At the end of *Böhmische Sprachlehre*, an appendix entitled 'Vom Uiberſetzen' ('On translation') is added, in which Tomsa criticised Pohl's tendency to create unnecessary neologisms and recommended that before a writer thinks of coining a new word, he should first search in old books and dictionaries or ask ordinary people.

Tomsa's inclination towards greater appreciation of the spoken usage increased throughout the years. In his later works (1804, 1805, 1812b), he recorded innumerable examples of words, collocations, utterances and short conversations from everyday communication.[21] They represent mostly utterances overheard from conversations of his relatives, domestic servants, or passers-by, for which Tomsa often accurately noted when, where and from whom he heard them. For example:

> Den 15. Aprill 1806. ward eine Bäurinn, als ſie ins Zimmer kam, gefragt, ob es noch kothig wäre; ſie antwortete: sucho je tam, ale zas se ženau mračna, kothig iſt es nicht, aber es kommen wieder eilig Wolken an, folglich wird es regnen und wieder kothig werden. (Tomsa 1812b: 37)
>
> 'On 15 April 1806, a farmwife who came inside was asked if it were still muddy outside; she answered: It is dry there, but the clouds are coming rapidly, that means it is going to rain, and it will be muddy again'.

Not only did he record colloquial language, but he also utilised it in his own translations. As an example, he himself explained how he used expressions from his father's texts to improve his translations of elementary school textbooks (Tomsa 1805: 19–20).

21 For more details see Fidlerová (2017, 2020b).

3.4 *Proposal of Josef Dobrovský*

3.4.1 *Dobrovský: Linguistic authority*

It seems unnecessary to discuss in detail the linguistic authority of Josef Dobrovský,[22] who was one of the most respected Slavic scholars at home and abroad already during his lifetime. His occupation as a tutor in the family of Franz Anton Josef Count of Nostitz-Rieneck enabled him to associate with prominent aristocrats, and together with his activities in the Royal Bohemian Society of Sciences (*Královská česká společnost nauk*) provided him with the opportunity to travel to many European countries, including Russia. His works in various fields (Slavic, Czech, Hebraic and Biblical studies, textology, literary criticism, political, cultural, literary and ecclesiastical history, theology etc.) attracted the attention of prominent scholars, and Dobrovský corresponded regularly with many of them. Although subsequent generations of Czech patriots disagreed with him in many respects, he represented a personality whose views had to be taken into account. The range and thematic diversity of his work is so great that it still lacks adequate evaluation and, for the most part, modern critical editions.

3.4.2 *Dobrovský's proposal: General characteristics*

Although Dobrovský's concept of Modern Standard Czech took into account the contemporary usage to some extent, it was based predominantly on the best literary tradition of the past. In this, Dobrovský was inspired by Adelung, among others, with whom he corresponded throughout his life. They exchanged, reviewed and corrected parts of each other's manuscripts (Svobodová 1955: 25–31). And Dobrovský shared

22 On Dobrovský see e.g. Havránek & Dolanský (1953); Otruba (1985); Vavřínek, Gladkova & Skwarska (2004).

with his correspondent, among other things, his aversion to the creation of new words (see Svobodová 1955: 31–2).

According to Gladkova (2004: 376), Dobrovský's goal was to reestablish language continuity at the level of the prestigious, written tradition only.[23] Consequently, he chose the language of written documents created in the last historical period when Czech was generally and actively used by all social classes, the Humanist period of the 1520s to 1620s. He labelled the Humanist period 'das ſchöne oder guldene Zeitalter der böhmiſchen Sprache' ('the beautiful or golden era of the Czech language'), and contrasted it to the following Baroque period, during which '[d]ie böhmiſche Sprache geräth in Verfall' ('the Czech language falls into decline') (Dobrovský 1792: 171, 193).

3.4.3 Dobrovský's proposal: Orthography

Prior to the publication of his grammar, Dobrovský systematically rejected any attempts to change the traditional orthography, valuing highly its antiquity and stability which enabled the readers to understand easily even books three hundred years old (Dobrovský 1780b: 111–2). However, in the preface to the first edition of his *Ausführliches Lehrgebäude der Böhmischen Sprache*, he himself proposed an orthography reform abandoning the general usage of <y>/<ý> after <c>, <z> and <ſ> and replacing it by <i>/<j> or <y>/<ý> according to etymology and morphological analogy. This change obviously did not simplify the spelling, but rather made it more difficult for the writer. However, it brought written Czech closer to its Old Czech appearance and enabled Dobrovský to describe individual morphemes more easily. Consequently, he originally did not intend it to be obligatory in general usage, but to serve for the sake of scholarship only (Dobrovský 1819: XVI–XVIII). Nevertheless, thanks especially to the enthusiasm of Dobrovský's student Václav Hanka (1791–1861), it eventually gained general acceptance in the early 1840s.

23 See also Fidlerová (2020b).

3.4.4 Dobrovský's proposal: Morphology

Among the authors mentioned, Dobrovský's morphology was the closest to the usage of the cultivated authors of the 'Golden Age' extended, for the sake of Comenius, until the 1670s. His account of Czech morphology was all the more archaic as he paid increased attention to the usage of Czech Biblical translation (Dobrovský 1809: 38; cf. Ďurovič 2004: 406). According to Gladkova (2004: 384), Dobrovský's grammar was in fact originally intended not as a codification manual, but as a description and a defence of the best achieved state of the Czech language, which was to function as a symbol of the nation.

In his grammar, Dobrovský listed only the older forms of 'hard' adjective endings and allowed for regional forms only if they were older than the classical usage and were already recorded in Old Czech manuscripts (Dobrovský 1819: 202–3). As for the Instr. pl. of the nouns, he rejected the variability of endings, especially the usage of newer forms, and gave only the original endings *-y/-i* for most masculina and neuters and *-ami/-emi* for most feminina (and small groups of masculina, which in Old Czech had the same declension as feminina). He reserved *-mi* generally for the declensions in which it occurred in Old Czech, and tentatively mentioned only rare exceptions, especially in case of imminent ambiguity of meaning (Dobrovský 1819: 163–93).

3.4.5 Dobrovský's proposal: Lexicon

According to Dobrovský, grammaticographers and lexicographers do not belong among the people entitled to coin new words. This privilege is reserved only for writers who in their mother tongue lack a word for what they need to express, and even then, they must do so sparingly and 'nach der Analogie der Grammatik' ('according to the grammatical analogy'). 'Der Grammatiker, und Lexikograph hingegen ſoll den Sprachgebrauch als das Grundgeſetz ſeiner Regeln anſehen, und hat ſchlechterdings mit der Erfindung neuer Wörter nichts zu thun'. ('The grammarian and the lexicographer, on the other hand, should regard the language usage as the basic law of his rules, and has nothing to do with the invention of new words') (Dobrovský 1780a: 106). However, although Dobrovský himself engaged in the occasional collecting of words among the people (when working in Hradisko, he took walks in the countryside and studied the language of local inhabitants, Gladkova 2004: 378), his main source of usage was Humanist and early Baroque literature. As he pointed out, it was a duty of a Czech lexicographer, 'daß er die Wörter, die man in klaſſiſchen Büchern, z. B. in der Bibel, in Hageks Kronik u. ſ. w. findet, nicht auslaſſe' ('that he does not omit the words that one finds in classical books, e.g. in the Bible, in Hájek's *Chronicle* etc.'). If such a lexicographer was to find a word of unknown meaning there, then it was advisable to ask rural speakers of Czech, whose language was less corrupted than that of the cities (Dobrovský 1779: 254–5).

Dobrovský repeatedly warned against inventing new words, for example: '[M]an hüte ſich ſorgfaltig, neue Wörter zu erfinden. Nichts iſt in einer Sprache nachtheiliger, als das unmäßige Dichten der Wörter' ('Be careful not to invent new words. Nothing is more detrimental in a language than the intemperate creation of words') (Dobrovský 1779: 255). Ironically, it was his detailed description of Czech word formation (Dobrovský 1791, 1809: 14–132; 1819: 12–160) that enabled the following generation of patriots to forge neologisms on a large scale. In this work, Dobrovský proceeded quite differently from Rosa and Pohl: he preferred form over meaning and tried to analyse all Czech words into what he perceived as their smallest units. Therefore, he listed only the ending *-jk* together with a list of examples (including e.g. *zahradnjk* 'gardener', *ſlužebnjk* 'servant', *ſennjk* 'hay shed') without any specification of its general meaning (Dobrovský 1819: 47).

4. Conclusions

History is written by the victors. In retrospect, we often tend to regard the actual course of events as the only possible one, disregarding other contemporary alternatives as less valuable. However, in historical investigation, we should evaluate all alternatives impartially, considering their pros and cons, as well as the social and power relations and historical contingencies. Thus, we should not, as has been common up to now, 'wie selbstverständlich anzunehmen, dass Dobrovskýs Weg der einzig richtige, ja der einzig mögliche gewesen sei' ('to assume as a matter of course that Dobrovský's way was the only right one, indeed the only possible one') (Berger 2008: 41). At the given time, the situation was far less simple than that.

In this chapter, I introduced three diverging approaches to the standardisation of Modern Czech. All three authors conceived of their work as bringing Czech back to its former perfection, but each of them understood this perfection differently. For Pohl it was the perfect inner structure, a rational system that had been distorted by historical accidents and especially by loans from other languages. As a language expert, he felt entitled to remove the admixtures and to fill in what was missing. According to both Tomsa and Dobrovský, it was necessary to search for (almost) forgotten ways of expression in ancient writings and in contemporary usage, but each of them perceived the relationship between these two sources differently. According to Tomsa, actual usage, especially of simple speakers from the countryside who had not undergone formal education in German schools, was as important as the language of the surviving written monuments. As a translator of texts intended for inexperienced learners, he preferred communicative efficiency to historical purity. Moreover, the gradual convergence of the Slavic languages was also important to him, at least to the extent of enabling their speakers to read each other's books. For Dobrovský, ancient written sources were primary, and usage was merely a pointer to what in the old writings had already become hopelessly archaic and could not be revived.

The degree of occasional mutual misunderstanding among these authors is interesting, most obviously in relation to Pohl. Although he explicitly stated in his writings that he was concerned with regularity, with restoring the rational linguistic system of Czech to its original perfection, most of his contemporaries dismissed his work as arbitrary and confused, i.e. the complete opposite of what Pohl intended. We can only speculate whether they really were not aware of his intention, whether it was already so foreign to their own understanding of language that they did not actually recognise it, or whether they considered the imperfections in its implementation, which could indeed give the impression of arbitrariness, more important. They may even have deliberately chosen to ignore it, because they deemed a dispute about the philosophy of language too abstract and irrelevant.

As Krämer, Vogl & Kolehmainen (2022: 8, 10) remind us, '[p]rocesses of standardisation can be one out of many surface forms of Language Making', which in turn 'are intimately linked to power relations'. Although all of the authors discussed in this chapter disposed of some kind of linguistic authority, the scholarly prestige and social position of Dobrovský significantly contributed to the acceptance of his proposal. However, of at least equal importance was the fact that his proposal found a dedicated, energetic and determined promoter in Václav Hanka, who fought hard for the adoption of Dobrovský's spelling reform and adapted his grammar into Czech, thus bringing it to the general public's attention. Had it not been for his commitment, and for the fact that the members of the younger generation around Josef Jungmann (1773–1847) concentrated on the expansion of vocabulary – greatly aided by Dobrovský's theory of word formation – and were not really interested in reforming phonology and morphology, Dobrovský's grammar could have remained only a well-founded description of the Middle Czech and the standardisation of Modern Czech might have taken different paths.

On the other hand, Dobrovský's proposal cannot be described as the only winner of this dispute. In the long run its historicising tendency proved partly unpractical (first in theatre plays and fiction, later in newspapers) and had to be gradually adjusted to contemporary usage, as Tomsa had intended. This process was rather slow and resulted in Modern Standard Czech combining archaic and modern features. Neither did Pohl's concept prove to be a complete failure: some elements of his orthography reform were eventually adopted, although not until many years after his death.

Bibliography

Primary sources

Dobrovský, Josef. 1779. [Um den Bogen zu füllen…]. *Böhmische Litteratur auf das Jahr 1779*. Prague: Mangoldische Buchhandlung. 250–6.

Dobrovský, Josef. 1780a. Revision. *Böhmische und Mährische Litteratur auf das Jahr 1780* II/2. Prague: Mangoldische Buchhandlung. 93–111.

Dobrovský, Josef. 1780b. Apologie für die alte, einmal angenommene, in ganz Böhmen übliche böhmische Orthographie. *Böhmische und Mährische Litteratur auf das Jahr 1780* II/2. Prague: Mangoldische Buchhandlung. 111–28.

Dobrovský, Josef. 1791. Über den Ursprung und die Bildung der slavischen und insbesondere der böhmischen Sprache. In František Jan Tomsa, *Vollständiges Wörterbuch der böhmisch- deutsch- und lateinischen Sprache. Mit einer Vorrede begleitet von Herrn Rektor Joseph Dobrowsky*. Prague: von Schönfeld-Meißnerische Handlung. 11–32.

Dobrovský, Josef. 1792. *Geschichte der Böhmischen Sprache und Litteratur*. Prague: Calve.

Dobrovský, Josef. 1803. [Review] Modlitby pro Krestany Katolické; pracý a nákladem Frantisska Tomsy, d. i. Gebethe für katholische Christen; verfasst und verlegt von Franz Tomsa. *Annalen der Literatur und Kunst in den österreichischen Staaten* II. Jahrganges I. Band, Nr. 85, VII. Stück. 693–5.

Dobrovský, Josef. 1809. *Ausführliches Lehrgebäude der Böhmischen Sprache zur gründlichen Erlernung derselben für Deutsche, zur vollkommenern Kenntniß für Böhmen*. Prague: Herrl.

Dobrovský, Josef. 1819. *Lehrgebäude der Böhmischen Sprache. Zum Theile verkürzt, zum Theile umgearbeitet und vermehrt*. Prague: Haase.

Justi, Johann Heinrich Gottlob. 1750. *Abhandlung von dem Zusammenhang der Vollkom[m]enheit der Sprache mit dem blühenden Zustand der Wissenschaften*. Vienna: Trattner.

Nejedlý, Jan. 1828. *Widerlegung der sogenannten analogisch-orthographischen Neuerungen in der böhmischen Sprache*. Prague: Sommersche Buchdruckerey.

Pohl, Jan Václav. 1756. *Grammatica Linguae Bohemicae, Oder Die Böhmische Sprachkunst*. Vienna, Prague, Trieste: Trattner. 2nd edn 1764, Vienna: Trattner.

Pohl, Jan Václav. 1773. *Neuverbesserte Böhmische Grammatik, mit all erforderlichen tüchtigen Grundsätzen*. Vienna: van Ghelen's heirs; Prague, Brno, Lvov: Frag- und Kundschaftsamt. 2nd edn 1783, Vienna: Trattner.

Pohl, Jan Václav. 1776. *Die böhmische Sprachkunst*. Vienna: Gräfferische Buchhandlung.

Pohl, Jan Václav. 1786a. *Prawopisnost Řeči Čechské*. Vienna: Trattner.
Pohl, Jan Václav. 1786b. *Wahre gegründete böhmische Rechtschreibart*. Vienna: Hraschanzky.
Rosa, Václav Jan. 1672. *Grammatica Linguae Bohemicae / Cžechořečnost*. Prague: Arnolt z Dobroslavína sen.
Tomsa, František Jan. 1782. *Böhmische Sprachlehre*. Prague: Verlag der k. k. Normalschule.
Tomsa, František Jan. 1784. *Elementarwerk der böhmisch- deutsch- und lateinischen Sprache*. Erstes Bändchen. Prague: K. k. Normalschulbuchdruckerei.
Tomsa, František Jan. 1791. *Vollständiges Wörterbuch der böhmisch- deutsch- und lateinischen Sprache. Mit einer Vorrede begleitet von Herrn Rektor Joseph Dobrowsky*. Prague: von Schönfeld-Meißnerische Handlung.
[Tomsa, František Jan]. 1800. *Naučenj, gak se má dobře česky psát. Pro české sskoly*. Prague: Cís. král. normální škola.
Tomsa, František Jan. 1802. *Über die čechische Rechtschreibung mit einem Anhange, welcher dreizehn čechische Gedichte enthält*. Prague: self-published.
Tomsa, František Jan. 1804. *Über die Bedeutung, Abwandlung und Gebrauch der čechischen Zeitwörter*. Prague: self-published.
Tomsa, František Jan. 1805. *Über die Veränderungen der čechischen Sprache, nebst einer čechischen Chrestomathie seit dem dreizehntem Jahrhunderte biß jetzt*. Prague: self-published.
Tomsa, František Jan. 1812a. *Grössere čechische Orthographie, gemeiniglich böhmische Orthographie genannt*. Prague: self-published.
Tomsa, František Jan. 1812b. *Von den Vorzügen der čechischen Sprache, oder über die Billigkeit und den Nutzen, die čechische Sprache zu erhalten, empor zu bringen, und über die Mittel dazu*. Prague: self-published.

Secondary sources

Ayres-Bennett, Wendy. 2021. Modelling language standardization. In Wendy Ayres-Bennett & John Bellamy (eds). *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 27–64.
Berger, Tilman. 1999. *Z dějin českého pravopisu: Jan Václav Pól*. <https://homepages.uni-tuebingen.de/tilman.berger/Publikationen/Pohl.pdf> (accessed 30 November 2024.)
Berger, Tilman. 2000. Tschechischunterricht in der Habsburgerfamilie ab 1526. *Wiener Slavistisches Jahrbuch* 46, 61–71.

Berger, Tilman. 2008. Der Beitrag von Johann Wenzel Pohl zur Entwicklung der slavischen Sprachwissenschaft. In Sebastian Kempgen, Karl Gutschmidt, Ulrike Jekutsch & Ludger Udolph (eds). *Deutsche Beiträge zum 14. Internationalen Slavistenkongress Ohrid 2008*. München: Sagner. 39–52.

Berger, Tilman. 2010. Argumentace reformátorů českého pravopisu v letech 1780–1850. In Michaela Čornejová, Lucie Rychnovská & Jana Zemanová (eds). *Dějiny českého pravopisu (do r. 1902). Sborník příspěvků z mezinárodní konference Dějiny českého pravopisu (do r. 1902) 23.–25. září 2010, Brno, Česká republika. / History of Czech Orthography (up to 1902). Proceedings of the International Conference History of Czech Orthography (up to 1902) 23.–25. September 2010, Brno, Czech Republic*. Brno: Host/Masarykova univerzita. 343–59.

Berger, Tilman. 2012. Odkaz Jána Václava Póla dnešku. In Světla Čmejrková, Jana Hoffmannová & Jana Klímová (eds). *Čeština v pohledu synchronním a diachronním*. Prague: Karolinum. 235–9.

Čenský, Ferdinand. 1876. K dějinám řeči a literatury české v XVIII. století. *Osvěta* 6, 81–98, 251–67.

Ďurovič, Ľubomír. 2004. Josef Dobrovský a mluvnice Pavla Doležala. In Vladimír Vavřínek, Hana Gladkova & Karolína Skwarska (eds). *Josef Dobrovský. Fundator studiorum slavicorum. Příspěvky z mezinárodní vědecké konference v Praze 10.–13. června 2003.* Prague: Slovanský ústav AV ČR. 403–13.

Fidlerová, Alena A. 2017. Jak mluvil (a psal) český venkovan na přelomu 18. a 19. století? *Historie – Otázky – Problémy* 9(1), 48–65.

Fidlerová, Alena A. 2018a. Teaching Czech in a plurilingual community in the Age of Enlightenment: The case of František Jan Tomsa. In Nicola McLelland & Richard Smith (eds). *The History of Language Learning and Teaching I. 16th–18th Century Europe*. Cambridge: Legenda, Modern Humanities Research Association. 110–28.

Fidlerová, Alena A. 2018b. The fate of the letter <g> in the history of Czech orthography. *Beiträge zur Geschichte der Sprachwissenschaft* 28(2), 217–62.

Fidlerová, Alena A. 2020a. Orthography and Group Identity. A Comparative Approach to Studying Orthographic Systems in Early Modern Czech Printed and Handwritten Texts (c. 1560–1710). In Marco Condorelli (ed.). *Advances in Historical Orthography, c. 1500–1800*. Cambridge: Cambridge University Press. 154–75.

Fidlerová, Alena A. 2020b. *Sprache der Nähe* vs. *Sprache der Distanz* and the 19th Century Codification of Modern Czech. In Émilie Aussant & Jean-Michel Fortis (eds). *History of Linguistics 2017. Selected papers from the 14th International Conference on the History of the Language Sciences (ICHoLS 14), Paris, 28 August – 1 September [SiHoLS 127].* Amsterdam, Philadelphia: Benjamins. 205–19.

Fishman, Joshua. 1967. Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues* 23(2), 29–38.

Gladkova, Hana. 2004. Ještě jednou ke koncepci spisovného jazyka u Josefa Dobrovského. In Vladimír Vavřínek, Hana Gladkova & Karolína Skwarska (eds). 2004. *Josef Dobrovský. Fundator studiorum slavicorum. Příspěvky z mezinárodní vědecké konference v Praze 10.–13. června 2003*. Prague: Slovanský ústav AV ČR. 373–92.

Haugen, Einar. 1987. *Blessings of Babel: Bilingualism and Language Planning. Problems and Pleasures*. Berlin, New York, Amsterdam: de Gruyter.

Havinga, Anna D. 2018. *Invisibilising Austrian German: On the Effect of Linguistic Prescriptions and Educational Reforms on Writing Practices in 18th-Century Austria*. Berlin, Boston: de Gruyter.

Havránek, Bohuslav & Julius Dolanský (eds). 1953. *Josef Dobrovský 1753–1953: sborník studií k dvoustému výročí narození*. Prague: Nakladatelství ČSAV.

Hroch, Miroslav. 1985. *Social Preconditions of National Revival in Europe: A Comparative Analysis of the Social Composition of Patriotic Groups among the Smaller European Nations*. Cambridge: Cambridge University Press.

Keipert, Helmut. 1991. Die 'Wiener Anleitung' in der slavischen Grammatikographie des ausgehenden 18. Jahrhunderts. *Zeitschrift für Slavische Philologie* 51, 23–59.

Koupil, Ondřej. 2015. *Grammatykáři. Gramatografická a kulturní reflexe češtiny 1533–1672*. 2nd edn. Prague: Karolinum.

Krämer, Philipp, Ulrike Vogl & Leena Kolehmainen. 2022. What is 'Language Making'? *International Journal of the Sociology of Language* 274, 1–27.

Newerkla, Stefan Michael. 2000. Tschechischunterricht in Wien und Wiener Neustadt bis 1775. *Wiener Slavistisches Jahrbuch* 46, 73–84.

Newerkla, Stefan Michael. 2001. Toulka příspěvky vídeňských učitelů češtiny k národnímu obrození (k 225. výročí založení bohemistiky ve Vídni). *Listy filologické / Folia philologica* 124(1/2), 107–13.

Newerkla, Stefan Michael. 2004a. Josef Valentin Zlobický v kruhu svých předchůdců a současníků. Josef Valentin Zlobický im Kreise seiner Vorgänger und Zeitgenossen. In Josef Vintr & Jana Pleskalová (eds). *Vídeňský podíl na počátcích českého národního obrození – J. V. Zlobický (1743–1810) a současníci: život, dílo, korespondence. / Wiener Anteil an den Anfängen der tschechischen nationalen Erneuerung – J. V. Zlobický (1743–1810) und Zeitgenossen: Leben, Werk, Korrespondenz*. Prague: Academia. 42–60, 137–58.

Newerkla, Stefan Michael. 2004b. *J. V. Pohl: Překlad náboženské příručky – J. W. Pohl: Übersetzung einer religiösen Fibel*. In Josef Vintr & Jana Pleskalová (eds). *Vídeňský podíl na počátcích českého národního obrození – J. V. Zlobický (1743–1810) a současníci: život, dílo, korespondence. / Wiener Anteil an den*

Anfängen der tschechischen nationalen Erneuerung – J. V. Zlobický (1743–1810) und Zeitgenossen: Leben, Werk, Korrespondenz. Prague: Academia. 249–54.

Newerkla, Stefan Michael. 2010. Recepce pravopisu Jana Václava Póla jeho vídeňskými a vídeňskonovoměstskými současníky. In Michaela Čornejová, Lucie Rychnovská & Jana Zemanová (eds). *Dějiny českého pravopisu (do r. 1902). Sborník příspěvků z mezinárodní konference Dějiny českého pravopisu (do r. 1902) 23.–25. září 2010, Brno, Česká republika. / History of Czech Orthography (up to 1902). Proceedings of the International Conference History of Czech Orthography (up to 1902) 23.–25. September 2010, Brno, Czech Republic*. Brno: Host/Masarykova univerzita. 360–75.

Newerkla, Stefan Michael. 2012. Kontinuität und Wandel in der Minderheitensituation – Tschechisch in Wien bis 1775. In Claudia Woldt (ed.). *Tschechisch bis 1775 – historische Kontinuität oder Geschichte mit Sollbruchstellen? Beiträge zum 5. Bohemicum Dresdense. 12. November 2010*. München, Berlin, Washington: Sagner. 67–80.

Newerkla, Stefan Michael. 2016. Vídeňské počátky českého národního probuzení. In Ivo Cerman (ed.). *Habsburkové. Vznikání občanské společnosti 1740–1918*. Prague: Nakladatelství Lidové noviny. 109–23.

Otruba, Mojmír. 1985. Josef Dobrovský. In Vladimír Forst et al. (eds). *Lexikon české literatury. Osobnosti, díla instituce 1, A–G*. Prague: Academia. 560–71.

Pavkovič, Aleksander. 2011. *Tschechisch und Slovenisch im Vergleich ihrer schrift- und standardsprachlichen Entwicklung*. München, Berlin: Sagner.

Reichel, Walter. 2004. Josef Valentin Zlobický – první profesor českého jazyka a literatury: Život, působení a zásluhy na pozadí osvícenství. Josef Valentin Zlobický – erster Professor für böhmische Sprache und Literatur: Leben, Wirken und Verdienste vor dem Hintergrund der Aufklärung. In Josef Vintr & Jana Pleskalová (eds). *Vídeňský podíl na počátcích českého národního obrození – J. V. Zlobický (1743–1810) a současníci: život, dílo, korespondence. / Wiener Anteil an den Anfängen der tschechischen nationalen Erneuerung – J. V. Zlobický (1743–1810) und Zeitgenossen: Leben, Werk, Korrespondenz*. Prague: Academia. 24–41, 115–36.

Rössler, Paul. 1997. *Die deutschen Grammatiken der zweiten Hälfte des 18. Jahrhunderts in Österreich. Ein Beitrag zur Reform der deutschen Schriftsprache*. Frankfurt am Main: Lang.

Schamschula, Walter. 1973. *Die Anfänge der tschechischen Erneuerung und das deutsche Geistesleben (1740–1800)*. München: Fink.

Svobodová, Zdeňka. 1955. *Dobrovský a německá filologie*. Rozpravy Československé akademie věd 65(2). Prague: Nakladatelství ČSAV.

Šlosar, Dušan, Radoslav Večerka, Jan Dvořák & Petr Malčík. 2009. *Spisovný jazyk v dějinách české společnosti*. Brno: Host.

Vavřínek, Vladimír, Hana Gladkova & Karolína Skwarska (eds). 2004. *Josef Dobrovský. Fundator studiorum slavicorum. Příspěvky z mezinárodní vědecké konference v Praze 10.–13. června 2003*. Prague: Slovanský ústav AV ČR.

Vintr, Josef. 2004. Josef Valentin Zlobický – zapomenutý český vlastenec z osvícenské Vídně. Josef Valentin Zlobický – ein vergessener tschechischer Patriot aus dem Wien der Aufklärung. In Josef Vintr & Jana Pleskalová (eds). *Vídeňský podíl na počátcích českého národního obrození – J. V. Zlobický (1743–1810) a současníci: život, dílo, korespondence. / Wiener Anteil an den Anfängen der tschechischen nationalen Erneuerung – J. V. Zlobický (1743–1810) und Zeitgenossen: Leben, Werk, Korrespondenz*. Prague: Academia, 13–23, 101–14.

Vogl, Ulrike & Truus de Wilde. 2022. Teachers as foreign Language Makers: on standard language ideology, authenticity and language expertise. *International Journal of the Sociology of Language* 274, 107–31.

Wiesinger, Peter. 2014. *Das österreichische Deutsch in Gegenwart und Geschichte*. 3rd edn. Wien, Berlin: LIT.

ALENKA JELOVŠEK

5 Slovenian norms of address in the first half of the nineteenth century: Between purism and politeness

ABSTRACT

In this contribution, I present the discrepancies in the Slovenian language between the codified norms of polite address in the first half of the nineteenth century and their actual usage in the texts written in the same period. Building on my previous research on the development of the Slovenian address system, I highlight the specific period when, as the result of grammarians' interventions based on their puristic views on the language, the third-person plural address (*onikanje*), which was used in Slovenian towns, was deliberately suppressed in favour of the second-person plural address (*vikanje*), which was used in the countryside and expressed a lower degree of respect in the established Slovenian address system. As this was contrary to the social norms of address in the aristocratic and bourgeois classes, even the grammarians themselves used the third-person plural address when addressing people of higher ranks or German-speaking audiences.

1. Introduction*

Modern Slovenian has a bipartite system of address consisting of the pronouns for second-person singular and second-person plural, which is similar to other modern Slavic languages except for Polish. The situation

* This contribution was written as part of the research programme IMAGE – WORD – KNOWLEDGE. The Transmission and Transformation of Ideas on the Territory Between the Eastern Alps and the North Adriatic 1400-1800 (P6-0437), which was financially supported by the Slovenian Research and Innovation Agency.

in the Slavic languages is the result of convergent historical developments (Betsch 2000: 11). In some Western Slavic languages and Slovenian, the bipartite address systems, which had developed around the fifteenth century, were expanded to multipartite address systems, presumably under the influence of the neighbouring languages, especially German. Those systems were again simplified from the eighteenth century onwards, mostly by omitting the indirect forms of address (Berger 1998: 314).

The Slovenian (pronominal) address system had historically followed (and slightly lagged behind) its German counterpart until the beginning of the nineteenth century, when romantic nationalism and linguistic purism resulted in a deliberate remodelling of Slovenian norms of address to differentiate them from prevalent German address forms. The use of the third-person plural pronoun *oni* (used in predominantly German-speaking towns and modelled on the contemporary[1] German polite address pronoun *Sie*, cf. Reindl 2007) in polite address was deprecated as a Germanism by Slovenian grammarians in favour of the older form of address, the second-person plural pronoun *vi* (still used in predominantly Slovenian-speaking rural areas). This caused the collapse of the established system of address and, for nearly half a century, fostered uncertainty in usage, since grammatical rules were not aligned with the social norms of address in the aristocratic and bourgeois classes, and there was thus a danger that an addressee could perceive the use of *vi* address as impolite.

The aim of this chapter is to examine the discrepancies between the codified norms of polite address in the first half of the nineteenth century and their actual usage in the texts written in the same period. The chapter is structured as follows: Section 2 provides a short historical overview of the development of the Slovenian (pronominal) address system in comparison with the German address system until the end of the eighteenth century. In Section 3, a comparative description of grammarians' theoretical instructions in their grammars of Slovenian that were published in the first half of the nineteenth century and the use of pronominal address pronouns in their other texts (e.g. sample dialogues in the grammars, personal correspondence

1 The term *contemporary* is used in the meaning 'occurring at the same time' throughout the chapter.

in Slovenian, literary works) will be presented. Its aim is to determine which factors influenced grammarians' choice of pronominal form (e.g. text type, addressee's social status), and at what point in the analysed period the use of the second-person plural[2] pronoun *vi* became a generally socially acceptable form of address in reality and not only in grammatical codification. The conclusion considers the main findings of the study and outlines the development of the address system in the following periods.

2. The development of the Slovenian (pronominal) address system until the end of the eighteenth century in comparison with the German address system

It is assumed that the Slovenian system of address evolved under the heavy influence of German (cf. Reindl 2007: 156;[3] Jelovšek 2011), which was, together with Latin, the prestige language in the region for centuries (Ahačič 2014: 17–20). Through the centuries, the adoption of German-language elements into Slovenian took place at all levels of the language system, particularly through contact of speakers of the respective languages (influenced by historical, political, geographic and cultural circumstances spanning centuries), either in the process of importation (the adoption of foreign linguistic elements from the donor language into the recipient language) or – such as with forms of address – through substitution (replacement with similar elements typical of the recipient language) (Legan Ravnikar 2017: 36).

2 In this chapter, the abbreviations second pl. (for the second-person plural pronouns and verb endings), third pl. (for the third-person plural pronouns and verb endings), etc., are used.

3 Donald Reindl (2007: 156) substantiates this with the fact that the establishment of the same marked linguistic form, such as the use of the plural form for addressing a single person and, even more so, the use of the third person in the same situation, in languages in contact is more likely a sign of borrowing than independent evolution.

The use of *ir*, the polite second pl. form to address a single person, is first attested in German in the ninth century (Besch 1998: 92), and the bipartite system of forms of address became definitively established in the High Middle Ages (eleventh to thirteenth century). In the fourteenth century, nominal expressions consisting of a possessive pronoun and an abstract noun (*Euer Liebden*, *Euer Gnaden*) also entered the system. At first, these forms of address agreed with second pl. verbs and were contextually replaced by second pl. pronouns, but, in the sixteenth century, they started to agree with third sg. or pl. verbs and were contextually replaced with third-person pronouns. In the late sixteenth century, the next type of noun phrases of address evolved, mainly using the noun *Herr* ('sir'): this type mostly agreed with third sg. verbs; its placement on the status ladder was between *Ihr* and other forms of indirect address with abstract nouns. Besch (1998: 94) explains this by referring to the influence of absolutist intensification of etiquette, with people avoiding using direct address with *you* (singular or plural), opting instead for a form expressing greater distance.

In the seventeenth century, use of the *der Herr* type of indirect address increased, and a process of replacing nominal forms of address with anaphoric pronouns began. Since the 1620s, the third sg. personal pronoun could be used on its own instead of the noun *Herr*. For indirect forms of address of the *Euer Gnaden* type, replacement with a third pl. personal pronoun was possible since the 1680s or 1690s. This can be understood as intensifying the polite use of *Er/Sie* through plural. In the eighteenth century, third-person pronouns started to supplant the competing nominal forms of address. The pronominal address with *Sie* (third pl.) was definitively established by the 1730s. Compared with previous periods, the status of *Er/Sie* and *Ihr* started to decline rapidly in this period, further enhancing the importance of the plural *Sie*. In the following decades, the status of *Er/Sie* declined so much that it was even lower than the status of *Ihr*, but this was only a transitional stage before both forms finally disappeared from the standard system of address in the nineteenth century[4] (Betsch 2000: 51–2). Since then, *Du* and *Sie* have formed a new bipartite system still characteristic of German today.

In Slovenian, sources for the older period are scarce (cf. Mikhailov 1998); the oldest preserved document in a proto-Slovenian language, the Freising Manuscripts (972–1039),[5] contains a formula for general confession, but the confessor is still addressed with the second sg. personal pronoun (Jelovšek 2011: 197). The use of the second pl. *vi* pronoun of politeness is first attested in the thirteenth century and in circumstances that allow the assumption that it was directly influenced by the German norm of address, as it occurred in a short Slovenian phrase in the poetry collection by Ulrich von Liechtenstein entitled *Frauendienst* (approx. 1227).[6] It can also be found in a fifteenth-century text meant for monolingual speakers of Slovenian, in a formula for the general confession, which means it can be assumed that a bipartite pronominal system of address had become established in the Slovenian language by then. In the second

4 Both *Ihr* and *Er/Sie* were still used in early twentieth-century dialects of German (cf. Grober-Glück 1994 for an overview). In Switzerland and the South-West of Germany a high degree of polite second-person plural forms (*Ihr* or morpho-phonological variants thereof) can be found, and some areas of Northern Germany show third-person singular forms (Simon 2003: 106).

5 The manuscripts, written in the late tenth or the early eleventh century, are the oldest Latin-script continuous text in a Slavic language, which could be described as Alpine Slavic (ca. 800–1000) or early Proto-Slovenian (ca. 1000–1200) (Šekli 2018: 27ss).

6 In this collection, Ulrich presents his journey from Venice to Vienna. When he arrived in Friesach/Breže in Carinthia on 1 May 1227, the Carinthian duke Bernhard II of Spanheim greeted him with the words '*Buge **waz** sprimi, gralva Venus*' ('May God receive **you**, royal Venus'; Grdina 1999: 23). Since the duke was a member of the high nobility, which mostly spoke German only, it is very likely that this is a case of direct transfer of the contemporary German norm of address to Slovenian. In light of the later prevailing attitude towards Slovenian by the nobility and intellectuals as the uncouth language of the peasants (Grdina 1999: 186), it seems surprising that the Carinthian duke greeted his guest in Slovenian, the 'language of the peasants'. A possible explanation is the regional pride of the Carinthian noblemen, who liked to stress that their dukedom originated in Slavic Carantania to justify a special status within the German Empire (Štih & Simoniti 1995: 91).

half of the sixteenth century, the use of honorific titles is found in printed sources; in Protestant translations from German, syntactic agreement of the verb with the title in person, gender and number is predominant, reflecting the use of address in originals; on the other hand, the plural is predominant in verbs in non-translated texts from the late sixteenth and early seventeenth centuries. The third pl. pronoun *oni* became established as the contextual replacement for titles in the seventeenth century, and it started to be used independently in the early seventeenth century. A tri- or quadripartite (if the use of titles as the politest form is taken into account) system of address was thus formed, enabling a more precise expression of social differences.

Such usage at the end of the eighteenth century is best illustrated by the Slovenian source from the period that is richest in forms of address: Anton Tomaž Linhart's comedy *Ta veseli dan ali Matiček se ženi* ('This Happy Day or the Marriage of Matiček'; 1790), an adaptation of Beaumarchais's *Le marriage de Figaro*. The play can be assumed to provide a complete model of the (bourgeois) system of address of the time. The use of forms of address is schematised in Figure 5.1.

The baron and his wife occupy the highest social rank among the characters; the others address them with *vaša gnada* and third pl. verbs,

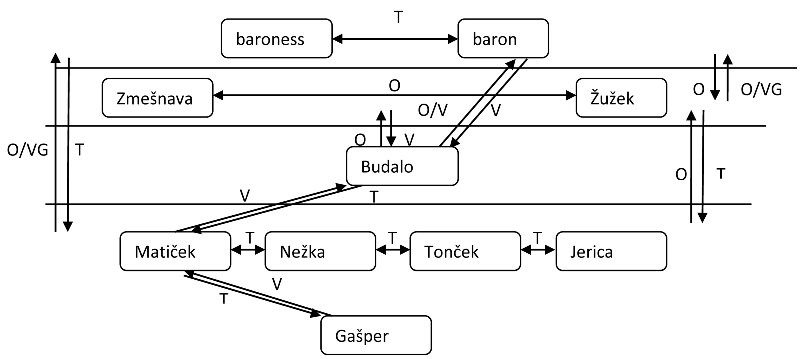


Figure 5.1 Address relations in *Ta veseli dan ali Matiček se ženi* (1790) (Jelovšek 2011: 202). O = third pl., V = second pl., T = second sg., VG = using the *Vaša gnada* title + plural address

occasionally using only third pl.[7] The title is contextually replaced by third pl. pronouns. The next rank is occupied by representatives of the class of clerks: Žužek,[8] a chancellor, and Zmešnava, a lawyer, who are addressed with third pl. both by their superiors – the baron and his wife – and their subordinates, while they use second sg. to address their subordinates. They use third pl. to address each other. An interesting case is Budalo, a scribe, who occupies an intermediate position between the higher and lower classes; his superiors use second pl. to address him, while he uses third pl. to address them, which indicates his lower social rank through language. He is also addressed with second pl. by his subordinates, and semantic agreement – with the auxiliary verb in plural but the participle in singular (cf. footnote 17) – is used.[9] The lower class is comprised of the castle servants as well as the baroness' godchild, Tonček, and the mayor's daughter, Jerca. These people mostly use second sg. to address one another, but there are also differences in status between them; for example, Gašper, a worker, uses second pl. to address Matiček, who was a gardener until recently but is now supposed to be the baron's personal servant; consequently, Gašper occupies the lowest place in the social hierarchy, and Matiček uses second sg. to address him. This work thus outlines a quadripartite system of address: second sg. – second pl. – third pl. – addressing with *vaša gnada* (in which the title agrees with third pl. verbs and is contextually replaced by the pronoun *oni* 'they').

To summarise, until the nineteenth century, the evolution of the second pl. and third pl. address in Slovenian and German was almost parallel (as can be seen in Figures 5.2 and 5.3), with the Slovenian evolution lagging behind only slightly. However, in Slovenian, there is no evidence

7 For examples on the usage of the forms of address in the play, see Jelovšek (2011: 202–3).

8 While the baron uses second pl. to address Žužek in one case, this is probably an inconsistency as third pl. is used in all other cases.

9 Supposedly, this used to be characteristic of addressing subordinate people (Toporišič 1992: 122), but that is not the case here as Budalo uses second sg. to address Matiček and clearly occupies a higher social rank.

of the gradual development of the *on* ('he') pronoun of address,[10] which arose as a replacement for the title *gospod* ('sir'); this shows that Slovenian did not borrow the full German system of address and did not adopt all its innovations, instead developing a system of its own. However, this system was undoubtedly influenced by German, as we can conclude on the basis of the prevalent translations from German and the social status of the languages.

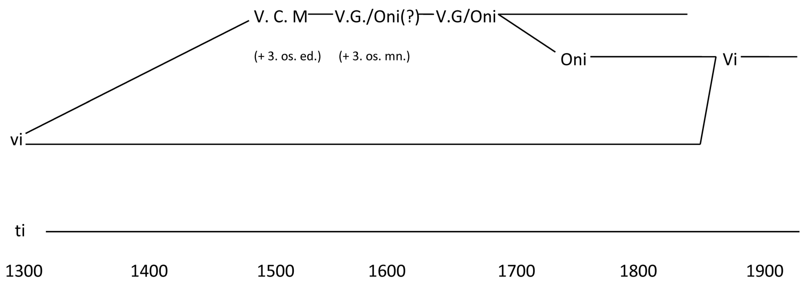


Figure 5.2 Evolution of the Slovenian system of address (Jelovšek 2011: 207). *V. C. M* = addressing with *Vaša cesarska majesteta* 'Your Imperial Majesty', *V. G.* = addressing with *Vaša gnada* 'Your Lordship'

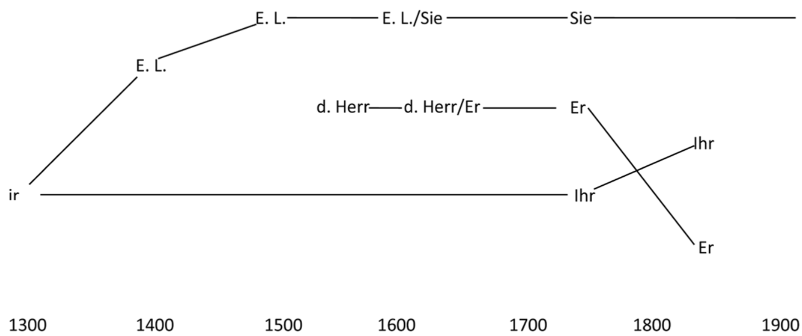


Figure 5.3 Evolution of the German system of address (Betsch 2000: 52). *E. L.* = addressing with *Euer Liebden* etc., *d. Herr* = addressing with *der Herr*

10 The only attestation of this form of address in Slovenian which I discovered appears in a nineteenth-century grammar and will be presented in Section 3.3.

The complex systems of address developed in both languages were appropriate for a time of great social differences, when there was a need for precise differentiation of social ranks, but afterwards – like the systems in some other languages[11] – it entered a process of simplification. However, the evolution of the Slovenian address system is typologically unusual, as one of the universal tendencies in the development of these systems of address is that a pronominal form used to express greater respect or social distance is never replaced by a form expressing less respect or social distance (Head 1978: 192). Thus, with the simplification of the German system of address, the intermediate forms – *Ihr* and *Er* – were abandoned, and *Sie* was preserved as the pronoun of politeness in the new bipartite system. In Slovenian, such a development was prevented by the purist views of nineteenth-century grammarians, who started to suppress the third pl. address as a Germanism.

3. The remodelling of Slovenian norms of address in the first half of the nineteenth century and the discrepancies between grammatical norms and actual usage

3.1 Jernej Kopitar

The new attitude towards address with the third pl. *oni* in Slovenian was first expressed by Jernej Kopitar[12] in his *Grammatik der slavischen*

11 The simplification of address systems is attested e.g. for English and Dutch (cf. Aalberse & Stoop 2015) and Portuguese (Lara-Bermejo, Bruno Guilherme 2021: 40); in Swedish, such a development is more recent (Jucker & Taavitsainen 2003). Some other Slavic languages also underwent similar developments as Slovenian, e.g. Czech and – to some degree – Slovak (cf. Kretzenbacher, Hajek, Lagerberg & Bresin 2019).

12 The use of second pl. and third pl. for polite address had been previously described in eighteenth-century grammars (Pohlin 1768: 138; Pohlin 1783: 176; Gutsman 1777: 90) as equal alternatives to the German polite address with third pl. *Sie* and no differences in their use had been mentioned.

Sprache in Krain, Kärnten und Steiermark (1809) and was based on the principles of romantic philosophy (especially that of Herder). He adopted the assessment of the language of the peasants as being purer and having fewer foreign influences compared to the language of the bourgeoisie and intellectuals (cf. Orožen 1981; Merchiers 2001: 110). Since the bourgeois modes of address were the same as in the German of the time, while rural areas still used the earlier form of address *Ihr* that was dying out in Standard German, he claimed the third pl. address was a Germanism (1):[13]

> (1) Der Krainiſche Landmann redet einen ältern, dem er Ehrfurcht ſchuldig iſt, mit vi (ihr) an, z B. *kakó* ***ſte***$_{\text{AUX-2PL}}$ ***rèkli***$_{\text{PART-PL}}$*, ozha, wie sagtet ihr, Vater? shèna, kaj bi* ***radi***$_{\text{ADJ-PL}}$ *Frau, was wünschtet ihr? Mósh, lepó* ***vaſ***$_{\text{PRON-2PL-ACC}}$ *próſim,* ***pomagájte***$_{\text{IMP-2PL}}$ *mi, Mann, ich bitte euch schon, helfet mir. Kaj* ***vam***$_{\text{PRON-2PL-DAT}}$ *je, mati? Al* ***ste***$_{\text{AUX-2PL}}$ **bolní**$_{\text{ADJ-PL}}$*? was ist euch, Mutter, seid ihr krank?*
>
> Die Krainer in der Stadt überſetzen das Deutſche Sie (oní), und gebrauchen es wie die Deutſchen*: kakó ſe* ***pozhútijo***$_{\text{PRS-3PL}}$*, wie befinden Sie sich?* ***Njìm***$_{\text{PRON-3PL-DAT}}$ *ſe pòzh dòbro godí! Ihnen geht es wohl gut! Kaj* ***márajo***$_{\text{PRS-3PL}}$ ***oní***$_{\text{PRON-3PL}}$ *Was kümmern Sie sich?*
>
> Eben dieſe Städter gebrauchen denn auch das vi gegen Dienſtbothen, Bauern etc. nach Deutſcher Art, und gewöhnlich ſo, daß, wie beym Italiäniſchen und Franzöſiſchen Höflichkeits *voi* und *vous*, das dazu gehörende Adjectiv etc. im Singular ſtehe: ſo hört man die Hausfrau zut Köchinn sagen: *Predragó* ***ſte***$_{\text{AUX-2PL}}$ ***kupíla***$_{\text{PART-SG.FEM}}$ *ihr habt zu theuer gekauft*; oder auf dem Markte zum Bauer*: O!* ***ſte***$_{\text{AUX-2PL}}$ ***predrág***$_{\text{ADJ-SG.MASC}}$ *O! ihr seyd zu theuer.* (Kopitar 1809: 288–9)
>
> ‘A Carniolan peasant addresses an elder that deserves his respect with *you* (*ihr*), e.g. *what did you say, father? what would you like, woman? Man, could you please help me? What is it, Mother? Are you ill?*
>
> Carniolans in towns are adopting the German *Sie* (*they*) and use it the same as Germans: *How are you feeling? You are doing very well! Why do you care?*
>
> These same townsfolk use *you* when addressing servants, farmers etc. the German way, usually so that, like with the Italian and French polite *voi* and *vous*, the related adjective is singular: thus, one can hear a mistress tell her cook: *You*

13 It is worth mentioning that even some German linguists rejected the third pl. address. As late as 1847, Jacob Grimm described it as unnatural and eccentric, expressing a belief that this form had not originated in the people, but was forced upon them by the higher classes (Besch 1998: 95).

bought this for too much money; or in the marketplace to a farmer: *Oh! You are too expensive*'.

Even though he did not expressly reject the use of third pl. pronoun for polite usage and advocate for the second pl. *vi* to be used in rural areas, this negative attitude could be inferred based on his generally purist attitude towards German and the emphasis on the language of the peasants that permeated his grammar, which combined xenophobic (anti-German) and ethnographic purism, as well as the archaic purism that marked the development of the Slovenian language (Thomas 1997: 135–6).

In Kopitar's rare Slovenian letters (cf. Suhadolnik 1981), we can find the second pl. address for his fellow linguist Valentin Vodnik (2), but he addresses his noble patron, Baron Sigismund Zois, with third pl. combined with the honorific title *Vaša gnada* 'Your Lordship', parallel with the German *Euer Gnaden* (3).

(2) Ki vém, de kar Goſpod Baronu od Slovenſkih rezhi piſhem, **vi**$_{\text{PRON-2PL-NOM}}$ tudi **ſliſhite**$_{\text{PRS-2PL}}$, inu ki vém, de ne **marate**$_{\text{PRS-2PL}}$ sa dolge piſma, inu ki tudi ne vtegnem, sato **vam**$_{\text{PRON-2PL-DAT}}$ ne piſhem. Bog **vam**$_{\text{PRON-2PL-DAT}}$ lonej, ki **ſte**$_{\text{AUX-2PL}}$ ſe s' mojo Grammatiko **vpirali**$_{\text{PART-PL}}$: zhe **vam**$_{\text{PRON-2PL-DAT}}$ morem kadaj per Beſediſhu poverniti, bom veſél. (Prijatelj 1926: 121)

'As I know that you hear all that I write to the Baron pertaining to Slovenian, and I know that you do not care for long letters, and as I also lack the time, I do not write to you. God repay you for your work for my grammar. If I can repay you by helping with your thesaurus, I will be glad'.

(3) Al **ſo**$_{\text{AUX-3PL}}$ **Vaſha gnada** vunder smiram sdravi? Skoraj bi to ménil, ki mi vunder tolkokrat **piſhejo**$_{\text{PRS-3PL}}$, kar nikoli niſim vréden. Pa tudi vſako verſto kuſhnem, ktiro ſo **njih**$_{\text{PRON-3PL-GEN}}$ perſti piſáli. (Kidrič 1939: 152)

'Is Your Lordship always well? I would almost think so, as you write to me so often, of which I am not worthy. I kiss every line your fingers have written'.

3.2 Valentin Vodnik

A similar distribution can be found in the works of Valentin Vodnik, the author of the first Slovenian grammar with only Slovenian as the metalanguage (Legan Ravnikar 2015). In it, we find a direct condemnation of the use of third pl. as a Germanism that should only be used when addressing 'Germans' in Slovenian (4).

> (4) Viſhi in imenitneji ludi, ktire poſebno poſhtujemo, vſelej imenujemo s vi nameſt ti, oni nameſt on, poſtavim: ozhe! **vi**$_{\text{PRON-2PL-NOM}}$ ſte rekli, de me **bote**$_{\text{AUX-2PL}}$ v' ſholo **dali**$_{\text{PART-PL}}$. Mati$_{\text{SG}}$ goſtje **napravlajo**$_{\text{PRS-3PL}}$, sa to ki **bodo**$_{\text{AUX-3PL}}$ moj ſtriz$_{\text{SG}}$ k' nam **priſhli**$_{\text{PART-PL}}$.[14] To ſe pravi, de jih ne tíkamo ampak víkamo, kadar ſo prizho; ino kadar jih tukaj ni, jih oníkamo. Prizho koga oníkati, je némſhvanje ino proti Nemzam navadno. V'pridigah ino viſokih govorih je li víkanje in tikanje v' navadi. (Vodnik 1811: 120)
>
> 'We always address higher-class and distinguished people we have particular respect for with *you* instead of *thou*, *they* instead of *he*, for example: *Father! You said you would send me to school. Mother is preparing to receive guests, as my uncle is coming to see us.* Specifically, instead of second sg., we use second pl. when they are present; when they are not here, we use third pl. To use third pl. to address someone who is present is a Germanism and usual to Germans. In sermons and lofty speech, the use of second pl. and second sg. is the convention'.

In his correspondence with fellow Slovenian linguists Kopitar (5) and Janez Nepomuk Primic (6), Vodnik used the recommended second pl. address, but he used the third pl. address in his 1813 letter to the son of the nobleman Anton Rudež, Jožef (7).

14 The example in this sentence illustrates not the third pl. address but the polite plural reference to third persons which was typical for Slovenian in the nineteenth century, as well as for a number of Slavic dialects and non-standard varieties in a more or less continuous area from Slovenian, across Kajkavian Croatian, Burgenland Croatian, Moravian and Silesian Czech, Slovak, southern and eastern Polish, to western and central Ukrainian; its use in the literay languages, at least to a certain degree, has been limited to Slovenian, Slovak and perhaps Ukrainian (Houtzagers 2018: 22). Vodnik and other Slovenian grammarians generally did not distinguish between the third pl. address and the polite plural reference to third persons and often used examples of both in their discussions of *onikanje*.

(5) Goſpod Kopitar! Velika nadloga mi je s' grammatiko isrozhéna, vuner dodélana je, ras'un mojih korenin inu **vaſhe**$_{\text{PRON-POSS-2PL-GEN.SG}}$ table; savol ktire ſim v sadrégi, ki **vaſhih**$_{\text{PRON-POSS-2PL-GEN.PL}}$ miſli ne vém. (Vodnik 1988: 279–80)[15]

'Mister Kopitar! A great task has been given to me with this grammar, but it is finished, except my roots and your table, because of which I am in a dilemma, not knowing your thoughts'.

(6) R. je tkoj obduhal, da **ste**$_{\text{AUX-2PL}}$ pervo polo mashnih bukev **brali**$_{\text{PART-PL}}$. (Vodnik 1988: 289)[16]

'R. immediately sniffed out that you have read the first folio of the missal'.

(7) Naj mi **naberejo**$_{\text{PRS-3PL}}$, kar **morejo**$_{\text{PRS-3PL}}$; jeſ sdaj ſnópje ſkladam v' koſlìz: kar ſhe ne bom védel, **jih**$_{\text{PRON-3PL-ACC}}$ bom po nemſhkih iménih zhes nekaj kédnov praſhal [...] Krüniza **bodo**$_{\text{AUX-3PL}}$ **iméli**$_{\text{PART-PL}}$, kakor **shelío**$_{\text{PRS-3PL}}$. (Vodnik 1988: 289–90)

'You should collect what you can for me; I am now weaving the threads together: what I will not know yet, I will ask you about the German names in a few weeks [...] You shall have Krüniz, as you wish'.

Vacillation between the use of second pl. and third pl. in correspondence can also be found in the letters Vodnik himself received from the same correspondent, Janez Nepomuk Primic. In a preserved letter from 3 September 1808 (8), Primic addressed Vodnik with the second pl., but in letters from 4 September 1810 (9) and 13 November 1810 (10), he used third pl.

(8) Viſoku zhaſtitlivi Goſpod Uzhenik! Kér ſe mi je saj lépa perloshnoſz naklúzila, hozhem **Vam**$_{\text{PRON-2PL-DAT}}$ tedaj en malu v' ſlavénskim jésiku piſati [...] Kakú ſe **Vam**$_{\text{PRON-2PL-DAT}}$ kaj godí? Upam, de **ſte**$_{\text{AUX-2PL}}$ sméram **sdrávi**$_{\text{ADJ-PL}}$. (Kidrič 1934: 15)

'Highly esteemed scholar! Because I have been presented with a good opportunity, I shall write to you a few words in the Slovenian language [...] How are you doing? I hope you are still well'.

(9) Viſòko Zhaſtitlivi Goſpod Valant! Vupam, de **ſo**$_{\text{AUX-3PL}}$ moje piſmize p. m. ſkus G. Prof. Sicherl gotòvo **prejéli**$_{\text{PART-PL}}$ [...] letó **prejmejo**$_{\text{PRS-3PL}}$ ſkus G. Kriviza dosdajſhniga ſluſhavza pravíze; **Njim**$_{\text{PRON-3PL-DAT}}$ ga prav dobro perporozhim. (Kidrič 1934: 38)

'Highly esteemed Mister Valant! I hope you have received my letter through Prof Sicherl [...] you receive through Mr Krivic the rights of the erstwhile listener; I recommend him highly to you'.

15 The letter was written between 7 and 9 December 1808.

16 The original letter has been lost, but the quote is preserved in Primic's letter to J. Zupan from 11 February 1813.

(10) Viſoko Zhastítlivi Goſpod! [...] Kadar **bodo**$_{\text{AUX-3PL}}$ li kólkaj zhaſa imeli, **Njih**$_{\text{PRON-3PL-ACC}}$ prósim mi poménik notri poſlanih némſhkih beſedi poſlati [...] Naj **sdravi**$_{\text{ADJ-PL}}$ **bodo**$_{\text{AUX-3PL}}$ [sic!] ino naj ná-me ne **posábijo**$_{\text{PRS-3PL}}$. (Kidrič 1934: 41–2)

'Highly esteemed sir! [...] When you have some time, I ask you to send me the meanings of the German words enclosed [...] May you be well and may you not forget me'.

It should be noted that the first letter, the content of which reveals that it represents one of his first attempts at writing letters to other intellectuals in Slovenian and which adheres to the norm set by Kopitar, was written before the publication of Kopitar's grammar (though it is possible that Primic read the text before its publication). The German norm of address, which he was more used to, prevailed in the second and third letters, written after the publication of Kopitar's grammar, but before the publication of Vodnik's own grammar.[17]

In Vodnik's literary works, polite address is seldom found. In the majority of cases, the second pl. is used, e.g. in his short anecdote published in 1798 (11).

(11) Ena dekla je kašo na mizo nesla, inu čez prag padla de so se čepíne inu kaša po hiši razletęle. Gospodina pravi: tako znam jez tudi! – Dękla odgovori: lahko **znate**$_{\text{PRS-2PL}}$, k' sim **vam**$_{\text{PRON-2PL-DAT}}$ pokazala. (Vodnik 1988: 222)

'A maid carried the porridge to the table, and she fell over the threshold, and the shards and the porridge flew across the house. The mistress says: I could do that myself. The maid answers: Of course, you can do it now, after I have showed it to you'.

When he adapted Linhart's *Ta veseli dan ali Matiček se ženi* for publication in Vincent Franul de Weißenthurn's grammar *Saggio grammaticale italiano-cragnolino* (1811), Vodnik, despite his general advocation of second pl. address, retained all of Linhart's various forms of address. They were adapted to the new concept of Slovenian polite address only thirty years later, when Andrej Smole and France Prešeren published the new adaptation of the play (see Table 5.1).[18]

17 In a German letter from one year earlier (1809), he addresses Vodnik with third pl.: '*Ich bitte um Erlaubniß,* **Sie** *auf einige Augenblicke in Ihren Geschäften zu unterbrechen*' ('I beg permission to interrupt you in your business for a few moments'; Kidrič 1934: 17).

18 Smole and Prešeren altered the language of the work significantly and adapted it to the contemporary language norm; in the purist spirit of the time, they

Table 5.1 Comparison of address in the 1790, 1811 and 1840 editions of the comedy *Ta veseli dan ali Matiček se ženi* by Anton Tomaž Linhart.

	1790 edition	Vodnik's adaptation (1811)	1840 edition
Nežka to the baron	Vaſha gnada, kaj **sapovedó**$_{\text{PRS-3PL}}$? kè bi **jih**$_{\text{PRON-3PL-ACC}}$ kdu per meni najdel? (p. 15) ('Your Lordship, what do you command? Imagine someone found you with me!')	Vaſha gnada, kaj **sapovedó**$_{\text{PRS-3PL}}$? kè bi **jih**$_{\text{PRON-3PL-ACC}}$ kdu per meni najdel? (p. 210)	Kaj **sapovéſte**$_{\text{PRS-2PL}}$, miloſljivi goſpod? Kaj bi bilo, ko bi **vaſ**$_{\text{PRON-2PL-ACC.}}$ kdo per mêni najdel (p. 16)? ('What do you command, gracious lord? What would happen if someone found you with me?')
Baron to Zmešnava	**Oní**$_{\text{PRON-3PL-NOM}}$ naj **ſtoję**$_{\text{PRS-3PL}}$ pred to utizo, inu naj obénga vùn ne **pustę**$_{\text{PRS-3PL}}$. (p. 140) ('You should stand before this shed and not let anyone out'.)	**Oni**$_{\text{PRON-3PL-NOM}}$ naj **stojé**$_{\text{PRS-3PL}}$ pred to utizo, ino naj nobêniga vùn ne **puſté**$_{\text{PRS-3PL}}$. (p. 294)	Pred tém ſhotoram **obſtojte**$_{\text{IMP-2PL}}$, ino ne **puſtite**$_{\text{IMP-2PL}}$ nikogar vùn. (p. 134) ('Stand before this tent and do not let anyone out'.)
Gašper to Matiček	Ho, ho, Matizhek! – tok **ſte**$_{\text{AUX-2PL}}$ ta zhas grosnu **sraſil**$_{\text{PART-SG}}$; **ſte**$_{\text{AUX-2PL}}$ **bil**$_{\text{PART-SG}}$ vſe bol **majhèn**$_{\text{ADJ-SG.}}$ inu **tènàk**$_{\text{ADJ-SG}}$ vidit. (p. 65) ('Ho, ho, Matiček! – You have grown terribly in the meantime; you seemed much shorter and slimmer'.)	Ho, ho, Matizhik! – tak **ste**$_{\text{AUX-2PL}}$ ta zhas grosno **sraſil**$_{\text{PART-SG}}$; **ste**$_{\text{AUX-2PL PL.}}$ **bil**$_{\text{AUX-2PL}}$ vse bol **majhèn**$_{\text{ADJ-SG}}$ inu **tènak**$_{\text{ADJ-SG}}$ vidit. (p. 244)	[...] grôsno **ſte**$_{\text{AUX-2PL}}$ **sraſtli**$_{\text{PART-PL}}$ ta zhàs; **manjſhi**$_{\text{ADJ-PL}}$ ino **ſhibkeji**$_{\text{ADJ-PL}}$ **ſte**$_{\text{AUX-2PL}}$ ſe mi **sdéli**$_{\text{PART-PL}}$. (p. 63) ('[...] you have grown terribly in the meantime; you seemed much shorter and slimmer to me'.)

3.3 Leopold Šmigoc

Only a year after the publication of Vodnik's grammar, the first grammar of the East Styrian variant of the Slovenian literary language[19] was published, written by government official Leopold Šmigoc and primarily aimed at Germans who wished to learn Slovenian, as evidenced by numerous exercises, sample conversations and a short dictionary.

In remarks in the chapter on personal pronouns (12), Šmigoc presented a new criterion for distinguishing between the use of second pl.

removed numerous Germanisms in particular. In addition to various lexical Germanisms (for example, they changed the honorific *vaša gnada* to *milostljivi gospod/milostljiva gospa* 'gracious lord/lady'), condemnation also befell the third pl. address, which, in line with grammatical recommendations, was consistently replaced by the second pl. address. Also, semantic agreement in the second pl. address was changed to syntactic agreement with the adjectives and participles in plural. Such an intervention resulted in less hierarchisation between the characters compared to the original: the linguistically expressed relationship between, for example, Žužek and Matiček was now the same as the one between Matiček and Gašper, while Budalo's position, in terms of language, became the same as the positions of Žužek and Zmešnava, as the bipartite pronominal system of address could no longer express the precise hierarchic stratification of the characters. In the comedy, this also curtailed both the sociolectal identification of a speaker through the expression of language as the reflection of the social and psychological response to a given situation and the authenticity or verism of the events (Orel 2006: 579).

19 The East Styrian literary language began to develop in the middle of the eighteenth century as a result of increased differentiation in Slovenian dialects. It was a supradialectal formation that existed until the middle of the nineteenth century, when a compromise standard was accepted, incorporating some of the East Styrian features into the new common Slovenian standard (cf. Jesenšek 2009). In terms of pronouns for polite address, the East Styrian literary language did not differ from the central standard.

and third pl., which can be equated with Brown's and Gilman's (1960: 257–8) notion of solidarity.[20]

(12) Die Slowenen machen einen Unterſchied zwiſchen den[sic!] *vi* ihr, Sie, und den[sic!] *oní* ſie. Wenn ſie mit einer Perſon reden, welcher ſie zwar Ehrfurcht erweiſen, aber doch mit ihr vertraut ſind, ſo brauchen ſie das *vi*, und ſetzen dann alles, was auf dieſes *vi* Bezug hat, in den plur. des männlichen Geſchlechtes, als z. B. *kéj* ***ste***$_{\text{AUX-2PL}}$ ***bili***$_{\text{PART-PL}}$ *ozha*? wo seid ihr gewesen Vater? das *oní* aber brauchen ſie, wenn ſie mit Perſonen reden, denen ſie auch Erfurcht ſchuldig ſind, aber keinen vertrauten Umgang mit ihnen pflegen, beſonders bey Perſonen höheren Standes, z. B. *Góſpod shlahten proſim, naj* ***vſliſhijo***$_{\text{PRS-3PL}}$ *mòjo proſhnjo*, gestrenger Herr, ich bitte Sie, erhören Sie meine Bitte. (Šmigoc 1812: 59)

'Slovenians distinguish between *you* (*ihr*, *Sie*) and *they* (*sie*). When speaking to a person whom they respect but still have a confidential relationship with, they use *you* and everything associated with that person in the masculine plural, e.g. *where have you been, father? They* is used when speaking with persons to whom they also owe respect, but with whom they do not have a confidential relationship, especially with persons of a higher status, e.g. *Noble sir, I ask you to hear my request'*.

Despite having described the functional differences between the two address pronouns, Šmigoc warned at the end that the use of third pl. was not proper Slovenian (13).

(13) Dieſe Art des Ausdruckes iſt nicht echt Sloweniſch, ſie hat ſich durch die Länge der Zeit aus der Deutſchen in unſere Sprache eingeſchlichen. (Šmigoc 1812: 59)

'This style of expression is not authentically Slovenian but has crept into Slovenian in the past from German'.

Such a warning could be attributed to the influence of Kopitar and Vodnik's grammars, as Šmigoc's sample conversations in the grammar show that he had no compunction about using third pl., especially as they

20 Brown and Gilman define solidarity as symmetrical relationship that can be produced by frequency of contact as well as by objective similarities, which create a 'like-mindedness that seems to be the core of the solidarity semantic' (Brown & Gilman 1960: 257).

were intended for foreign language speakers of the bourgeois class. The third pl. was used not only for the conversations between 'gentlemen' (14), but also for those between a shopkeeper and his customer (15), while second pl. was used only for addressing domestic help (16), in accordance with the description of town usage in Kopitar's grammar and in contradiction with Šmigoc's own rules for using the polite pronouns.

(14) Kak ſe **imájo**$_{\text{PRS}}$? – Kak **njim**$_{\text{PRON-3PL-DAT}}$ gré? (Wie befinden sie sich?)
Kak pa **njim**$_{\text{PRON-3PL-DAT.}}$ gré? – Kak pa ſe **oni**$_{\text{PRON-3PL-NOM}}$ **imájo**$_{\text{PRS-3PL}}$? (Und wie geht es Ihnen mein Herr?)
Dober vezhér, perjatel. Kam tak hitro? – Kam tak hitro **gréjo**$_{\text{PRS-3PL}}$? – Kam je tak hitro **njihov**$_{\text{PRON-POSS-3PL-NOM.SG}}$ pót? (Guten Abend mein Freund. Wo gehen Sie so geschwind hin?) (Šmigoc 1812: 195)
'How are you?'
'And how are you, dear sir?'
'Good evening, my friend. Where are you going in such a hurry?'
(15) a. Góſpod mèni je blo povédano, de **oni**$_{\text{PRON-3PL-NOM}}$ vſake ſorte ſukno **imájo**$_{\text{PRS-3PL}}$ – Góſpod jas ſem ſliſhil, de ſe per **njih**$_{\text{PRON-3PL-LOC}}$ vſake ſorte ſukno odáva. (Šmigoc 1812: 226)
'Sir, I have been told that you have all kinds of cloth. – Sir, I have heard all sorts of cloth are sold by you'
b. Naj ſi ſami **svólijo**$_{\text{PRS-3PL}}$, katéro ſe **njim**$_{\text{PRON-3PL-DAT}}$ nar bole dopáde. (Šmigoc 1812: 227)
'You should decide for yourself which is the most agreeable to you'
(16) Kaj **sapovéjo**$_{\text{PRS-3PL}}$? – Kaj **shelijo**$_{\text{PRS-3PL}}$? (Was befehlen Sie.)
She sa eniga **pokríte**$_{\text{IMP-2PL}}$, ino kuharzi **povete**$_{\text{IMP-2PL}}$, de nám jéſti dá. (Legt noch ein Ged[e]ck auf, und ſagt der Köchin, daß sie anrichte.) (Šmigoc 1812: 231)
'What do you command?
– Set the table for another person and tell the cook to serve us'.

On the other hand, in a conversation between a tailor (T) and a customer (C), the latter uses third sg. (17), which was not attested in any of the other written Slovenian sources examined in the course of this research.

(17) *T:* Góſpod mi blo povédano, de s' mènoj govoriti **ozhejo**$_{\text{PRS-3PL}}$. (Mein Herr man hat mir gesagt, Sie verlangten mich zu sprechen.)

> *C:* **Je**$_{\text{PRS-3SG}}$ leta góſpod ſambol, od katériga ſo góſpod J. s' mènoj govorili. (Ist das der Herr Kleidermacher, von welchem der Herr N. mit mir geſprochen hat.)
> *T:* Ja, góſpod, **ſem**$_{\text{PRS-1SG}}$ taiſti. (Ja, mein Herr zu dienen.)
> *C:* **Ozhe**$_{\text{PRS-3SG}}$ meni méro vséti na eno opravílo. (Will er mir das Maß zu meinem Kleide nehmen.) (Šmigoc 1812: 229)
> ‘*T:* Sir, I have been told that you wish to speak with me.
> *C:* Are you the tailor Mr J. spoke to me about?
> *T:* Yes, sir, I am the one.
> *C:* Would you take my measurements for a suit?’

This example is a complete transfer of the German mode of address using *Er* (‘he’), the status of which also declined in this period relative to the second pl. address.

3.4 Peter Dajnko

In another grammar of the Eastern Styrian variant of literary Slovenian, *Lehrbuch der Windischen Sprache* (1824), its author Peter Dajnko also emphasised that the only authentically Slovenian form of address was the second pl. pronoun, but that Slovenians who are in contact with the Germans would use of third pl. through their influence (18, 19).

> (18) Der reine Slowene ſpricht jenen, dem er eine Ehrfurcht oder Titulatur bezeigen will, mit *vi* ihr, an, z. B. ***Vi***$_{\text{PRON-2PL-NOM}}$ *oчa! kaj pà* ***vi***$_{\text{PRON-2PL-NOM}}$ *na to* ***reчete***$_{\text{PRS-2PL}}$*?* Ihr Vater! was sagt denn ihr dazu? (Dajnko 1824: 162)
> ‘The true Slovenian addresses everyone to whom he wishes to pay respect or wishes to address with his title with *you* (ihr), e.g. *You, Father, what do you say to that?*’
> (19) Wo aber der Slawe mit Deutſchen zusammenſtoßt, macht er es nach Deutſchem Sprachgebrauche, z. B. *Moj dragi gospod! kaj mi* ***velíjo***$_{\text{PRS-3PL}}$*?* Mein theurer Herr! was befehlen Sie mir? (Dajnko 1824: 163)
> ‘Where the Slavs are in contact with the Germans, they behave according to German-language usage, e.g. *My dear Sir, what do you order?*’

In the texts published in the grammar, he predominantly used second pl. for polite address (20).

(20) En gospód so svojega hlapca za nemarnosti volo jako bili. Hlapec je zaчél zvati: Zakaj pà mè **bíete**$_{\text{PRS-2PL}}$, se sem ne niч vчinil. Ravno za tega volo, so odgovorili gospod, **tè**$_{\text{PRON-2SG-ACC}}$ bíjem, da **si**$_{\text{AUX-2SG}}$ niч ne **vчinil**$_{\text{PART-SG}}$. (Dajnko 1824: 318)
'A gentleman beat his servant severely for negligence. The servant began to ask: Why do you beat me, when I have done nothing? That is why I beat thee, answered the gentleman, because thou have done nothing'.

The only aberration is the addressing of an innkeeper (I) with third pl., while he addresses his guest (G) with second pl (21).

(21) I. Vam$_{\text{PRON-2PL-DAT}}$ smem staro vino natoчiti, ali te$_{\text{AUX-2PL}}$ novo pili$_{\text{PART-PL}}$
G. Po чim kerчmarijo$_{\text{PRS-3PL}}$ staro? [...]
G. Naj mì prinesejo$_{\text{PRS-3PL}}$ tiчas masel novega, poznej pà poliч starega.
I. Xe dobro. Te$_{\text{AUX-2PL}}$ tydi kaj jeli$_{\text{PART-PL}}$?
G. Kaj majo$_{\text{PRS-3PL}}$ za obed? (Dajnko 1824: 303)
'I. May I pour you the old wine, or will you drink the new one?'
G. How much do you want for the old one? [...]
Bring me a pint of the new one, and afterwards a bottle of the old one.
I. All right. Will you also have anything to eat?
G. What do you offer?'

3.5 Anton Murko

The most detailed grammatical description of the Slovenian address system in the first half of the nineteenth century can be found in two editions of Anton Murko's grammar, the first published in 1832 (*Theoretisch-praktische Slowenische Sprachlehre für Deutsche*), and the second in 1843 (*Theoretisch-practische Grammatik der slowenischen Sprache in Steiermark, Kärnten, Krain und dem illyrischen Küstenlande*). Murko described similar differences in the uses of second pl. and third pl. as Šmigoc (see 3.2), differentiating them according to the degree of solidarity between interlocutors, but adding that the use of third pl. is more frequent in towns and that even there second pl. was sometimes used (22).

> (22) Zwiſchen den Höflichkeitspartikeln *vi* ihr, und *óni* oder *oní* ſie, findet ein Unterſchied Statt, der nicht außer Acht zu laſſen iſt. Leute ſeines Gleichen, als Vater, Mutter und andere ſeines Standes, denen er jedoch Hochachtung und inniges Vetrauen zugleich ſchuldig iſt, redet der Slowene durchgehends mit *vi* ihr, an; ſpricht er hingegen von vornehmern Perſonen, als weltlichen und geiſtlichen Obrigkeiten, oder überhaupt von ſolchen, die er ihres Standes, Ranges, oder einer andern rühmlichen Eigenſchaft wegen achtet, aber keinen vertrauten Umgang mit ihnen pflegt, ſo iſt, beſonders in den Städten und deren Umgebungen, das *óni* oder *oní* üblicher, obwohl auch in dieſem Falle das *vi* nicht ganz außer Gebrauche. (Murko 1832: 44–5)
>
> 'There is a difference between the polite pronouns *you* (*ihr*) and *they* (*Sie*) that must be observed. A Slovenian always addresses persons of similar social standing, such as father, mother and other persons of the same rank, to whom he is due respect and whom he is on confidential terms with *you*; but when he speaks to more distinguished persons, such as gentry and clergy, or in general with those with whom it is necessary to consider their position, rank or some other honourable characteristic, but with whom he does not have confidential relations, especially in cities and their surroundings, *they* is more common, although, even in these cases, *you* is not completely out of use'.

He also denounced the use of third pl. as non-Slavic (23).

> (23) Der ächte, den Geiſt ſeiner Sprache kennende, und bewahrende, Slawe kennt das *òni* als Höflichkeitspartikel nicht, ſonder ſpricht immer *vi*. (Murko 1832: 45)
>
> 'The genuine Slav, who knows and preserves the spirit of his language, does not know the *òni* as a polite particle, but always uses *vi*'.

In the second edition, Murko devoted even more attention to the topic of address. He described the use of second sg., second pl. and third pl. Although he specifically marked the use of third pl. as a Germanism, he no longer advised against it (24).

> (24) *Oni* Sie, iſt eigentlich nicht slawiſch, ſondern das überſetzte deutſche Sie, und wird, beſonders in der Stadt und deren Umgebungen, gebraucht, wenn man vornehmere Perſonen anredet, denen man wegen ihres Standes oder Ranges zwar Ehrfurcht ſchuldig, aber mit ihnen nicht zugleich vertraut iſt; aber auch in dieſem

> Falle iſt das ächte ſloweniſche *vi* Ihr nicht ganz außer Gebrauch. (Murko 1843: 58–9)
>
> ‘*They* (*Sie*) is actually not Slavic but the translated German *Sie*, and is used, especially in the city and its surroundings, when addressing noble people, whom one owes respect because of their status or rank, but with whom one does not have confidential relations; but even in this case, the real Slovenian *you* (*Ihr*) is not completely out of use’.

The reason for this could be the fact that, in bourgeois circles, the use of second pl. was still considered impolite, as Murko himself noted (25):

> (25) Zu merken aber iſt, daß der Unterſchied zwiſchen *ví* und *oní* erſt in neueſter Zeit im Sloweniſchen feſtzuſetzen ſich angefangen hat. Einſt redete der Slowene in allen Fällen, wo nach den obigen Beſtimmungen das *ti* nicht paſſend iſt, Jedermann mit *ví* an, wie der Franzoſe ſein *vous*, der Engländer ſein *you* ohne Unterſchied des Standes der angeredeten Perſon gebraucht. Aber jetzt würde es, beſonders in der Stadt und deren Umgebungen, für Grobheit gelten, auch für die höfliche Anrede ſich des ächt ſloweniſchen *ví* zu bedienen. (Murko 1843: 60)
>
> ‘It should be noted, however, that the difference between *you* and *they* has only recently begun to be established in Slovenian. Once upon a time, the Slovenian addressed everyone with *you* in all cases where the *thou* is not appropriate according to the above provision, as the French use their *vous*, the English use their *you* regardless of the status of the person addressed. But now, especially in the town and its surroundings, it would be considered rude to use the authentic Slovenian *you* for polite address’.

Therefore, Murko was still forced to predominantly use the third pl. address in his sample dialogues in both editions (28). In the first edition, only in the first dialogue, containing general greetings and polite phrases, both variants were consistently used (26), but some variants appear also in other dialogues (27).

(26) Dòbro jútro **njìm**$_{\text{PRON-3PL-DAT}}$ (**vàm**$_{\text{PRON-2PL-DAT}}$) Bóg daj. (Murko 1832: 170)
‘God grant you a good morning’.

(27) **So**$_{\text{AUX-3PL}}$ ſe kaj dòbro **iméli**$_{\text{PART-PL}}$? – **Ste**$_{\text{AUX-2PL}}$ kaj dòbre vólje **bilí**$_{\text{PART-PL}}$?
‘Did you have fun? – Were you in a good mood?’ (Murko 1832: 174)

(28) **Njìh**$_{\text{PRON-3PL-GEN}}$ dobrotnoſt me je mozhnó geníla. (Murko 1832: 173)
‘Your goodness moved me’.

In the 1843 edition, the situation is significantly different. Although in the theoretical part he emphasised the necessity of using 'Slavic' *vi* less, he strengthened it in the practical part. At the beginning, he warned in a note that 'for the sake of practice' (Murko 1843: 227) he would alternately use *oni*, *vi* and *ti* in sample conversations and stated, regarding their use, that the user should follow the rules in the grammar section. Therefore, in the cases where the interlocutors were social equals, in twelve dialogues third pl. was used, in ten dialogues second pl. was used and in five dialogues both variants were provided (29, 30).

(29) Ali mi ne **morejo**$_{\text{PRS-3PL}}$ povédati, kde zdravnik J. prebiva? (Könnten Sie mir gefälligst sagen, wo der Arzt R. wohnt?)
Ostánite$_{\text{IMP-2PL}}$, le **ostánite**$_{\text{IMP-2PL}}$. – Naj le **ostánejo**$_{\text{PRS-3PL}}$. (Bleiben Sie sitzen.)
Ali **ste**$_{\text{AUX-2PL}}$ se splòh dobro **počútili**$_{\text{PART-PL}}$? (Haben Sie sich immer gut befunden?) (Murko 1843: 229)
'Can't you tell me where Dr R. lives?'
'Stay, please stay – Remain seated'.
'Have you been well in general?'
(30) Nekaj še **ostanite**$_{\text{IMP-2PL}}$
Večkrat me **počástite**$_{\text{IMP-2PL}}$. – Večkrat me naj **počastijo**$_{\text{PRS-3PL}}$.
(Murko 1843: 230).
'Stay for a while longer'.
'Honour me more often – You should honour me more often'.

But in the dialogues when the interlocutors were a master and a servant (31), the latter always used the third pl. address.

(31) M: Britve ne réžejo, jih **moraš**$_{\text{PRS-3SG}}$ še enkrat nabrusit dati.
S: Ali **bodo**$_{\text{AUX-3PL}}$ dans bélo súknjo **oblékli**$_{\text{PART-PL}}$? [...] Žlahtni gospód, naj ne **pozábijo**$_{\text{PRS-3PL}}$, de **so**$_{\text{AUX-3PL}}$ dnés k' J. J. povábleni. (Murko 1843: 247)
'The razors do not cut; you have to have them sharpened'.
'Will you wear the white jacket today? [...] Noble sir you should not forget that you have been invited to go to J. J. today'.

We can conclude that even though the use of second pl. was gaining prestige in the circles of Slovenian intellectuals, which is proven by its use in the first Slovenian tale (*povest*) *Sreča v nesreči* by Janez Cigler (1836) and the already mentioned new edition of Linhart's *Ta veseli dan ali Matiček se ženi* (1840; see Section 2), it was still not universally accepted in bourgeois circles in the middle of the nineteenth century.

3.6 Others

A new impetus for the creation of Slovenian grammars was given by the March revolution of 1848, during which Slovenians demanded a new administrative unit in the Empire, which would unite all Slovenians in a united Slovenia, in which all public life in schools, offices and courts would be conducted in their language (cf. Melik 2002: 38–43). Because their German and Germanophile opponents[21] challenged Slovenians that the Slovenian language was not capable of taking on such functions,[22] quite a few grammars were published during this period, which had little

21 In the nineteenth century, the Austro-Hungarian Empire sought to assimilate members of non-German and non-Magyar ethnic groups. Accordingly, they hindered their development in the field of culture, education and the use of minority languages in all areas of public life (Klemenčič 2017: 131).

22 An example of such an attitude towards Slovenian and other Slavic languages can be found in a quote by the Carinthian politician and poet Adolf Ignaz von Tschabuschnigg: 'Gäb' es nur eine Sprache unter den Menschen wer weiß ob sie nicht die der Götter wäre: mit den slavischen aber lassen sie mich jedenfalls verschont. Ihre Laute zerreissen mein Ohr, sie tönen wie Klagen aus dem Tartarus, wie der Iammer des Gedankens, der hinter den rauhen Eisenstäben unzähliger Konsonanten verschmachtet. Wenn jener Beweis des Slavomanen wahr wäre, Adam habe im Paradiese slavisch gesprochen, so that er dies gewiß nur mit der Schlange. Was soll uns eine Sprache die keine Literatur hat[?]' ('If there were only one language among men, who knows if it wouldn't be that of the gods: but they'll spare me the Slavic ones anyway. Their sounds tear my ears apart, they sound like lamentations from Tartarus, like the lamentation of thought that languishes behind the rough iron bars of innumerable consonants. If that claim of the Slavophiles were true, that Adam spoke Slavic in Paradise, he certainly did so only with the serpent. What's the use of a language that has no literature?'; Tschabuschnigg 1846, quoted in Hladnik 2019: 290).

scientific value, but were written with the aim of proving to their opponents that Slovenian was suitable for use in new situations. The majority of them also dealt with the topic of polite address and, in the spirit of the period, their puristic attitude was generally strengthened.

For example, Matija Majar Ziljski in his *Pravila kako izobraževati ilirsko narečje i u obče slavenski jezik* (1848) wrote: 'Po slavensko se ali tika ali vika; – onikati je neslavensko i gnjusno' ('In Slavic one can use *thou* or *you* – to use *they* is non-Slavic and abhorrent'; Majar 1848: 89). In *Slovnica za Slovence* (1850), he advised teachers regarding their students: 'Navadi jih víkati in tíkati in povej jim, da je onikati gerdo in němškutarsko' ('Teach them to use *thou* and *you*, and tell them that to use they is ugly and 'Deutschthümlerisch';[23] Majar 1850: 13). In his first grammar, he also advised on how to finish the letter in true Slavic manner (32):

> (32) Lěpo se **Vam**$_{\text{PRON-2PL-DAT}}$, žlahtni gospodine! priporučim. Bog daj **Vam**$_{\text{PRON-2PL-DAT}}$ sreću i ljubo zdravje! (Majar 1850: 89).
>
> 'I recommend myself to you, gentle sir! God give you happiness and health!'

Similarly, Fran Malavašič (*Slovenska slovnica za prve slovenske šole v mestih in na deželah*, 1849) mentioned only the use of second pl. in his grammar, while Ivan Navratil (*Kurze Sprachlehre mit einer möglichst vollständigen Rechtschreibung der slovenische Sprache*, 1850) mentioned the use of third pl. and marked it as a Germanism (33).

> (33) Anredewörter haben die Slowenen eigentlich, gleich den übrigen Slaven, nur 2: *ti* du und *Vi* Ihr. (*Oni* iſt dem deutſchen Sie nachgebildet worden.) (Navratil 1850: 27)
>
> 'Slovenians, similar to other Slavs, have only two forms of address: *thou* and *you*. (*They* is modelled after the German *Sie*.)'

Anton Janežič in his *Kurzer leichtfasslicher Unterricht in der Slovenischen Sprache: für Deutsche* (1849), which was aimed at Germans, mentioned only the second pl. address (34), despite his claim that he included the

23 *Deutschthümlerisch*, Slovenian *nemškutarsko*, pertains to a person of non-German origin who advocates Germanisation (SSKJ2, <www.fran.si>).

most general and useful features of the grammar, so that the grammar would not lose its usefulness due to an excessive tendency towards purity (Janežič 1849: 29).

> (34) Man unterſcheidet nur zwei Anredewörter, nämlich; *ti* (Du) und *Vi* (Ihr, Sie). *Ti* hört man nur in der Sprache der Vertraulichkeit im Tone der Herrſchaft, oder auch der Verachtung. – In allen übrigen Fällen wird *Vi* gebraucht. (Janežič 1849: 32)
>
> ‘Only two forms of address are used: *thou* (*Du*) and *you* (*Ihr*, *Sie*). *Thou* is heard only in confidential conversations or in authoritative or pejorative tone, in all other cases *you* is used’.

That he was aware that the third pl. address was still present in Slovenian, is proven by Janežič’s Slovenian grammar, published five years later, *Slovenska slovnica s kratkim pregledom slovenskega slovstva* (1854), in which he warned his Slovenian students against its use (35).

> (35) Kader Slovenec z ljudmi govori, kterim je spoštovanje dolžen, jih vika; kader pa od njih govori, jih pogosto onika [...] Ljubi oče? Kam se **bote**$_{\text{AUX-2PL}}$ jutre **podali**$_{\text{PART-PL}}$? Mati! zakaj **ste**$_{\text{AUX-2PL}}$ **žalostni**$_{\text{ADJ-PL}}$? Osebe, s kterimi govorimo onikati ni slovensko. (Janežič 1854: 105)
>
> ‘When a Slovenian talks to people to whom he owes respect, he uses second pl. to address them; when he talks about them, he often uses third pl. [...] *Dear Father, where will you go tomorrow? Mother, why are you sad?* To use third pl. for someone when speaking to them is not Slovenian’.

The use of the third pl. address in the middle of the nineteenth century is also attested in Blaž Potočnik’s *Grammatik der slowenischen Sprache* (1849). Unlike other grammarians of the period, the author did not evaluate it, but merely pointed out its German origin and its general acceptability among speakers (36):

> (36) Auch die deutſche Art, höhere Perſonen mit Sie anzureden, hat ziemlich allgemein Eingang gefunden; z. B. *gospa, ali* ***bodo***$_{\text{AUX-3PL}}$ ***prishli***$_{\text{PART-PL}}$ *kaj k nam?* Frau, werden Sie wohl zu uns kommen? (Potočnik 1849: 84)
>
> ‘The German manner of address of people of higher rank with *they* is generally accepted, e.g. *Lady, will you come to us?*’

His statement confirmed the conclusion based on an earlier similar perception by Anton Murko that, despite all the efforts of the Slovenian purists and advice in various grammars of the first half of the nineteenth century, they were unable to eradicate the use of third pl. in Slovenian. This is corroborated by some examples of letters in Malavašič's grammar: while he generally used the second pl. address[24] for addressees of various social standings, even for a nobleman (37), exclusively third pl. is used in a sample letter to a priest (38).

(37) Kolar ponudi grajšaku svoje délo. Milostljivi gospod! Po smerti tukajšniga kolarskiga mojstra, ki je **Vam**$_{\text{PRON-2PL-DAT}}$ delati čast imel, se je v meni želja zbudila, zdaj na njegovo mesto stopiti in si s svojim natanjčnim délam **Vaše**$_{\text{PRON-POSS-2PL-ACC.SG}}$ zaupanje pridobiti. (Malavašič 1849: 160)

'Wheelwright offers the owner of the manor his work. Gracious Sir! After the death of the local wheelwright, who had the honour to work for you, a wish awoke in me to take his place and earn your trust with my precise work'.

(38) Častitljivi gospod fajmošter nej **bi**$_{\text{AUX}}$ mojo prošnjo milostljivo **uslišali**$_{\text{PART-PL}}$ in kadar bo mogoče, moja dva sina brez plačevanja v šolo **vzéli**$_{\text{PART-PL}}$. Častitljivi gospod, ki **so**$_{\text{AUX-3PL}}$ znan dobrotnik mladosti, ki v svoji fari milostivo in polni ljubezni ubogim soseske pomoč **dajo**$_{\text{PRS-3PL}}$ in vse dobro po moči **podperajo**$_{\text{PRS-3PL}}$, **bodo**$_{\text{AUX-3PL}}$ tim bolj prošnjo očeta **uslišali**$_{\text{PART-PL}}$, ki bi z svojih sinov rad koristne in verle možé zredil. (Malavašič 1849: 142)

'Honourable Pastor, graciously hear my request and, when possible, take my two sons in school for free. Honourable Sir, you who are known as a benefactor of the youth, who in his parish graciously and full of love offer help to the poor of the neighbourhood and support all that is well, hear the request of a father who wishes to raise his sons to become useful and virtuous men'.

This type of usage was preserved well into the twentieth century, until the Second World War and even beyond that in the countryside, and

24 Occasional third pl. can be found in some letters: 'Častiti gospod! Ker **STE**$_{\text{AUX-2PL}}$ mi mojo prošnjo zastran posodbe denarjev tako urno **SPOLNILI**$_{\text{PART-PL}}$, **STE**$_{\text{AUX-2PL}}$ me skerbi **REŠILI**$_{\text{PART-PL}}$, ki je mojimu sercu velik nepokoj delala. **NJIH**$_{\text{PRON-3PL-GEN}}$ človekoljubnost mi je mogoče storila [...]' ('Honourable Sir. As you have completed my request regarding the loan so promptly, you have saved me from worry which caused much unrest to my heart. Your philanthropy enabled me [...]'; Malavašič 1849: 143).

it also applied mainly to priests. Weiss (2003: 207) cites informants' reports that, in the Nova Štifta parish, some parishioners were still addressing the new priest with third pl. before the Second World War, and the parish manager of the Šmartno ob Dreti parish, during the visit of the Maribor bishop in 1970, gave instructions to his parishioners that they should simply use second pl. and not third pl. when addressing the bishop.

4. Conclusion

After the intervention of Slovenian grammarians in the first half of the nineteenth century, the Slovenian address system – which had until then, due to close contact with German, in general followed the development of the German system – shifted in a typologically unexpected manner (cf. Head 1978: 192), gradually replacing the pronoun expressing a higher degree of respect with a pronoun expressing a lower degree of respect. The inventory of forms of address for expressing social differences in the standard language was thus reduced to the bipartite system (*ti* – *vi*) that exists in modern Slovenian. This was the result of growing anti-German purism, which prompted Slovenian grammarians to advocate against the use of the third pl. address due to it being perceived as a Germanism. Their intervention disturbed the already established address system and caused uncertainty in usage for over half a century, as grammatical and social norms regarding the address of higher classes contradicted each other. Even the grammarians themselves, though condemning *onikanje* in theory, often used it themselves in other texts, especially in letters to persons of a higher rank or in sample conversations in grammars aimed primarily at Germans who wished to learn Slovenian.

By 1850, the third pl. address had mostly been banished from literary works, and a bipartite system of address had been established once again in the literary language, but grammarians still reported that the use of third pl. remained present in the language; in Malavašič's grammar, it was limited to the address of the clergy.

The Slovenian grammars in the first half of the twentieth century usually did not even mention the third pl. address. The exception is *Pregled slovnic slovenskega knjižnega jezika* by Stanko Bunc (1940: 98) where it was still marked as a Germanism. After the Second World War, the third pl. address was reported to have 'passed completely out of use' (Toporišič 1976: 326), and the form has been marked as archaic and its use was largely unknown (Kess & Jurčić 1978: 298, cited in Reindl 2007: 155); however, Reindl (2007: 155) reports that younger informants generally recount that their grandparents and even parents used to speak this way, and that such forms remain in active use by some of the middle generation. I have made similar observations and have heard the form used especially in the older generation's address of the clergy.

In the twentieth century, the role of the third pl. address in Slovenian changed significantly: it was more commonly heard in the countryside than in towns, but it did not belong to any particular dialect, since it was most often reported to be used for the address of priests, for which standard language is usually expected (or attempted). It was either not mentioned in normative manuals (for example, in most of grammars by Anton Breznik) or it was discouraged. Thus, it was a kind of archaic remnant of an older norm that did not belong to either the standard language or a dialect.

Bibliography

Primary sources

Cigler, Janez. 1836. *Şrezha v' nesrezhi, ali Popişvanje zhudne sgodbe dvéh dvojzhikov*. V' Ljubljani: Joshef Blasnik.

Dajnko, Peter. 1824. *Lehrbuch der Windischen Sprache*. Gradec: gedruckt und verlegt bey Johann Andreas Kienreich.

Gutsman, Ožbalt. 1777. *Windische Sprachlehre*. Celovec: Ignaz Aloys Kleinmayer.

Janežič, Anton. 1849. *Kurzer leichtfasslicher Unterricht in der Slovenischen Sprache: für Deutsche: Erster Kursus*. Celovec: J. Sigmund.

Janežič, Anton. 1854. *Slovenska slovnica s kratkim pregledom slovenskega slovstva ter z malim cirliskim in glagoliškim berilom za Slovence*. Celovec: založil Eduard Liegel.

Kidrič, France. 1934. *Korespondenca Janeza Nepomuka Primca 1808–1813*. Ljubljana: Znanstveno društvo.

Kidrič, France. 1939. *Zoisova korespondenca 1808–1809*. Ljubljana: Akademija znanosti in umetnosti.

Kopitar, Jernej. 1808. *Grammatik der Slavischen Sprache in Krain, Kärnten und Steyermark*. Ljubljana: Wilhelm Heinrich Korn.

Linhart, Anton Tomaž. 1790. *Ta veſsęli dan, ali: Matizhek ſe shęni*. <http://www.dlib.si/?URN=URN:NBN:SI:DOC-S2G9Q52R> (accessed 30 November 2024).

Linhart, Anton Tomaž. 1811. *Ta veſsęli dan, ali: Matizhek ſe shęni*. Ed. Valentin Vodnik. In Vincenc Franul Weissenthurn, *Saggio grammaticale italiano-cragnolino*. <http://www.dlib.si/?URN=URN:NBN:SI:DOC-4W1ZQPW2> (accessed 30 November 2024).

Linhart, Anton Tomaž. 1840. *Veséli dan ali Matiček se ženi: komédija v 5 délih*. Ed. Andrej Smole. Ljubljana: J. Blaznik.

Majar, Matija. 1848. *Pravila kako izobraževati ilirsko narečje i u obče slavenski jezik*. Ljubljana: J. Blaznik.

Majar, Matija. 1850. *Slovnica za Slovence*. Ljubljana: [samozal.] M. Majar.

Malavašič, Franc. 1849. *Slovenska slovnica za prve slovenske šole v mestih in na deželah*. Ljubljana: J. Giontini.

Murko, Anton. 1832. *Theoretisch-praktische Slowenische Sprachlehre für Deutsche, nach den Volkssprecharten der Slowenen in Steiermark, Kärnten, Krain und Ungarns westlichen Distrikten*. Gradec: Verlag der Franz Ferstlschen Buchhandlung, Johann Lorenz Greiner.

Murko, Anton. 1843. *Theoretisch-practische Grammatik der Slowenischen Sprache in Steiermark, Kärnten, Krain und dem illyrischen Küstenlande*. Gradec: Verlag der Fr. Ferstl'schen Buchhandlung.

Muršec, Jožef. 1847. *Kratka slovenska slovnica za pervence*. Gradec: Lajkamovi nasledniki.

Navratil, Ivan. 1850. *Kurze Sprachlehre mit einer möglichst vollständigen Rechtschreibung der slovenischen Sprache*. Ljubljana: J. Blasnik.

Pohlin, Marko. 1768. *Kraynska grammatika. das ist: Die crainerische Grammatik, oder Kunst die crainerische Sprach(e) regelrichtig zu reden, und zu schreiben*. Laybach.

Pohlin, Marko. 1783. *Kraynska grammatika. das ist: Die crainerische Grammatik, oder Kunst die crainerische Sprach(e) regelrichtig zu reden, und zu schreiben*. Laybach.

Šmigoc, Leopold. 1812. *Theoretisch-praktische Windische Sprachlehre*. Graz: Aloys Tusch.

Vodnik, Valentin. 1811. *Pismenost ali gramatika sa perve shole*. Ljubljana: Leopold Eger.

Vodnik, Valentin. 1988. *Zbrano delo*. Ljubljana: Državna založba Slovenije.

Secondary sources

Aalberse, Suzanne & Stoop, Wessel. 2015. The exceptional loss of the pronoun T. *Journal of Pragmatics* 88, 190–201.

Ahačič, Kozma. 2014. *The History of Linguistic Thought and Language Use in 16th Century Slovenia*. Frankfurt am Main: Lang.

Berger, Tilman. 1998. Äußere Einflüsse und interne Faktoren bei der Herausbildung der slavischen Anredesysteme. *Die Welt der Slaven* XLIII, 307–22.

Besch, Werner. 1998. *Duzen, Siezen, Titulieren: zur Anrede im Deutschen heute und gestern*. Göttingen: Vandenhoeck & Ruprecht.

Betsch, Michael. 2000. *Diskontinuität und Tradition im System der tschechischen Anredepronomina (1700–1850)*. München: Sagner.

Brown, Roger & Albert Gilman. 1960. The pronouns of power and solidarity. In Thomas A. Sebeok (ed.). *Style in Language*. Cambridge, Massachusetts: M.I.T. Press. 253–76.

Bunc, Stanko. 1940. *Pregled slovnice slovenskega knjižnega jezika*. Ljubljana: Jugoslovanska knjigarna.

Grdina, Igor. 1999. *Od Brižinskih spomenikov do razsvetljenstva*. Maribor: Založba Obzorja.

Grober-Glück, Gerda. 1994. *Die Anrede des Bauern und seiner Frau durch das Gesinde in Deutschland um 1930 unter volkskundlichen und soziolinguistischen Aspekten nach Materialien des Atlas der deutschen Volkskunde*. Frankfurt am Main: Lang.

Head, Brian. 1978. Respect degrees in pronominal reference. In Joseph H. Greenberg (ed.). *Universals of Human Language*. Vol. 3: *Word Structure*. Stanford: Stanford University Press. 151–211.

Hladnik, Miran. 2019. Literarische Zweisprachigkeit im slowenischen Teil der Habsburger Monarchie. In Andreas Leben & Alenka Koron (eds). *Literarische Mehrsprachigkeit im österreichischen und slowenischen Kontext*. Tübingen: Narr Francke Attempto. 289–304.

Houtzagers, Peter. 2018. The honorific third-person plural in Slavic. *Russian Linguistics* 42(1–2), 1–26.

Jelovšek, Alenka. 2011. Razvoj zaimenskega ogovornega sistema v slovenskih pisnih virih do leta 1850. *Slavistična revija* 59(2), 195–211.

Jesenšek, Marko. 2009. Nastanek in razvoj vzhodnoštajerskega knjižnega jezika. *Časopis za zgodovino in narodopisje* 80–45(1), 91–110.

Klemenčič, Matjaž. 2017. German-Slovene relations in the Slovene lands from the mid-19th century until today. *Onomàstica* 3, 127–60.

Kretzenbacher, Heinz L., John Hajek, Robert Lagerberg & Agnese Bresin. 2019. *Now you Sie me, now you don't.* The history and remnants of the 3pl V address pronoun calque in Slovak (onikanie) and in Czech (onikání). In Bettina Kluge & María Irene Moyna (eds). *It's Not All About You: New Perspectives on Address Research*. Amsterdam, Philadelphia: Benjamins. 125–37.

Lara-Bermejo, Víctor & Ana Rita Bruno Guilherme. 2021. The diachrony of pronouns of address in 20th-century European Portuguese. *Studies in Hispanic and Lusophone Linguistics* 14(1), 39–79.

Legan Ravnikar, Andreja. 2015. Valentin Vodnik. Pismenost ali Gramatika sa Perve Shole. In Kozma Ahačič (ed.). *Slovenske slovnice in pravopisi: spletišče slovenskih slovnic in pravopisov od 1584 do danes*. <https://www.fran.si/slovnice-in-pravopisi/11/1811-vodnik> (accessed 30 November 2024).

Legan Ravnikar, Andreja. 2017. K problematiki vpliva stičnega jezika – nemščine na semantične spremembe in stilno vrednost najstarejše slovenske knjižne leksike (16. stoletje). *Slovenski jezik – Slovene linguistic studies* 11, 35–53.

Melik, Vasilij. 2002. *Slovenci 1848–1918*. Maribor: Litera.

Merchiers, Ingrid. 2001. *Cultural Nationalism in the South Slav Habsburg Lands in the Early Nineteenth Century: The Scholarly Network of Jernej Kopitar (1780–1844)*. München: Sagner.

Mikhailov, Nikolai. 1998. *Frühslowenische Sprachdenkmäler: die handschriftlische Periode der slowenischen Sprache (XIV. Jh bis 1550)*. Amsterdam, Atlanta GA: Rodopi.

Orel, Irena. 2006. Jezikovne posodobitve Linhartove komedije Ta veseli dan ali Matiček se ženi skozi čas. In Marko Jesenšek & Zinka Zorko (eds). *Jezikovna predanost: Akademiku prof. dr. Jožetu Toporišiču ob 80-letnici*. Maribor: Slavistično društvo Maribor, SAZU Ljubljana. 572–84.

Orožen, Martina. 1981. Kako pojmovati Kopitarjevo opredelitev za ljudsko osnovo slovenskega knjižnega jezika. *Slavistična revija* 29(2), 187–200.

Prijatelj, Ivan. 1926. Korespondenca med Vodnikom in Kopitarjem. *Časopis za slovenski jezik, književnost in zgodovino* V. 121–43.

Reindl, Donald F. 2007. Slovene Ultra-Formal Address: Borrowing, Innovation, and Analysis. *Slovenski jezik* 6, 151–68.

Simon, Horst J. 2003. From pragmatics to grammar. Tracing the development of respect in the history of the German pronouns of address. In Irma Taavitsainen & Andreas H. Jucker (eds). *Diachronic Perspectives on Address Term Systems*. Amsterdam, Philadelphia: Benjamins. 85–23.

Slovar slovenskega knjižnega jezika, druga, dopolnjena in deloma prenovljena izdaja, 2014. <www.fran.si> (accessed 30 November 2024).

Suhadolnik, Stane. 1981. Kopitarjeva (dopisovalna) slovenščina. *Slavistična revija* 29(2), 171–86.

Šekli, Matej. 2018. *Tipologija lingvogenez slovanskih jezikov*. Ljubljana: ZRC SAZU, Inštitut za slovenski jezik Frana Ramovša, Založba ZRC.

Štih, Peter & Vasko Simoniti. 1995. *Slovenska zgodovina do razsvetljenstva*. Ljubljana: Korotan, Klagenfurt: Mohorjeva družba.

Thomas, George. 1997. The impact of purism on the development of the Slovene standard language. *Slovenski jezik* 1, 133–52.

Toporišič, Jože. 1976. *Slovenska slovnica*. Maribor: Obzorja.

Toporišič, Jože. 1992. *Enciklopedija slovenskega jezika*. Ljubljana: Cankarjeva založba.

Weiss, Peter. 2003. Načini ogovarjanja in govorjenja o odsotni osebi v govorih spodnje Zadrečke doline. In Zinka Zorko & Mihaela Koletnik (eds). *Glasoslovje, besedoslovje in besedotvorje v delih Jakoba Riglerja*. Maribor: Slavistično društvo. 199–215.

JUDIT VÁRI AND MARCO TAMBURELLI

6 Double standards in speakers' minds? An evaluation of standard varieties in Luxembourg and Belgium

ABSTRACT
We investigate how double-standard norms manifest at individuals' perceptual level. To do so, we study attitudes towards the (non-)standard varieties in two speech communities, i.e. the Belgische Eifel in Belgium and the Canton Clervaux in Luxembourg. In these communities, differences in codified norms emerge from differing standardisation processes: Standard Luxembourgish shows lower levels of codification and implementation, compared to the highly standardised *Hochdeutsch* present in Belgium. In contrast, norms of usage are more similar, since in both speech communities, individuals possess concepts of model speakers and model texts of their respective standards. In this paper, we hypothesise that differences in codified norms will manifest in the type of explicit attitudes that previous studies have found to require more cognitive effort. On the other hand, similarities in norms of usage will manifest in implicit attitudes requiring less cognitive effort. Implicit attitudes were tested with Implicit Association Tests, while explicit attitudes were measured with questionnaires.

1. Introduction

1.1. (Non-) standard language varieties in perceptual dialectology

Numerous empirical studies explore theories of language standardisation in individuals' language production, tracing concepts such as regional standard languages and colloquial standard languages (Lenz 2003; Lenz & Plewnia 2010; Auer 2011; Kehrein 2012; Elspaß & Kleiner 2019). However, fewer studies have empirically investigated theories of language standardisation in speakers' perception especially in Germanic speech communities, for example Austria, Switzerland, Germany, Luxembourg and Belgium (Huesmann 1998; Lenz 2003; Schmidlin 2011; Herrgen 2015; Schmidlin et al. 2017; Kiesewalter 2019). Studies which explore speakers' perspectives belong to the research field of perceptual dialectology (Preston 1999, 2010; Preston & Niedzielski 2013; Hundt 2018). These studies distinguish between standard varieties and non-standard varieties based on speakers' opinions about how these varieties differ (Preston 2016, 2018). Perceptual dialectological studies explore speakers' minds in order to identify dialectal areas as well as categories and attributions of (non-) standard languages varieties (Preston 1989; Purschke 2018; Kiesewalter 2019; Hoffmeister, 2021). Therefore, perceptual dialectological studies can be said to test language standardisation theories empirically by applying a cognitive perspective. Similarly, the present study which is situated in the field of perceptual dialectology aims to test language standardisation theory in speakers' minds.

Three studies on perceptions of Germanic (non-) standard varieties highlight methodological approaches which previously have been applied to empirically test theories of language standardisation (Lenz 2003; Schmidlin et al. 2017; Kiesewalter 2019). Schmidlin et al. (2017) explored the concept of pluricentricity in individual speakers' cognitive reality. According to language standardisation theories, pluricentric languages possess different but equally important centres – often in different nations – each with its own standard varieties (Clyne 1992). Using questionnaires to investigate corrective behaviour of teachers, Schmidlin et al. (2017) found hardly any evidence that teachers as norm authorities possess multiple norms such as Austrian and Swiss standard norms in addition to German standard norms. A study by Kiesewalter (2019) used auditory stimuli containing phonetic variations which were previously identified in the literature to be markers of various regional dialects in Germany. Participants from different dialectal regions listened to the stimuli and rated them on perceived vernacularity on a continuous scale.

Results show strong agreement in evaluating the (non-) standardness of speech, thus providing evidence for the interregional 'macro-synchronisations' of (non-) standard varieties posited by Schmidt & Herrgen (2011). The study by Lenz (2003), while mainly focused on exploring language standardisation theories in speakers' production, also includes speakers' perceptions, as interviewees comment on when in their lives they started using the standard variety instead of the non-standard variety.

In the last decades, perceptual dialectological studies started to include seminal work from (social) psychology on mechanisms of (social) cognition (Kahneman & Tversky 1972; Rosch et al. 1976; Rosch & Lloyd 1978; Chaiken 1980; Tajfel 1982; Tversky & Kahneman 2004; Turner 2010). Cognitive frameworks of language categorisation emerged in perceptual dialectology investigating speakers' mental representations, see Section 1.2. (Campbell-Kibler 2012; Speelman et al. 2013; Rosseel et al. 2018). Cognitive approaches to perceptual dialectology also introduced new methodologies based on implicit and explicit processes of speakers' perceptions, see 1.2.2. (Campbell-Kibler 2016; Rosseel & Grondelaers 2019). The next section (1.2.) will provide details on these approaches. Section 1.3. elaborates on how the present study applies these cognitive approaches to address its research questions.

1.2. Cognitive perspectives

Research in (social) psychology has found a plethora of evidence for mental representations of objects, people and concepts in individuals' minds (Kahneman & Tversky 1972; Rosch et al. 1976; Rosch & Lloyd 1978; Chaiken 1980; Tajfel 1982; Tversky & Kahneman 2004; Turner 2010). These mental representations help individuals to organise their (social) world in categories in order to facilitate quicker processing of stimuli (Chaiken 1980; Chen & Chaiken 1999). A category consists of more or less typical exemplifiers and the features they share (Rosch et al. 1976; Rosch & Lloyd 1978). For example, individuals possess mental representations of tables as part of this category, which is likely based on typical exemplifiers such as wooden dining tables and the feature 'four legs'. However, an office desk and even an operating table might be considered to be part of the category 'table' based on the feature 'four legs'. Similarly, individuals possess mental representations of social groups as categories such as specific demographics, e.g. 'young people', 'women' or specific professions such as 'librarians' (Kahneman & Tversky 1972; Tajfel 1982; Tversky & Kahneman 2004; Turner 2010). Mental representations include evaluative features – for example character traits such as 'shy' and 'introvert' in the case of librarians – and studies found evaluative features to be the basis for stereotyping based on factors such as ethnicity (Tajfel 1982; Fiske et al. 2002; Tversky & Kahneman 2004; Devine & Sharp 2009; Dovidio et al. 2010; Dolan 2014).

1.2.1. Mental representations of (non-) standard languages

Studies in perceptual dialectology have argued that speakers possess mental representations of (non-) standard language varieties similar to mental representations of objects and people in (social) categories (Grondelaers & van Hout 2016; Campbell-Kibler 2016; Hoffmeister 2021; Vári & Tamburelli 2023). In this view, language variety categories are based on typical exemplifiers and typical category features which include, on the one hand, the presence or absence of specific linguistic forms and, on the other hand, evaluative features. More specifically, studies show that individuals categorise specific lexical, morphological, syntactic and phonetic forms to be typical of their categories of 'standard language' to varying degrees (Lenz 2003; Kiesewalter 2019). In addition, evaluative features are also part of the mental representations for (non-) standard language, since studies have found evaluative features such as 'prestigious', 'warm' or 'clumsy' to be associated with (non-) standard language varieties, cf., for example, Preston (1999). Evaluative features of (non-) standard languages can also be attributed to individual speakers, with some studies exploring this attribution processes with, for example, the 'Speaker Evaluation Paradigm' (Ryan Bouchard et al. 1982; Zahn & Hopper 1985;

Campbell-Kibler 2012). Not only do attribution processes involve perceivers considering language varieties to be, for example, 'prestigious' or 'warm', but they also lead perceivers to consider speakers themselves as possessing these attributes. On the other hand, speakers' evaluative features also aid the categorisation of their speech. Studies show that possessing information about speakers' social group, including evaluative features (e.g. race, speakers' country of origin), impacts perceivers' categorisation of the speakers' utterances into, for example, standard and non-standard language or as different regional varieties (Walker et al. 2018; Kutlu et al. 2022).

1.2.2. Implicit and explicit evaluative processes

Studies on language attitudes provide valuable insights into the evaluative features of mental representations of (non-) standard languages (Campbell-Kibler 2012; Speelman et al. 2013; Rosseel et al. 2018; Kircher & Fox 2019). In addition, language attitude research highlights the distinction between two different types of evaluative processes, i.e. implicit and explicit processes, which are based on differences in cognitive underpinnings (Campbell-Kibler 2016; Rosseel & Grondelaers 2019). Similar to the evaluation processes of objects and people, evaluation of language varieties can be based on either implicit, automatic processes, or explicit, systematic processes (Chen & Chaiken 1999; Fazio & Towles-Schwen 1999; Wilson et al. 2000). Implicit cognitive processes require little time and cognitive resources, unlike explicit evaluation processes which require considerably more time and higher degrees of cognitive load (De Houwer & Moors 2007). Dual attitude models suggest that implicit as well as explicit evaluation processes are simultaneously present in individuals' minds, i.e. resulting in individuals holding simultaneously explicit as well as implicit attitudes towards the same attitude object (Chen & Chaiken 1999; Fazio & Towles-Schwen 1999; Wilson et al. 2000). Measuring these types of attitudinal processes also involves different methodologies. While explicit attitudes are elicited by direct methods such as questionnaires, reaction time experiments are used to tap into implicit attitudes (De Houwer & Moors 2007; De Houwer et al. 2009). For decades, research on language attitudes used only direct methods such as questionnaires to elicit explicit language attitudes, often as part-of-speech evaluation paradigms (Lambert et al. 1966; Kircher & Zipp 2022). However, more recent studies have begun applying reaction time experiments such as Implicit Association Tests or Affective Priming to tap into implicit evaluation processes (Campbell-Kibler 2012; Speelman et al. 2013; Rosseel 2022; Walker et al. 2022). Results show major differences between individuals' implicit and explicit attitudes towards (non-) standard varieties in many speech communities (Speelman et al. 2013; Rosseel et al. 2018; Vári & Tamburelli 2021, 2023). Studies on explicit attitudes frequently demonstrate either egalitarian attitudes towards (non-)standard varieties or a preference for non-standard varieties (Weber 2009; Vári & Tamburelli 2021). These results are often interpreted as signs of a rising appreciation for linguistic diversity (Costa et al. 2018). In contrast, implicit attitude measures tend to show more negative attitudes towards non-standard varieties and their speakers (Campbell-Kibler 2012; Speelman et al. 2013; Vári & Tamburelli 2023).

To sum up, the research field of perceptual dialectology encompasses studies on speakers' perception of (non-) standard languages (Preston 2018; Hundt 2018). Some studies test concepts of language standardisation theories – such as for example pluricentricity – at a micro-level, by investigating speakers' minds (Schmidlin et al. 2017; Kiesewalter 2019). Cognitive studies in perceptual dialectology suggest that perceivers possess mental representations of standard as well as non-standard varieties as categories with typical exemplifiers, including typical linguistic forms and evaluative features. Perceivers access evaluative features of mental representations based on either implicit, automatic cognitive processes or explicit, systematic processes.

1.3. Research gap: Mental representations of double-standard norms in speakers' minds

Some perceptual dialectological studies explore whether theories of language standardisation are reflected in speakers' minds (Schmidlin 2011; Schmidlin et al. 2017; Kiesewalter 2019). However, these studies attempt to elicit speakers' mental representations via explicit processes, since they use questionnaires which require more time and cognitive load compared to reaction time experiments. Studies on individuals' implicit evaluation of (non-) standard language varieties differ in their results from studies on explicit evaluations (Speelman et al. 2013; Rosseel et al. 2018; Vári & Tamburelli 2023). We therefore suggest that implicit processes are equally important in investigating mental representations of (non-) standard languages and their overlap with concepts of language standardisation theories. Consequently, this paper explores individuals' mental representations of (non-) standard languages based on implicit as well as explicit cognitive processes. More specifically, we aim to focus on evaluative features of mental representations, and not, for example, the presence or absence of specific linguistic forms. This paper asks whether mental representations of (non-) standard languages overlap with the concept of double-standard norms as defined by Ammon (2003a). Overall, the concept of double-standard norms suggests the presence of two distinct norms of standard language varieties, i.e. norms of usage as well as codified norms (Ammon 2003a). In the following, we elaborate on the concept of double-standard norms within our two case studies, i.e. the speech communities of Belgium and Luxembourg.

1.4. Norms of usage and codified norms in Luxembourg and Belgium

We explore double-standard norms in speakers' minds in Belgium and Luxembourg by investigating norms of standard language usage as well as codified norms of standard language, both in relation to their related non-standard varieties.

Non-standard Moselle-Franconian varieties are spoken in both speech communities. As well as being the most widely spread non-standard varieties in the Belgische Eifel, a region in the south of the German-speaking community of Belgium, Moselle-Franconian varieties are also present as non-standard varieties in Canton Clervaux, a region in the North of Luxembourg that borders with Belgium. For the sake of brevity, we will refer to the two speech communities described above as 'Belgium' and 'Luxembourg' in the following sections of the paper.

1.4.1. Codified norms

Overall, we chose to investigate double-standard norms in Belgium and Luxembourg due to the strong genealogical similarities of their non-standard Moselle-Franconian varieties on the one hand, and the major differences regarding their choices of standard varieties, namely Standard Luxembourgish and Standard German, on the other. We propose that differences between these standard varieties are particularly prominent when considering codified norms in contrast to norms of usage especially from a perceptual perspective. Such a combination of differences and similarities between the speech communities enables us to tease apart double-standard norms at the cognitive level and investigate how norms of usage manifest separately from codified norms in speakers' mental representations.

Codified norms are standard varieties based on linguistic codices such as dictionaries, grammar books and pronunciation guides (Ammon 1996). After selecting one or more varieties to serve as a standard, codification can take place as one phase of standardisation processes (Haugen 1968; Haugen 1997; Coupland & Kristiansen 2011). However, codified standards often need to be further elaborated in functions and forms, thus elaboration constitutes an additional aspect of language standardisation processes. Finally, the last aspect of language standardisation processes involves the full implementation and acceptance of standard varieties by the speech community (Haugen 1968; Haugen 1997; Coupland & Kristiansen 2011).

In Luxembourg, standardisation processes in the twentieth century resulted in vernacular Moselle-Franconian being associated with Standard Luxembourgish, a standardised variety of Moselle-Franconian (Stell 2006; Gilles 2015). The creation of Standard Luxembourgish introduced Moselle-Franconian into the written domain, thereby elaborating its varieties' functions (Fehlen 2016; Gilles 2015). Codices such as dictionaries were created for Standard Luxembourgish, though a spoken standard variety was not codified, for example in the form of pronunciation guides (Newton 1996; Stell 2006; Gilles 2015). Standard Luxembourgish was legally recognised as an official language of Luxembourg in 1984, potentially bolstering its acceptance in the community (Horner 2009; Horner & Weber 2008; Gilles 2015). However, Moselle-Franconian speakers still lack extensive formal education in the codified norms of Standard Luxembourgish, which leads to limited knowledge of those norms emerging in, for example, variations in orthography and morphosyntax (Horner 2009; Entringer et al. 2018). Finally, codified norms show low degrees of acceptance and implementation in the speech community, since teachers with sound knowledge of codified norms accept non-standard variants of Moselle-Franconian in schoolwork (Horner & Weber 2015).

In Belgium, Standard German is the standard variety associated with Moselle-Franconian vernaculars (Weber 2009; Darquennes 2013; Möller 2017;). Standard German is a highly standardised variety with a much earlier onset of standardisation processes compared to Standard Luxembourgish, dating back to the sixteenth century (Durrell 1999; Mattheier 2003). Numerous codices of Standard German validate the written and spoken standard norm and corrective behaviour oriented towards it in Germany as well as in speech communities with Germanic varieties (Durrell 1999; Mattheier 2003). Furthermore, Standard German in Belgium shows a significantly higher degree of standardisation than Standard Luxembourgish does in Luxembourg, reaching the last stage of standardisation processes, i.e. implementation and acceptance in the community (Haugen 1997; Vári & Tamburelli 2021). Moselle-Franconian speakers in Belgium recognise not only the codified norms of Standard German as a standard variety for their vernaculars, but they also acquired these codified norms in formal education (Combuchen 2009; Weber 2009; Möller 2017). Unlike in Luxembourg, the standard variety is the language of instruction as well as a subject language in schools within the speech community.

1.4.2. Norms of usage

In contrast to codified norms, norms of usage consist of linguistic variants which are regularly used by model speakers or appear in model texts but are not codified (Ammon 2003a). According to Ammon (2003a), model speakers and writers are individuals who are regarded by others to be of higher status and who produce linguistic output in high prestige domains, such as politics, religion, education and media. This linguistic output can become a standard variety of a language by usage (Ammon 2003a, 2003b).

The complex process of how language usage establishes standard language norms is significantly under-researched from a perceptual cognitive perspective (Hoffmeister et al., 2021) and it is beyond the scope of this paper to discuss the current literature on norm development processes in depth. Overall, cognitive perceptual dialectology suggests that norms of standard language (codified norms as well as norms of usage) constitute social norms i.e. patterns of beliefs and behaviours. A plethora of studies has explored social norms in groups of individuals and identified different phases such as norm conception, diffusion, internalisation and institutionalisation (see for an overview Legros & Cislaghi 2020). Building on this research, we suggest that codified norms of standard language can be conceptualised as more institutionalised social norms. In contrast, norms of standard language by usage are more similar to internalised habitual social norms.

In general, evidence supporting norms of spoken standard language by usage comes from studies reporting that listeners show moderate to high degrees of consistency in those spoken language varieties and linguistic variants they consider to be more or less spoken standard language (Lenz 2003; Kiesewalter, 2019).

Perceptual studies have found evidence for standard norms of usage in Luxembourg (Fehlen 2008, 2009, 2016; Neises 2013; Entringer et al. 2018). Two studies report that speakers in Luxembourg believe in the existence of a variety of Luxembourgish / Moselle-Franconian which is spoken, not regionally marked, and has higher degrees of intelligibility and prestige (Neises 2013; Entringer et al. 2018), all characteristics which are typical of spoken standard norms of usage (Ammon 1989, 2003b, 2004). Participants locate this type of speech mainly in the Centre region such as Luxembourg City (Staader) and in the Alzette Valley (Entringer et al. 2018; Neises 2013). Luxembourgish speakers identify literature and administrative communications by the state or council rather than websites and social media posts to be examples of model texts for Standard Luxembourgish (Entringer et al. 2018). Equally, they identify specific model speakers such as politician Jean-Claude Juncker as well as linguist and author Alain Atten (Entringer et al. 2018).

Even if perceptual studies of our speech community in Belgium are scarce, the literature suggests that standard language norms of usage are also present (Nelde 1987; Nelde & Darquennes 2002; Combuchen 2009; Weber 2009; Darquennes 2013, 2019; Möller 2017). Examples of model texts of Standard German in Belgium are the daily newspaper Grenz-Echo, official communications from the political administration and a small literary tradition (Combuchen 2009; Möller et al. 2013). The proximity to the neighbouring state of Germany fosters constant exposure to a plethora of model texts from Germany in the form of literature and media (Möller 2017; Darquennes 2019). Model speakers of Standard German are found not only in German media such as private TV stations, but also in the TV and radio programmes of the BRF, the Belgian broadcasting station (Combuchen 2009).

1.4.3. Hypotheses regarding codified standard norms and norms of standard language usage

Overall, we identified differences between codified norms of both Standard Luxembourgish and Standard German, which appear evident in both, and the onset of standardisation processes as well as the degrees of standardisation. In contrast, we argued that norms of usage are more similar in the two speech communities when compared to codified norms. Evidence suggests that model speakers and model texts of standard language usage are equally present and equally rooted in Moselle-Franconian speakers' exposure to the usage of Standard Luxembourgish (for Luxembourg) or Standard German (for Belgium) in similar contexts such as media, civil administration and literature. Consequently, we hypothesise that the difference between speech communities will show more strongly in speakers' mental representations of codified norms compared to their mental representations of norms of standard language usage.

In addition, we suggest a conceptual link between different types of cognitive processes, i.e. explicit and implicit, and double-standard norms. We propose that differences and similarities in double-standard norms, i.e. codified norms and norms of usage, will manifest in speakers' mental representations that are accessed via explicit and implicit processes. On the one hand, mental representations of codified norms are based on knowledge of actual codices. When speakers recall knowledge of their codified standard norms, this process requires time and cognitive resources and thus is based on explicit underlying cognitive process. In contrast, we argue that individuals access their mental representation of standard norm of usage by implicit, automatic cognitive processes. Norms of usage are most likely acquired via habitual exposure and less via formal education, unlike codified standard varieties. Therefore, we suggest that pathways to mental representations of norms of usage require less cognitive effort and time compared to pathways to retrieve representations of codified norms. In the scope of this study, we focus only on the evaluative features of the mental representations which can be explored as explicit and implicit attitudes.

Based on the links we discussed above, we propose that differences between the speech communities of Belgium and Luxembourg will manifest more strongly in speakers' explicit attitudes (which we take to be indicative of codified norms) compared to their implicit attitudes (which we associate with norms of usage). Consequently, the difference between Belgian and Luxembourgish participants' explicit attitudes towards their respective standard varieties (Standard German vs. Standard Luxembourgish) will be significantly larger than the difference in their implicit attitudes. To our knowledge, the theoretical framework of double-standard norms has not yet been empirically investigated at a cognitive level by tapping into implicit evaluative processes.

2. Methodology

Participants were recruited with calls from local media and by collaborating with local community services such as social clubs. Sixty-one Belgian and sixty-eight Luxembourgish participants (seventy-six female, mean age forty-one years, range between nineteen and seventy-seven years) are part of the final data sets measuring participants' explicit and implicit attitudes towards their respective standard varieties in relation to their non-standard varieties. Participants were competent and authentic speakers of the non-standard varieties in the respective speech community, i.e. the non-standard Moselle-Franconian varieties in the Luxemburgish Canton Clervaux, namely the Éisleker varieties, as well as the non-standard Moselle-Franconian varieties in the Belgian region of the Belgische Eifel. Belgian participants rated their competence in Standard German and Moselle-Franconian as high on a five-point scale (from 0/not at all, to 4/perfect, Moselle-Franconian: mean = 3.27, s.d. = 0.70; Standard German: mean = 3.43, s.d. = 0.51), similar to their Luxembourgish peers' linguistic background (Éisleker mean: = 3.71, s.d. = 0.57; Standard Luxembourgish: mean = 3.30, s.d. = 0.71). Participants in both groups reported being nationals of the respective country and having spent the majority of their childhood in the respective regions (see 1.3). Participants from Belgium did not significantly differ from the Luxembourgish participants regarding their age and gender (BELG: thirty-seven females, overall: mean age = fourty-four years, range: twenty-one – seventy-four; LUX: thirty-eight females, overall: mean age = thirty-eight years, range: nineteen – seventy-seven), see also Figure 6.4 and 3.2. Results. Participants first took part in an Implicit Association Test (IAT) (study 1) and afterwards they filled out an Attitudes towards Language (AtoL) questionnaire (study 2). The two studies are briefly described below in further detail (see Vári & Tamburelli 2021, 2023).

2.1. Study 1: The IAT

The IAT is an established measure applied in sociolinguistic studies to explore implicit attitudes towards different language varieties in various speech communities, for example in the US, Canada (Quebec), Spain (Catalonia), Belgium (Pantos & Perkins 2012; Lehnert & Hörstermann 2019; Lehnert et al. 2018b; Rosseel et al. 2018; Ianos et al. 2020). IATs are based on implicit automatic processes since they require little degrees of cognitive load and time (De Houwer 2001; De Houwer et al. 2009). The underlying assumption of the IAT is that implicit attitudes can be inferred from how strongly participants associate two categories with each other such as 'standard variety' with 'positive valence'. The strength of association becomes evident mainly in participants' reaction times when categorising exemplifiers of the contrasting categories via key press. The IAT has been subjected to much scrutiny since its inception (Greenwald et al. 1998), with discussions revolving around the construct validity of this measure, its predictive value for individuals' behaviour, the numerous procedural features of the experimental paradigm and its stimuli as well as data scoring and analyses (Blair 2002; Blanton et al., 2009; Gast & Rothermund 2010; Agosta et al. 2011; Blanton et al. 2015; Wolsiefer et al. 2017; Meissner et al. 2019; Schimmack 2021; Cummins et al. 2022). It is beyond the scope of this paper to provide an overview of studies on the IAT from the last two decades. Overall, the methodology has improved significantly (Greenwald et al. 2022) and is considered to be a reliable measure of group differences between associations (Schimmack 2021). For a detailed overview of the use of IATs in linguistics, see Rosseel (2022).

2.1.1 Materials

Each IAT stimulus belongs to one category, namely the target and the attribute category. Each category has two sub-categories resulting in four stimuli groups. Target stimuli represent the target category, which is the object of study, in this case the category 'language variety'. The category 'language variety' is sub-divided into two groups of stimuli, one group representing the respective standard varieties, while the second stimulus group represents the non-standard varieties. Attribute stimuli are exemplifiers of attribute categories, most commonly – and in the case of our study – emotional valence, sub-divided into positive or negative valence (De Houwer 2001; Rosseel 2022). Similar to the target stimuli, attribute stimuli represent either the sub-category positive valence or negative valence. Stimuli can only belong to one of the four groups, i.e. standard variety, non-standard variety, positive valence or negative valence.

Audio recordings of six words in four different language varieties served as target stimuli in the present study. Participants listened to six different target words in two variations, i.e. six words in their respective standard variety (Standard German or Standard Luxembourgish) and the same six words in their non-standard variety (non-standard Moselle-Franconian from either Belgium or Luxembourg), totalling in twelve target stimuli per IAT, see word list Table 6.1. (Non-) standard variations were elicited by phonetic and lexical variants. Audio stimuli are available on the platform Open Science Framework.[1]

In order to avoid confounds, we controlled for frequency and neutral valence of target words by extensive study of previous analyses of

1 DOI 10.17605/OSF.IO/E8NMP (last accessed 21.04.2024).

Table 6.1 Target stimuli list for IAT – Standard German-based transcription for non-standard varieties.

Belgium		Luxembourg	
English	standard / non-standard	English	standard / non-standard
hat	*Mütze / Kapp*	*basket*	*Kuerf / Koref*
head	*Kopf / Kopp*	*eyes*	*Aan / Auen*
milk	*Milch / Melisch*	*navel*	*Bauschknöpfsche / Nu^{e}bel*
nine	*Neun / Nöng*	*people*	*Leit / Lekt*
pages, leaves	*Blätter / Bladder*	*sew*	*bitzen / beizen*
people	*Leute / Lökt*	*socks*	*Strömp / Huasen*

word frequency and sentiment in German, Dutch, English and French (Bradley & Lang 1999; Corson & Quistrebert 2000; Bonin et al. 2003; Bertels et al. 2009; Lahl et al. 2009; Võ et al. 2009; van Brysbaert et al. 2011; Moors et al. 2013; Warriner et al. 2013; Heuven et al. 2014). Finally, our small-scale norming study established frequency and valence ratings of non-standard words as well as speakers' perceived authenticity (Vári & Tamburelli 2023).

Attribute stimuli are exemplifiers for the attribute category which in the present study indexed general emotional valence, i.e. 'positive/negative'. We chose emotionally valent pictures as attribute stimuli in order to present participants with non-verbal stimuli to avoid language interferences. The sixteen pictures from the International Affective Picture System (Lang et al. 2008) consisted of eight positive pictures and eight negative pictures which showed no depictions of any words which could be semantically related to the target words. We established this absence of associative links by consulting data sets on individuals' word association norms (Kiss et al. 1973; Nelson et al. 1998; Melinger & Weber 2006)

A detailed description of the materials and procedures of the present study's IAT is beyond the scope of this contribution but are reported on in Vári & Tamburelli (2023).

Table 6.2 Overview of IAT block design: non-st = non-standard; st = standard.

	study type 1				study type 2			
block	amount of trials	stimulus type	left key 'E'	right key 'I'	amount of trials	stimulus type	left key 'E'	right key 'I'
1	20	audio	non-st	st	20	audio	st	non-st
2	20	image	positive	negative	20	image	positive	negative
3	20	both	non-st & positive	st & negative	20	both	st & positive	non-st & negative
4	40	both	**non-st** & positive	**st** & negative	40	both	**st** & positive	**non-st** & negative
5	20	audio	**st**	**non-st**	20	audio	**non-st**	**st**
6	20	both	st & positive	non-st & negative	20	both	non-st & positive	st & negative
7	40	both	st & positive	non-st & negative	40	both	non-st & positive	st & negative

2.1.2. Procedure

IATs were delivered online and lasted approximately twelve minutes on average. Participants sorted target auditory stimuli into the target categories 'standard' and 'vernacular' as well as visual attribute stimuli into the attribute categories 'positive' and 'negative' by pressing designated keys on the keyboard. Key mappings were indicated at the top of the screen by labels for language varieties or non-verbal labels, in the form of 'thumbs up' and 'thumbs down' pictograms for the attribute categories 'positive' and 'negative'. Labels indicating language varieties were 'Hochdeutsch' for Standard German, 'Lëz-St' for Standard Luxembourgish, 'Platt' for the Moselle-Franconian variety in Belgium and 'Éislek' for the Moselle-Franconian variety in Luxembourg. Vári & Tamburelli (2021) describe motivations for choosing these labels based on previous literature and the small-scale norming study. Visual stimuli

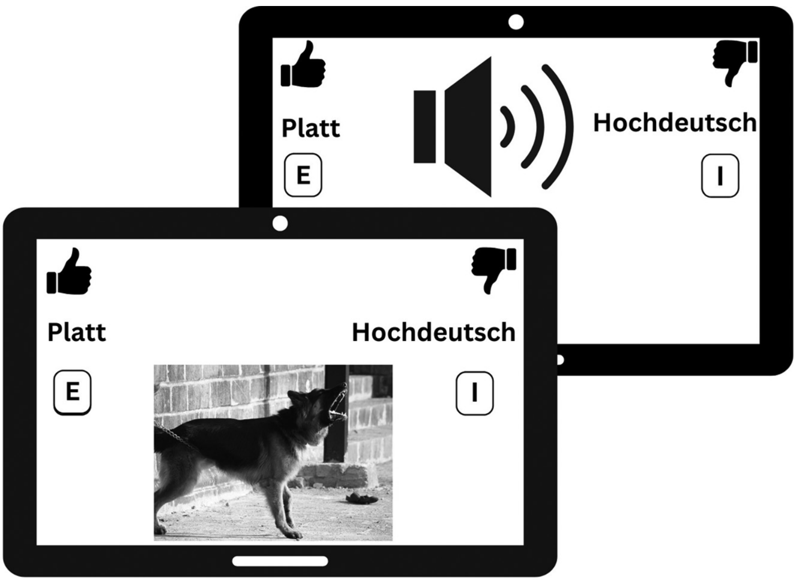


Figure 6.1 Two categorisation tasks alternate (categorisation via key press left key 'E' or right key 'I').

(emotionally valent picture) appeared in the middle of the screen and a pictogram of a loudspeaker indicated the playback of sound for auditory stimuli (see Figure 6.1).

Each trial consisted of participants categorising one stimulus, and participants' accuracy and reaction time (RT) were measured for each categorisation in milliseconds. Participants received negative feedback for a false categorisation in the form of a red cross. The IATs had seven blocks in which stimuli and key mapping of categories changed following the classical design of IATs (Greenwald et al. 1998; Greenwald et al. 2022; Rosseel 2022), see Table 6.2 for a summary of the experimental design. The key of this experimental design is a switch of key mappings after block 4. Target stimulus categories, i.e. 'standard or vernacular', change position, while the mapping of 'positive/negative' remains constant. Participants were randomly assigned to two study types with different sequences of key mapping combinations to counterbalance order effects.

2.2. Study 2: The Attitudes towards Language (AtoL) questionnaire

2.2.1 Materials

We measured participants' explicit attitudes towards their respective standard varieties by using the AtoL questionnaire (Schoel et al. 2012). Attitude measures using questionnaires give participants ample time to reflect on their answers, which requires more cognitive load (Kircher 2016; Rosseel & Grondelaers 2019). The AtoL is a multiscale questionnaire consisting of bipolar adjectives arranged on semantic differential scales (Osgood 1952). AtoL questionnaires were carefully constructed and validated by their application in various language varieties in different contexts, e.g. Bavarian, Saxonian, German, English, Chinese (Schoel et al. 2012). The questionnaire consists of three main factors which reflect the perception of languages: first, the dimensions of Sound (e.g. harsh – soft), second, Structure (e.g. precise – vague) and finally, Value (e.g. beautiful – ugly). We included only the factor Value with its five semantic differential scales in our AtoL questionnaire as a general measure of evaluative features. Vári & Tamburelli (2021) describe further reasons for this focus. One additional semantic differential scale complemented the five original scales in order to facilitate comparisons with previous adaptations of the AtoL questionnaire in Luxembourg (Lehnert et al. 2018a; Lehnert et al. 2018b). Previous research included bipolar adjective pairs in German, which the present study used for semantic differential scales of the Value dimension (Schoel et al. 2012a; Schoel et al. 2012b; Schoel & Stahlberg 2012; Lehnert et al. 2018b). The present study was the first to use Luxembourgish as a language of instruction in an AtoL questionnaire – besides German – which made it necessary to translate bipolar adjectives into Luxembourgish with the help of a native speaker. Table 6.3 below shows all semantic differential scales. In our AtoL questionnaire, the six semantic differential scales of the Value dimension were combined with labels representing the respective language varieties. In Belgium, three labels were presented to participants representing Standard German (*Hochdeutsch*), Moselle-Franconian (*Platt*) and French (*Französisch*).

Table 6.3 Bipolar adjective pairs of the AtoL semantic differential scales from the Value dimension with additions from Lehnert (2018), here (L).

English	German	Luxembourgish
beautiful – ugly	*schön – hässlich*	*schéin – ellen*
appealing – abhorrent	*ansprechend – abstoßend*	*uspriechend – ofstoussend*
pleasant – unpleasant	*angenehm – unangenehm*	*agreabel – desagreabel*
inelegant – elegant without style – with style	*unelegant – elegant*	*net elegant – elegant*
clumsy – graceful	*schwerfällig – anmutig*	*schwéierfälleg – liichtfälleg*
practical – impractical (L)	*unpraktisch – praktisch* (L)	*onpraktesch – praktesch* (L)

In Luxembourg, four labels indexed Standard Luxembourgish (*Staater / Standard Lëtzebuergesch*), Moselle-Franconian (*Éislek*), Standard German (*Hochdeutsch*) and French (*Französisch*).

2.2.2 Procedure

After participants completed the IAT described in Section 2.1., they completed online versions of AtoL questionnaires which lasted approximately fifteen to twenty minutes. Participants indicated for each language variety how they would describe it on six sematic differential scales of the Value dimension. More specifically, participants marked the positions between six bipolar adjective pairs ranging from 0/left adjective to 4/right adjective. Bipolar adjective pairs were presented in a randomised order and their positions on the two opposing sides of the semantic differential scales were pseudo-randomised. Negative and positive adjectives changed their position from the left to the right side of the scale or vice versa with every third semantic differential scale. Our aim was to ensure participants' attention and avoid potential position effects, which can affect the validity of responses (Dörnyei & Taguchi 2009).

3. Statistical analyses

3.1. Data pre-processing

Before statistical analysis, both attitude measures (i.e. IATs as well as AtoL questionnaire) required pre-processing of the data. The final IAT data set of 129 participants excluded data from participants participating multiple times and participants whose response times were less than 300 ms for more than 10% of their trials indicating a lack of engagement with the experiment. Equally, participants who reported a high amount of unknown stimulus words in the background questionnaire were excluded from the final analysis as this indicated potentially low language competence. Overall, these criteria resulted in the exclusion of thirty-six participants.

In addition, we followed standard procedures when matching the data of the two study types with different key mapping (i.e. study type 1 and 2), we excluded erroneous trials as well as trials with latencies above 10,000 ms (Greenwald et al. 2003). Finally, we calculated the standard (difference) 'd-score', which is an effect size measure of category associations (Greenwald et al. 2003). In our case, the d-score represents implicit attitudes towards non-standard Moselle-Franconian compared to attitudes towards the respective standard variety, i.e. Standard Luxembourgish or Standard German. We calculated the d-score by averaging each participant's RTs for the blocks with the 'non-standard' and 'positive' mapping' as well as the blocks with the inverted mapping, i.e. 'standard' and 'positive' and then subtracting the former mean RT from the latter mean RT. Differences in RTs based on averages of trial blocks with inverted key mapping are then divided by participants' pooled standard deviations.

The pre-processing of the AtoL scores also included screening for multiple participation normalising of the inverted semantic differential scales and calculating mean scores of the six sematic differential scales for each language variety. If data was discarded for IAT scores, participants' matching AtoL responses were discarded as well.

Two participants were excluded from the analysis since their contradictory responses for inverted items were indicative of superficial engagement with the task or positions effects. Both are factors potentially influencing the validity of the attitude measurement (Dörnyei and Taguchi 2009).

3.2. Results

All analyses were performed using R Statistical Software (version 4.3.1; R Core Team 2023) and the lme4 package (Bates et al. 2015). We used logistic regression to investigate our hypothesis stating a larger difference between Belgian and Luxembourgish participants' codified norms, i.e.

explicit attitudes towards their respective standard varieties (Standard German vs. Standard Luxembourgish), and compared their implicit attitudes towards their respective Standard varieties (norms of usage). We fitted a logistic model (estimated using Maximum Likelihood Estimation) to predict participants' speech community (Belgium vs. Luxembourg) with three fixed effects d-score (implicit attitude non-standard versus standard), AtoL_non_st_mean (explicit attitude non-standard), AtoL_St_mean (explicit attitude standard) which were all z-scored. In addition, models to converge included a random intercept for participants (pp).

The model's explanatory power is indicated by AIC 110.2 and BIC 124.5. Within this model only the effect of d-score (implicit attitude non-standard versus standard) and AtoL St mean (explicit attitude standard) are statistically significant (d-score: $\beta = 1.90$, 95% CI [0.50, 3.37], $p = 0.008$ and AtoL St mean: $\beta = -1.57$, 95% CI [-2.80, -0.32], $p = 0.014$) when holding the other covariates constant, see also Figure 6.4. The effect of the d-score is positive, meaning that the more positive the d-score, i.e. the more positive the implicit attitudes towards vernacular varieties and conversely, the more negative the implicit attitudes towards the standard variety, the more likely participants' speech community is Luxembourg and not Belgium . In contrast, the effect of AtoL St mean is negative, since negative ratings of explicit attitudes towards the standard variety increase the likelihood that participants are from Luxembourg and not Belgium (see also Figure 6.2 & Figure 6.3). All model estimates are reported in log-odds.

However, for the purposes of our study of differences between codified norms (explicit attitudes) and norms of usage (implicit attitudes), we are more interested in the standardised estimates and their confidence intervals to explore the magnitude of how attitude measures affect the likelihood that participants originate from Belgium or Luxembourg. Similarly, confidence intervals allow us to explore the variance of these estimates and thus make inferences about the consistency and homogeneity of participants' responses. Standardised estimates of effects of d-scores (implicit attitudes) deviate more from 0 (no impact) compared to standardised estimates of the AtoL scores (explicit attitudes) by 0.41, showing

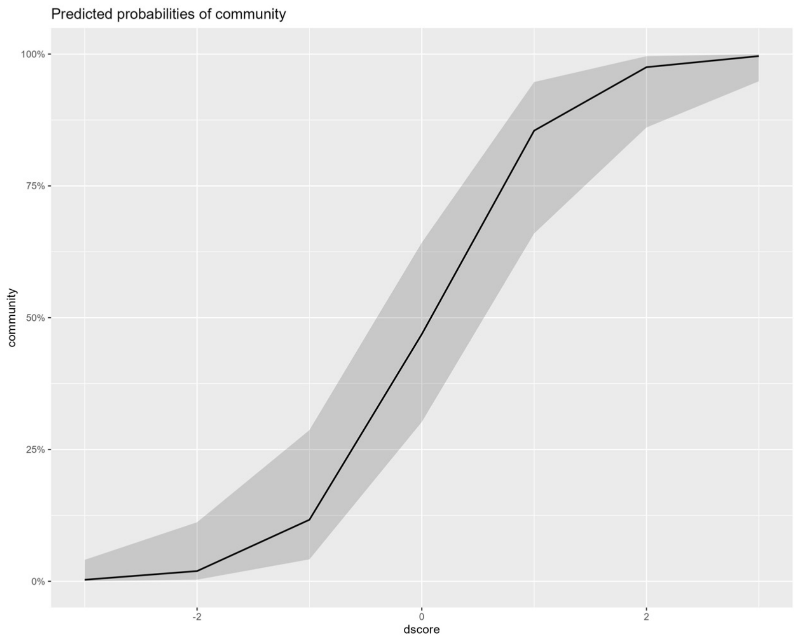


Figure 6.2 Predicted probability of participants belonging to speech community Belgium vs. Luxembourg predicted by d-score.

a stronger effect of d-scores (implicit attitudes) on the difference between speech communities. In contrast, variance around the estimated effect of the d-score (implicit attitudes) is slightly larger (95% CI [0.50, 3.37], 2.87) compared to the estimated effect of AtoL scores, i.e. explicit attitudes (95% CI [-2.80, -0.32]; 2.48). Standardised estimates of gender, age and AtoL non-standard ratings do not significantly predict the speech community of participants (see Figure 6.4).

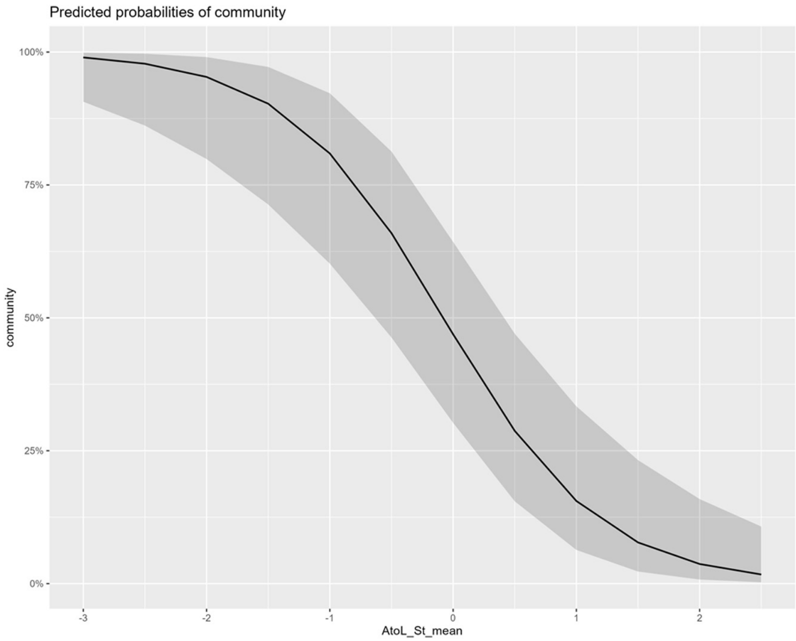


Figure 6.3 Predicted probability of participants belonging to speech community Belgium vs. Luxembourg predicted by AtoL ratings.

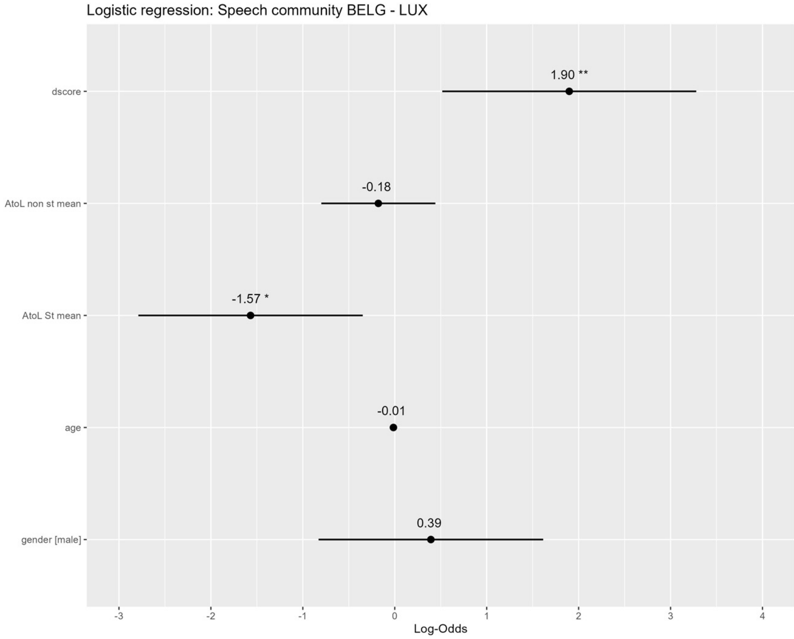


Figure 6.4 Untransformed estimates on the linear scale.

4. Discussion

This study set out to compare standard language norms in Luxembourg and Belgium at individuals' cognitive level of mental representations. While Moselle-Franconian speakers are present in both speech communities, their vernaculars are associated with different standards, namely Standard Luxembourgish in Luxembourg and Standard German in Belgium (Gilles 2015; Darquennes 2019). We argued that previous literature shows larger differences between the codified norms of Standard German and Luxembourgish compared to the differences found in their norms of usage (Entringer et al. 2018; Weber 2009; Neises 2013). We identified codified norms of Standard Luxembourgish to be different from codified norms of Standard German, since a standard variety for Luxembourgish/Moselle-Franconian was created significantly later than the emergence of Standard German (Mattheier 2003; Stell 2006). Overall, the later onset of the standardisation processes in Luxembourg led to lower degrees of standardisation for Standard Luxembourgish compared to Standard German (Haugen 1997; Vári & Tamburelli 2021). Evidence of these lower degrees of standardisation is found, for example, in the lack of implementation and acceptance of codified norms in Luxembourg (Horner & Weber 2015; Entringer et al. 2018).

On the other hand, we argued that previous research shows norms of usage to be relatively similar across Luxembourg and Belgium, compared to the differences we find in codified norms. In both speech communities, model speakers and model texts of standard language are present in similar contexts such as media, civil administration, political discourses and literature (Weber 2009; Neises 2013; Entringer et al. 2018).

Most importantly, we suggested that differences between codified norms and similarities between norms of usage will manifest at a cognitive level among Moselle-Franconian speakers. We proposed that speakers tap into their mental representations of (non-) standard languages via two types of cognitive processes, i.e. implicit and explicit processes. To recall mental representations of codified norms, speakers require considerable time and cognitive resources, since these norms are based on the knowledge of actual codices. Consequently, we hypothesised that codified norms and their mental representations are accessed via explicit processes. On the other hand, norms of usage are based on implicit cognitive processes requiring less time and cognitive load, since norms of usage do not require explicit knowledge of codices, being based instead on speakers' habitual exposure to model texts and speakers. We proposed to focus on evaluative features of mental representations in the form of implicit attitudes and explicit attitudes towards (non-) standard language varieties.

Overall, we hypothesised that differences between speech communities will be more prominent in speakers' explicit attitudes, as we take these to arise from mental representations of codified norms, while implicit attitudes, which we base on mental representations of norms of usage, will show less – if any – difference. We therefore expected that Belgian and Luxembourgish participants' explicit attitudes towards their respective standard varieties (Standard German vs. Standard Luxembourgish) would show significantly greater differences when compared to their implicit attitudes.

Contrary to our hypothesis, implicit attitudes were better predictors of participants' speech community compared to explicit attitudes. However, implicit attitudes showed more variation when predicting how likely participants are to belong to either the Belgian or Luxembourgish speech community, providing evidence for inconsistencies in effects of implicit attitudes. One possible interpretation of this result is the presence of potential confounds which we did not take into account in the present study. For example, we assumed norms of usage to be more similar than codified norms in the two speech communities based on the presence of comparably accessible model speakers and texts in similar domains. We suggested that the presence of model speakers and model texts in speakers' cognitive reality shows that norms of usage in Belgium as well as Luxembourg underwent one phase of the standardisation process, i.e. implementation. However, we failed to sufficiently consider both sides of this aspect of Haugen's (1997) analysis of standardisation processes which does not only include implementation but also acceptance of norms.

We did not consider that Belgian speakers might still differ from Luxembourgish speakers regarding acceptance of their respective norms of usage, which ultimately could play a role in evaluative features of mental representations. Speakers could be aware of model texts and model speakers which are considered to be prestigious in their speech community, but they might still not fully accept these norms in their individual cognitive reality. This conflict between implementation and acceptance of standard norms has previously been raised in discussions about implicit (language) attitudes, and specifically whether these types of attitudes are more indicative of individuals' preferences or of societal norms (Rosseel et al. 2018). Based on the scarce literature on speakers' evaluation of standard norms in Belgium and Luxembourg, we argued that norms of usage are more similarly implemented and accepted in both speech communities compared to codified norms. We hypothesised these similarities would be reflected in implicit attitudes in contrast to explicit attitudes. However, this hypothesis did not consider potential differences in how well norms of usage are accepted at the individual as well as the community level, and how these might affect implicit attitudes.

Consequently, we suggest that future studies that aim to investigate double-standard norms in speakers' cognitive reality should first focus only on implicit and explicit categorisation processes of (non-) standard linguistic features. Such studies could limit their focus to measuring the presence or absence of specific phonetic and semantic variants in speakers' mental representations without measuring evaluative features. The limited scope would still allow an investigation of language standardisation theories, such as double norms in speakers' minds, but would not introduce confounds of language evaluation. Later studies could then build on this work and explore evaluations of phonetic and semantic variants.

In addition, future studies would need to address the limitations of the present study regarding the modality of language norms. While implicit attitudes were measured in relation to spoken standard norms using IATs, explicit attitude measures were elicited based on exposure to written labels in questionnaires. However, due to an external factor (i.e. the absence of codified norms for a spoken standard variety in Luxembourg), this variation in modality might have considerably reduced the comparability of the two sets of results.

Limitations notwithstanding, the present study was the first to propose a link between double-standard norms and implicit and explicit cognitive processes in speakers' minds. This contribution highlights the importance of empirically testing the psychological reality of the concepts behind language standardisation theory in order to fully understand the workings of standard language norms and their effects.

Bibliography

Agosta, Sara, Valentina Ghirardi & Cristina Zogmaister. 2011. Detecting fakers of the autobiographical IAT. *Applied Cognitive Psychology* 25(2), 299–306.

Ammon, Ulrich. 1989. Towards a descriptive framework for the status/function (social position) of a language within a country. In Ulrich Ammon (ed.). *Status and Function of Languages and Language Varieties*. Berlin, New York: de Gruyter. 21–106.

Ammon, Ulrich. 1996. On comparing the centres of plurinational languages: The example of German. In Marlis Hellinger & Ulrich Ammon (eds). *Contrastive Sociolinguistics*. Berlin, New York: de Gruyter. 17–36.

Ammon, Ulrich. 2003a. On the social forces that determine what is standard in a language and on conditions of successful implementation. *Sociolinguistica* 17, 1–10.

Ammon, Ulrich. 2003b. *Sprachstandards: Language Standards*. Tübingen: Niemeyer.

Ammon, Ulrich. 2004. Standard variety / Standardvarietät. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds). *Sociolinguistics. An International Handbook of the Science of Language and Society.* Vol. 1. 2nd ed. Berlin, New York: de Gruyter. 273–83.

Auer, Peter. 2011. Dialect vs. standard: a typology of scenarios in Europe. In Bernd Kortmann & Johan van der Auwera (eds). *The Languages and Linguistics of Europe: A Comprehensive Guide*. Berlin, Boston: de Gruyter. 485–500.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1–48.

Bertels, Julie, Régine Kolonsky & José Morais. 2009. Norms of Emotional Valence, Arousal, Threat Value and Shock Value for 80 Spoken French Words: Comparison Between Neutral and Emotional Tones of Voice. *Psychologica Belgica* 49, 19–40.

Blair, Irene V. (2002). The Malleability of Automatic Stereotypes and Prejudice. *Personality and Social Psychology Review* 6(3), 242–61.

Blanton, Hart, James Jaccard & Christopgher N. Burrows. (2015). Implications of the Implicit Association Test D-Transformation for Psychological Assessment. *Assessment* 22(4), 429–40.

Blanton, Hart, James Jaccard, Jonathan Klick, Barbara Mellers, Gregory Mitchell & Philip E. Tetlock. (2009). Strong Claims and Weak Evidence: Reassessing the Predictive Validity of the IAT. *Journal of Applied Psychology* 94(3), 567–82.

Bonin, Patrick, Alain Méot, Lydie F. Aubert, Nathalie Malardier, Paula Niedenthal & Marie-Christin Capelle-Toczek. 2003. Normes de concrétude, de valeur d'imagerie, de fréquence subjective et de valence émotionnelle pour 866 mots. *L'Année psychologique* 103, 655–94.

Bradley, Margaret M. & Peter J. Lang. 1999. *Affective Norms for English Words (ANEW): Stimuli, Instruction Manual and Affective Ratings.* Technical report C-1, Gainesville, The Center for Research in Psychophysiology, University of Florida.

Brysbaert, Marc, Matthias Buchmeier, Markus Conrad, Arhur M. Jacobs, Jens Bölte & Andrea Böhl. 2011. The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology* 58, 412–24.

Campbell-Kibler, Kathryn. 2012. The Implicit Association Test and sociolinguistic meaning. *Lingua* 122, 753–63.

Campbell-Kibler, Kathryn. 2016. Towards a cognitively realistic model of meaningful sociolinguistic variation. In Anna M. Babel (ed.). *Awareness and Control in Sociolinguistic Research*. Cambridge: Cambridge University Press. 123–51.

Chaiken, Shelly. 1980. Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion. *Journal of Personality and Social Psychology* 39, 752–66.

Chen, Serena & Shelly Chaiken 1999. The Heuristic-Systematic Model in Its Broader Context. In Shelly Chaiken & Yaacov Trope (eds). *Dual-process Theories in Social Psychology*. New York: Guilford Press.

Clyne, Michael. 1992. *Pluricentric languages: Differing norms in different nations*. Berlin, New York: Mouton de Gruyter.

Combuchen, Jo. 2009. Deutsch in Belgien. *Lebende Sprachen* 53, 53–9.

Corson, Yves. & Quistrebert Virginie. 2000. Emotional judgment norms concerning the meanings of ambiguous words. *Current Psychology of Cognition:* 333–92.

Costa, James, Haley De Korne & Pia Lane. 2018. Standardising minority languages. Reinventing peripheral languages in the 21st century. In Pia Lane, James Costa & Haley De Korne (eds). *Standardizing Minority Languages: Competing Ideologies of Authority and Authenticity in the Global Periphery*. New York: Taylor & Francis. 1–23.

Coupland, Nikolas. & Tore Kristiansen. 2011. SLICE: Critical perspectives on language (de-) standardisation. In Tore Kristiansen & Nikolas Coupland (eds). *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus Press. 11–35.

Cummins, Jamie, Ian Hussey & Adriaan Spruyt. 2022. The role of attitude features in the reliability of IAT scores. *Journal of Experimental Social Psychology* 101, 1–8.

Darquennes, Jeroen. 2013. Deutsch als Muttersprache in Belgien: Forschungsstand und Forschungsperspektiven. In Karina Schneider-Wiejowski, Birte Kellermeier-Rehbein & Jakob Haselhuber (eds). *Vielfalt, Variation und Stellung der deutschen Sprache*. Berlin, Boston: de Gruyter. 343–62.

Darquennes, Jeroen. 2019. Komplexe Überdachung III: Belgien. In Joachim Herrgen & Jürgen Erich Schmidt (eds). *Language and Space: An International Handbook of Linguistic Variation* / Vol 4: *Deutsch*. Berlin, Boston: de Gruyter Mouton. 1060–76.

De Houwer, Jan. 2001. A Structural and Process Analysis of the Implicit Association Test. *Journal of Experimental Social Psychology* 37, 443–51.

De Houwer, Jan. & Agnes Moors. 2007. How to define and examine the implicitness of implicit measures. In Bernd Wittenbrink & Norbert Schwarz (eds). *Implicit Measures of Attitudes: Procedures and Controversies*. New York: Guilford Publications. 179–94.

De Houwer, Jan, Sarah Teige-Mocigemba, Adriaan Spruyt & Agnes Moors. 2009. Implicit measures: A normative analysis and review. *Psychological Bulletin* 135, 347–68.

Devine, Patricia. G. & Lindsay B. Sharp. 2009. Automaticity and control in stereotyping and prejudice. In Todd D. Nelson (ed.). *Handbook of Prejudice, Stereotyping, and Discrimination*. Hove: Taylor & Francis. 61–87.

Dolan, Kathleen. 2014. Gender stereotypes, candidate evaluations, and voting for women candidates: what really matters? *Political Research Quarterly* 67, 96–107.

Dörnyei, Zoltán & Tatsuya Taguchi. 2009. *Questionnaires in Second Language Research: Construction, Administration, and Processing*. Florence: Taylor & Francis.

Dovidio, John F., Miles Hewstone & Peter Glick. 2010. *The SAGE Handbook of Prejudice, Stereotyping and Discrimination*. London: SAGE Publications.

Durrell, Martin. 1999. Standardsprache in England und Deutschland. *Zeitschrift für Germanistische Linguistik* 27, 285–308.

Elspaß, Stephan & Stefan Kleiner. 2019. Forschungsergebnisse zur arealen Variation im Standarddeutschen. In Joachim Herrgen & Jürgen Erich Schmidt (eds). *Language and Space: An International Handbook of Linguistic Variation*. Vol 4: *Deutsch*. Berlin, Boston: de Gruyter Mouton. 159–84.

Entringer, Nathalie, Peter Gilles, Sara Martin & Christoph Purschke. 2018. [*Schnëssen-App – Är Sprooch fir d'Fuerschung*]. Unpublished raw data.

Fazio, Russel H. & Tamara Towles-Schwen. 1999. The MODE model of attitude-behavior processes. In Shelly Chaiken & Yaacov Trope (eds). *Dual-process Theories in Social Psychology*. New York: Guilford Press. 97–116.

Fehlen, Fernand. 2008. Multilingualismus und Sprachenpolitik. In Wolfgang H. Lorig (ed.). *Das politische System Luxemburgs. Eine Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften. 45–61.

Fehlen, Fernand. 2009. *BaleineBis: une enquête sur un marché linguistique multilingue en profonde mutation / Luxemburgs Sprachenmarkt im Wandel*. Luxembourg: SESOPI Centre intercommaunitaire.

Fehlen, Fernand. 2016. *Die Luxemburger Mehrsprachigkeit*. Bielefeld: transcript-Verlag.

Fiske, Susan T., Amy J. C. Cuddy, Peter Glick & Jun Xu. 2002. A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology* 82, 878–902.

Gast, Anne, & Klaus Rothermund (2010). When old and frail is not the same: Dissociating category and stimulus effects in four implicit attitude measurement methods. *Quarterly Journal of Experimental Psychology* 63(3), 479–98.

Gilles, Peter. 2015. From Status to Corpus: Codification and Implementation of Spelling Norms in Luxembourgish. In Winifred V. Davies & Evelyn Ziegler (eds). *Language Planning and Microlinguistics from Policy to Interaction and Vice Versa*. Basingstoke: Palgrave Macmillan. 128–50.

Greenwald, Anthony G., Brian A. Nosek & Mahzarin R. Banaji. 2003. Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology* 85, 197–216.

Greenwald, Anthony G., Debbie E. Mcghee & Jordan L. K. Schwartz. 1998. Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology* 74, 1464–80.

Greenwald, Anthony G., Miguel Brendl, Huajian Cai, Dario Cvencek, John F. Dovidio, Malte Friese, Adam Hahn, Eric Hehman, Wilhelm Hofmann, Sean Hughes, Ian Hussey, Christian Jordan, Teri. A. Kirby, Calvin K. Lai, Jonas W. B. Lang, Kristen P. Lindgren, Dominika Maison, Brian D. Ostafin, James R. Rae, Kate A. Ratliff, Adriaan Spruyt & Reinout W. Wiers. 2022. Best research practices for using the Implicit Association Test. *Behavior Research Methods* 54, 1161–80.

Grondelaers, Stefan & Roeland van Hout. 2016. How (in)coherent can standard languages be? A perceptual perspective on co-variation. *Lingua* 172–173, 62–71.

Haugen, Einar. 1968. Language planning in Modern Norway. In Joshua A. Fishman (ed.). *Readings in the Sociology of Language*. The Hague: Mouton. 673–87.

Haugen, Einar. 1997. Language standardization. In Nikolas Coupland & Adam Jaworski (eds). *Sociolinguistics: A Reader*. London: Macmillan. 341–52.

Herrgen, Joachim. 2015. Entnationalisierung des Standards. Eine perzeptionslinguistische Untersuchung zur deutschen Standardsprache in Deutschland, Österreich und der Schweiz. In Alexandra Lenz & Manfred M. Glauninger (eds). *Standarddeutsch im 21. Jahrhundert. Theoretische und empirische Ansätze mit einem Fokus auf Österreich*. Göttingen: Vienna University Press. 139–64.

Hoffmeister, Toke. 2021. Sprachkonzepte in der Öffentlichkeit. Kognitive Repräsentationen der deutschen Sprache. In Toke Hoffmeister, Markus

Hundt & Saskia Naths (eds). *Laien, Wissen, Sprache: Theoretische, methodische und domänenspezifische Perspektiven*. Berlin, Boston: de Gruyter. 71–104.

Hoffmeister, Toke, Markus Hundt & Saskia Naths. 2021. Theorien, Methoden und Domänen der Folk Linguistics im deutschsprachigen Raum. Einleitung in den Band. In Toke Hoffmeister, Markus Hundt & Saskia Naths (eds). *Laien, Wissen, Sprache: Theoretische, methodische und domänenspezifische Perspektiven*. Berlin, Boston: de Gruyter, VII–XXII.

Horner, Kristine. 2009. Language policy mechanisms and social practices in multilingual Luxembourg. *Language Problems and Language Planning* 33, 101–11.

Horner, Kristine & Jean-Jacques Weber. 2008. The Language Situation in Luxembourg. *Current Issues in Language Planning* 9, 69–128.

Horner, Kristine & Jean-Jacques Weber. 2015. Multilingual education and the politics of language in Luxembourg. In Catharina Peersman, Gijsbert Rutten & Rik Vosters (eds). *Past, Present and Future of a Language Border. Germanic-Romance Encounters in the Low Countries*. Berlin, Boston: de Gruyter. 233–53.

Huesmann, Anette. 1998. *Zwischen Dialekt und Standard. Empirische Untersuchung zur Soziolinguistik des Varietätenspektrums im Deutschen*. Tübingen: Niemeyer.

Hundt, Markus. 2018. Wahrnehmungsdialektologie – quo vadis? In Alexandra N. Lenz & Albrecht Plewnia (eds). *Variation – Normen – Identitäten*. Berlin, Boston: de Gruyter. 99–126.

Ianos, Maria Adelina, Andrei Rusu, Àngel Huguet & Cecilio Lapresta-Rey. 2020. Implicit language attitudes in Catalonia (Spain): investigating preferences for Catalan or Spanish using the Implicit Association Test. *Journal of Multilingual and Multicultural Development*: 1–16.

Kahneman, Daniel & Amos Tversky. 1972. Subjective probability: A judgment of representativeness. *Cognitive Psychology* 3, 430–54.

Kehrein, Roland. 2012. *Regionalsprachliche Spektren im Raum: Zur linguistischen Struktur der Vertikale*. Stuttgart: Steiner.

Kiesewalter, Carolin. 2019. *Zur subjektiven Dialektalität regiolektaler Aussprachemerkmale des Deutschen*. Stuttgart: Steiner.

Kircher, Ruth. 2016. The Matched-Guise Technique In Zhu Hua (ed.). *Research Methods in Intercultural Communication. A Practical Guide*. Chichester: Wiley. 196–211.

Kircher, Ruth & Lena Zipp. 2022. *Research Methods in Language Attitudes*. Cambridge: Cambridge University Press.

Kircher, Ruth & Sue Fox. 2019. Attitudes towards Multicultural London English: implications for attitude theory and language planning. *Journal of Multilingual and Multicultural Development* 40, 847–64.

Kiss, Gábor R., Christine A. Armstrong, R. Milroy & J. Piper. 1973. An associative thesaurus of English and its computer analysis. In Adam J. Aitken, Richard W. Bailey & Neil Hamilton-Smith (eds). *The Computer and Literary Studies*. Edinburgh: University Press. 153–65.

Kutlu, Ethan, Mehrgol Tiv, Stefanie Wulff & Debra Titone. 2022. Does race impact speech perception? An account of accented speech in two different multilingual locales. *Cognitive Research: Principles and Implications* 7, 7.

Lahl, Olaf, Anja S. Göritz, Reinhard Pietrowsky & Jessica Rosenberg. 2009. Using the World-Wide Web to obtain large-scale word norms: 190,212 ratings on a set of 2,654 German nouns. *Behavior Research Methods* 41, 13–9.

Lambert, Wallace E., Hannah Frankel & G. Richard Tucker. 1966. Judging Personality Through Spech: A French-Canadian Example. *Journal of Communication* 16, 305–21.

Lang, Peter J., Margaret M. Bradley & Bruce N. Cuthbert. 2008. *International affective picture system (IAPS): Affective Ratings of Pictures and Instruction Manual*. Technical Report A-8. Gainesville: University of Florida.

Legros, Sophie & Beniamino Cislaghi. 2020. Mapping the social-norms literature: An overview of reviews. *Perspectives on Psychological Science* 15(1), 62–80.

Lehnert, Tessa Elisabeth & Thomas Hörstermann. 2019. The Role of Implicit Nationality Preference in Speaker Evaluations in the Multilingual Context of Montreal. *Journal of Language and Social Psychology* 38, 283–304.

Lehnert, Tessa Elisabeth, Sabine Krolak-Schwerdt & Thomas Hörstermann. 2018a. Explicit and implicit speaker evaluations and their differential attitudinal determinants. *Language Sciences* 69, 68–79.

Lehnert, Tessa Elisabeth, Sabine Krolak-Schwerdt & Thomas Hörstermann. 2018b. Language and nationality attitudes as distinct factors that influence speaker evaluations: Explicit versus implicit attitudes in Luxembourg. *Language & Communication* 61, 58–70.

Lenz, Alexandra N. 2003. *Struktur und Dynamik des Substandards*. Stuttgart: Steiner.

Lenz, Alexandra N. & Albrecht Plewnia (eds). 2010. *Grammar between Norm and Variation*. Frankfurt am Main: Lang.

Mattheier, Klaus J. 2003. German. In Ana Deumert & Wim Vandenbussche (eds). *Germanic Standardizations: Past to Present*. Amsterdam, Philadelphia: Benjamins. 211–44.

Meissner, Franziska, Laura Anne Grigutsch, Nicolaus Koranyi, Florian Müller & Klaus Rothermund. 2019. Predicting Behavior with Implicit Measures: Disillusioning Findings, Reasonable Explanations, and Sophisticated Solutions. *Frontiers in Psychology* 10, 1–16.

Melinger, Alissa & Andrea Weber. 2006. *Database of Noun Associations for German*. <http://www.psycholing.es.uni-tuebingen.de/nag/index.php> accessed 8 June 2024.

Möller, Robert. 2017. Deutsch in Ostbelgien – ostbelgisches Deutsch? In Melanie Wagner, Winifred Davies, Annelies Häcki Buhofer, Regula Schmidlin & Eva Lia Wyss (eds). *Standardsprache zwischen Norm und Praxis. Theoretische Betrachtungen, empirische Studien und sprachdidaktische Ausblicke*. Tübingen: Narr. 89–123.

Möller, Robert, Sandra Weber, Sabrina Lander & Melanie Wirtz 2013. *Dialektatlas der Deutschsprachigen Gemeinschaft*. <https://www.ostbelgienkulturerbe.be/desktopdefault.aspx/tabid-3532/> accessed 8 June 2024.

Moors, Agnes, Jan De Houwer, Dirk Hermans, Sabine Wanmaker, Kevin Van Schie, Anne-Laura Van Harmelen, Maarten De Schryver, Jeffrey De Winne & Marc Brysbaert. 2013. Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods* 45, 169–77.

Neises, Diane. 2013. *Levelling toward a Higher Standard? A Study on Dialect Perception and its Potential Implications for Language Change in Luxembourg*. Unpublished Master's Thesis, University of York.

Nelde, Peter & Jeroen Darquennes. 2002. German in Belgium: Linguistic Variation from a Contact Linguistic Point of View. *Journal of Multilingual and Multicultural Development* 23, 65–79.

Nelde, Peter H. 1987. Language contact means language conflict. *Journal of Multilingual and Multicultural Development* 8, 33–42.

Nelson, Douglas L., Cathy L. Mcevoy & Thomas A. Schreiber. 1998. *The University of South Florida Word Association, Rhyme, and Word Fragment Norms*. <http://www.usf.edu/FreeAssociation/> (accessed 30 November 2024).

Newton, Gerald. 1996. *Luxembourg and Lëtzebuergesch*. Oxford: Clarendon Press.

Osgood, Charles E. 1952. The nature and measurement of meaning. *Psychological Bulletin* 49, 197–237.

Pantos, Andrew J. & Perkins, Andrew W. 2012. Measuring implicit and explicit attitudes toward foreign accented speech. *Journal of Language and Social Psychology* 32, 3–20.

Preston, Dennis R. 1989. Standard English spoken here: The geographical loci of linguistic norms. In Ulrich Ammon (ed.). *Status and Function of Languages and Language Varieties*. Berlin, New York: de Gruyter. 324–54.

Preston, Dennis R. 1999. A Language Attitude Approach to the Perception of Regional Variety. In Dennis R. Preston (ed.). *Handbook of Perceptual Dialectology*. Amsterdam: Benjamins. 359–73.

Preston, Dennis R. 2010. Language, space and the folk. In Peter Auer & Jürgen Erich Schmidt (eds). *Language and Space*. Vol 1. *Theories and Methods*. Berlin, Boston: de Gruyter. 179–201.

Preston, Dennis R. 2016. Whaddayaknow now? In Anna M. Babel (ed.). *Awareness and Control in Sociolinguistic Research*. Cambridge: Cambridge University Press. 177–99.

Preston, Dennis R. 2018. Language Regard: What, why, how, whither? In Betsy E. Evans, Erica J. Benson & James N. Stanford (eds). *Language Regard: Methods, Variation and Change*. Cambridge: Cambridge University Press. 3–28.

Preston, Dennis R. & Nancy A. Niedzielski. 2013. Approaches to the study of Language Regard. In Tore Kristiansen (ed.). *Language (De)standardisation in Late Modern Europe: Experimental Studies*. Oslo: Novus Press. 307–29.

Purschke, Christoph. 2011. *Regionalsprache und Hörerurteil. Grundzüge einer perzeptiven Variationslinguistik*. Stuttgart: Steiner.

Purschke, Christoph. 2018. Language Regard and cultural practice: Variation, evaluation, and change in the German regional languages. In Betsy E. Evans, Erica J. Benson & James N. Stanford (eds). *Language Regard: Methods, Variation and Change*. Cambridge: Cambridge University Press. 249–65.

Rosch, E. & Lloyd, B. B. 1978. *Cognition and Categorization*, Hillsdale: Wiley.

Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson & Penny Boyes-Braem. 1976. Basic objects in natural categories. *Cognitive Psychology* 8, 382–439.

Rosseel, Laura. 2022. The Implicit Association Test Paradigm. In Ruth Kircher & Lena Zipp (eds). *Research Methods in Language Attitudes*. Cambridge: Cambridge University Press, 250–68.

Rosseel, Laura & Stefan Grondelaers. 2019. Implicitness and experimental methods in language variation research. *Linguistics Vanguard* 5, 1–6.

Rosseel, Laura, Dirk Speelman & Dirk Geeraerts. 2018. Measuring language attitudes using the Personalized Implicit Association Test: A case study on regional varieties of Dutch in Belgium. *Journal of Linguistic Geography* 6, 20–39.

Ryan Bouchard, Ellen, Howard Giles & Richard J. Sebastian. 1982. An integrative perspective for the study of attitudes toward language variation. In Ellen Ryan Bouchard & Howard Giles (eds). *Attitudes towards Language Variation. Social and Applied Contexts*. London: E. Arnold. 1–20.

Schimmack, Ulrich. (2021). The Implicit Association Test: A Method in Search of a Construct. *Perspectives on Psychological Science* 16(2), 396–414.

Schmidlin, Regula. 2011. *Die Vielfalt des Deutschen: Standard und Variation Gebrauch, Einschätzung und Kodifizierung einer plurizentrischen Sprache*. Berlin, Boston: de Gruyter.

Schmidlin, Regula, Eva Lia Wyss, & Winifred Davies. 2017. Plurizentrik revisited – aktuelle Perspektiven auf die Variation der deutschen Standardsprache. In Melanie Wagner, Winifred Davies, Annelies Häcki Buhofer, Regula Schmidlin & Eva Lia Wyss (eds). *Standardsprache zwischen Norm und Praxis. Theoretische Betrachtungen, empirische Studien und sprachdidaktische Ausblicke*. Tübingen: Narr. 7–23.

Schmidt, Jürgen Erich & Joachim Herrgen. 2011. *Sprachdynamik: Eine Einführung in die moderne Regionalsprachenforschung*. Berlin: Schmidt.

Schoel, Christiane, Jennifer Eck, Janin Roessel & Dagmar Stahlberg. 2012. Spracheinstellungen aus sozialpsychologischer Perspektive I: Deutsch und Fremdsprachen. In Ludwig M. Eichinger & Gerhard Stickel (eds). *Sprache und Einstellungen Spracheinstellungen aus sprachwissenschaftlicher und sozialpsychologischer Perspektive*. Tübingen: Narr. 163–204.

Schoel, Christiane & Dagmar Stahlberg. 2012. Spracheinstellungen aus sozialpsychologischer Perspektive II: Dialekte. In Ludwig M. Eichinger & Gerhard Stickel (eds). *Sprache und Einstellungen Spracheinstellungen aus sprachwissenschaftlicher und sozialpsychologischer Perspektive*. Tübingen: Narr. 205–27.

Schoel, Christiane, Janin Roessel, Jennifer Eck, Jana Janssen, Branislava Petrovic, Astrid Rothe, Selma Carolin Rudert & Dagmar Stahlberg. 2012. 'Attitudes Towards Languages' (AToL) Scale. *Journal of Language and Social Psychology* 32, 21–45.

Speelman, Dirk, Adriaan Spruyt, Leen Impe & Dirk Geeraerts. 2013. Language attitudes revisited: Auditory affective priming. *Journal of Pragmatics* 52, 83–92.

Stell, Gerald. 2006. *Luxembourgish Standardization. Context, Ideology and Comparability with the Case of West Frisian*. Louvain-la-Neuve: Peeters.

Tajfel, Henri. 1982. Social Psychology of Intergroup Relations. *Annual Review of Psychology* 33, 1–39.

Turner, John C. 2010. Towards a cognitive redefinition of the social group. In Tom Postmes & Nyla R. Branscombe (eds). *Rediscovering Social Identity*. New York: Psychology Press. 210–34.

Tversky, Amos & Daniel Kahneman. 2004. Judgement under Uncertainty. Heuristics and Biases. In Elizabeth J. Marsh & David A. Balota (eds). *Cognitive Psychology: Key Readings*. Hove: Psychology Press. 631–42.

Van Heuven, Walter J. B., Pawel Mandera & Emmanuel Keuleers. 2014. SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology* 67, 1176–90.

Vári, Judit & Marco Tamburelli. 2021. Accepting a 'New' Standard Variety: Comparing Explicit Attitudes in Luxembourg and Belgium. *Languages* 6, 134.

Vári, Judit & Marco Tamburelli. 2023. Standardisation: bolstering positive attitudes towards endangered language varieties? Evidence from implicit attitudes. *Journal of Multilingual and Multicultural Development* 44(6), 447–66.

Võ, Melissa, Markus Conrad, Lars Kuchinke, Karolina Urton, Markus Hofmann & Arthur Jacobs. 2009. The Berlin Affective Word List Reloaded (BAWL-R). *Behavior Research Methods* 41, 534–8.

Walker, Abbie, Jen Hay, Katie Drager & Kauyumari Sanchez. 2018. Divergence in speech perception. *Linguistics* 56, 257–78.

Walker, Abbie, Katie Drager & Jen Hay. 2022. The use of priming in language attitudes research. In Ruth Kircher & Lena Zipp (eds). *Research Methods in Language Attitudes*. Cambridge: Cambridge University Press. 313–29.

Warriner, Amy Beth, Victor Kuperman. & Marc Brysbaert. 2013. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods* 45, 1191–207.

Weber, Sandra. 2009. *Dialekt in Ostbelgien, Nordrhein-Westfalen und Rheinland-Pfalz. Eine Untersuchung zu regionalen und nationalen Unterschieden in der Verbreitung des Dialekts und den Dialektattitüden, verglichen mit der Sprachsituation in Luxemburg.* Unpublished master's thesis, Université de Liège.

Wilson, Timothy D., Samuel Lindsey & Tonya Y. Schooler. 2000. A model of dual attitudes. *Psychological Review* 107, 101–26.

Wolsiefer, Katie, Jacob Westfall & Charles M. Judd. 2017. Modeling stimulus variation in three common implicit attitude tasks. *Behavior Research Methods* 49(4), 1193–209.

Zahn, Christopher J. & Robert Hopper. 1985. Measuring language Attitudes: The Speech Evaluation Instrument. *Journal of Language and Social Psychology* 4, 113–23.

VIKTORIJA KOSTADINOVA

7 'Double standards' in American English

ABSTRACT

The present study compares standards of usage prescribed in the metalinguistic literature with patterns of actual use of a number of linguistic features in American English in order to investigate the extent to which prescriptions align with actual language use. Specifically, the study investigates a collection of usage guides in order to identify specific prescriptions distinguishing standard from non-standard linguistic features and subsequently explores the patterns of use of these features in a set of corpora. In addition to exploring how prescribed standards compare with standards of use across traditional written texts, the study also includes the perspective of spoken, as well as online language use. The results suggest that prescriptions sometimes align with actual use, which happens to be the case either for features which are salient as non-standard, or for features which do not vary greatly in the first place. For other features, a situation of 'double standards' arises, which can be observed through a comparison between written edited texts and spoken or online language use rather than by comparing prescriptions and usage. Finally, one case also provides evidence of changing prescriptions in relation to changes in language use.

1. Introduction[1]

The American English variety, which followed the post-colonial path of development (Schneider 2007: 251–306), with distinctive linguistic

1 I would like to acknowledge the support of the Bridging the Unbridgeable project at Leiden University, which I was a part of between 2012 and 2016, in obtaining the raw data for the Hyper Usage Guide of English database and purchasing the full-text data of COCA and COHA. I thank the two anonymous reviewers for their useful and constructive comments on a previous version of this paper, and my colleagues Imogen Cohen and Rose van der Zwaard for providing stylistic advice. Any remaining errors are my own.

features arising through processes of koineisation, supraregionalisation, and levelling, followed by diversification, became reified into a standard variety through a process of standardisation. The emergence of the ideology of the standard, and its inextricable link to the emergence of a new national identity, was critical in giving the new standard its legitimacy, very much as it did in the European context (Haugen 1966). The role that ideology plays in the rise of standard languages in general, and in the emergence of Standard American English in particular, explains why Standard American English is seen as a myth (Lippi-Green 2012) and an ideological construct (Kretzschmar 2010), tied to a set of norms which, in the most extreme view, are seen as being at odds with the way in which speakers actually use language. However, Standard American English is also used to refer to the variety which has been investigated in descriptive analyses of language variation and change (e.g. Mair 2006; Grieve 2012; Hinrichs et al. 2015). In this context, then, the variety is not an ideological construct that does not exist in reality – or at least not more so than any other variety in a conceptual linguistic sense (cf. Schneider 2021). Pointing out these two different ways in which Standard American English can be conceptualised is not meant to suggest a preference for either, but to make the point that language standards can be both ideological and usage based. Furthermore, the ideas about standard variants and the patterns of use of these variants are continually in interaction. It is this interaction between prescribed standards and standards in language use in American English that is the focus of the present paper.

By prescribed standards, I am referring to 'rules/judgements that aim to promote and enforce standardisation and "standard" usage' (Curzan 2014: 24), which can be identified in normative or metalinguistic sources during the stage of prescription, as opposed to the stage of codification. The stage of prescription is one of the stages in the process of standardisation, usually following the stage of codification (cf. Haugen 1966; Milroy & Milroy 1985). Though distinct, these two processes are closely related, and in some cases known to overlap. Ayres-Bennett (2020: 185), for instance, observes that codification and prescription are not explicitly distinguished in the first version of Haugen's model. Milroy & Milroy (1985), on the other hand, distinguish between codification and prescription, and, following this, historical-sociolinguistic analyses of the development of Standard British English, for instance, have distinguished between the stage of codification and the stage of prescription (Tieken-Boon van Ostade 2019: 8; Yáñez-Bouza 2021).

When it comes to examining the relationship between standards and language use, different sets of questions tend to be relevant in the codification stage as opposed to the prescription stage. While studying the stage of codification may be concerned predominantly with the selection of standards and principles upon which such selection takes place, studying the development of standard varieties during the prescription stage raises a different set of questions. Once the standards are codified, and the idea of the standard variety has become entrenched, other processes unfold. These include the development of and participation in language prescription, as well as dealing with changes in the standard, both ideological and linguistic.

In the context of American English, the codification stage is associated most notably with the efforts of Noah Webster, especially with respect to his influence on orthography, and Lindley Murray in the context of grammatical codification (Kretzschmar 2010; Kretzschmar & Meyer 2012). The prescription stage is characteristic of the twentieth and the early twenty-first century and the proliferation of usage guides. The stage of prescription is characterised by the vitality of the ideology of the standard, and, as Kretzschmar & Meyer (2012: 140) put it, Americans' participation 'in the cultural practice of believing in a language standard'. Though the ideology of the standard is still strong during the prescription stage, both the ideology and the use of language are constantly in flux. While at an aggregate level the standard ideology may remain robust, specific instances of linguistic features associated with the standard, as well as the context of use in which standard features are used, may change.

One type of change arises from the fact that the prescribed standards are not always observed in actual language use, which can then lead to a case of 'double standards': a split between the norms that are codified as constituting the standard variety and the norms of language in actual use. This has emerged as one of the main findings of a number of studies exploring the connection between codified and prescribed norms and patterns of use in the context of evaluating the success of prescriptivism in American English (e.g. Albakry 2007; Anderwald 2012, 2013, 2014; Hinrichs et al. 2015; Kostadinova 2020). Many of these analyses have shown that there is a mismatch between prescriptions and use, with some evidence of alignment between prescriptions and practice in specific written edited genres or text types. This can be linked to double standards: a situation where what is described as the standard and what is actually used are different.

The ideas about what constitutes the standard can also change, raising the question of potential changes in prescriptions. As Hickey (2012: 15–6) notes, 'it is a truism that standard forms of English continually change' which can 'have implications for prescriptivists defending the immutability of standard English'. Addressing this issue in some way, Dant (2012) examines the extent to which prescriptions in the *Chicago Manual of Style* align with usage and, based on this, makes recommendations for how these prescriptions should be changed. With respect to a few of the prescriptions, Dant suggests that they be dropped altogether as they appear to be universally disregarded in actual use, which also indicates the existence of double standards. While it is too early to obtain evidence of the extent to which usage guide writers are following Dant's suggestions, there are other cases that provide an indication of prescriptions changing in relation to changing patterns of use. Kostadinova (2020) has also shown that usage guide pronouncements on the split infinitive, for instance, have changed over time towards a greater acceptance of the construction as part of the standard. This means that any investigation of the interaction between prescribed standards and actual language use needs to take into account a bidirectional relationship between the two.

Taking these points of departure, the present study aims to contribute empirical insights into the relationship between standard norms prescribed in the metalinguistic literature and patterns of language use with respect to a set of linguistic features. The research questions are (a) what is considered standard in the metalinguistic literature? And (b) are these observations in line with actual use? In order to answer the first question, I investigate a collection of usage guides, identifying how standard and non-standard linguistic variants are described. For the second question, I explore the usage patterns of these features in a set of corpora. In addition to exploring how prescribed standards compare with standards of usage across traditional written texts, I also include the perspective of spoken, as well as online language use. In present-day American English, online language use has emerged as a specific context of use in its own right, which creates new contexts in which people from different regions interact, providing the opportunity for investigating the extent to which speakers observe standards in informal written communication.

2. Method

The analysis is based on a comparison between prescriptions identified in metalinguistic texts about a set of linguistic variants and the patterns of actual use of those variants in language corpora, described in more detail below. The inclusion of evidence from online language is meant to provide insights into how language is used in interaction in informal written contexts, which are similar to colloquial speech to some extent. Since many of the linguistic variants involved in delimiting standard from non-standard usage are morphosyntactic in nature and, as such, tend to be low in frequency, they can be difficult to investigate with traditional spoken corpora. Online language use can serve as a good alternative source of data, since datasets of sufficient size can be extracted and used for analysing morphosyntactic variants.

The source of prescriptions in metalinguistic texts used for the present analysis is the *Hyper Usage Guide of English* database, a collection of usage guide entries covering 123 usage problems (i.e. linguistic features or variants) extracted from seventy-seven usage guides, British or American, published between 1847 and 2008. There are 6,247 entries (or 572,871 tokens) in the database, 2,855 of which are from American usage guides, which is the subset explored in the present study. The database has some limitations worth keeping in mind. For instance, some usage problems (e.g. *that/which*, or *shall/will*) are covered much more often than others (e.g. the use of *momentarily*) and in some cases the same entry may cover multiple usage problems. Despite these limitations, the database provides a good source of data for the comparative analysis of metalinguistic commentary across a high number of linguistic features.

In order to identify which of the 123 linguistic features tend to be described in terms of the standard/non-standard distinction, I adopt an approach which involves the analysis of the use of metalinguistic terms similar to Chapman (2019). This entails searching the entries for a set of metalinguistic terms which are often used in prescriptions of the standard rules and norms. These items were identified in a bottom-up way by first extracting all nouns and adjectives from the texts (using the Python library 'spaCy' for NLP-based text annotation) and subsequently selecting manually those nouns and adjectives (given in Table 7.1) which refer to the rules of the standard language. As Table 7.1 shows, I identify four categories: (a) terms which directly distinguish between standard and non-standard variants (standardness), (b) terms which distinguish between correct and incorrect variants (correctness), (c) terms which refer to rules (rules) and d) terms which distinguish levels of formality (formality). As Chapman (2019: 40) notes, the dimension of formality can be seen as the level of the 'hyperstandard': in addition to the basic level at which the rules of the standard are established as the variety used by educated speakers, there is an extra level at which rules are distinguished as being observed by 'some speakers'. In other words,

Table 7.1 Terms searched in the collection of entries.

STANDARDNESS	CORRECTNESS	RULES	FORMALITY
dialect	*correct*	*conventions*	*colloquial*
dialectal	*corrected*	*error*	*educated*
dialects	*incorrect*	*grammatical*	*formal*
region		*mistake*	*informal*
regional		*proper*	*speech*
regions		*rule*	*uneducated*
standard		*rules*	*writing*
standards		*wrong*	*written*
non-standard			
substandard			

within the rules of the standard we can distinguish a level that defines the rules of careful standard speakers. For this reason, in the analysis I combine the terms in the first three columns in Table 7.1 under the standardness label and distinguish those from the formality terms in the fourth column.

In this way, 4,765 occurrences of these terms were extracted from the database texts, and the linguistic features most often described by these terms were identified. This is an analysis of prescriptive or normative metalanguage (cf. Sundby et al. 1991; Peters & Young 1997), which is an account of prescriptivism in expression, but not in intention (cf. Ayres-Bennett 2020). In other words, while this approach may miss other ways in which standard forms are discussed, I am using it here as a way of tapping into the variation in metalinguistic discourse, and of arriving at an initial selection of features. This is not to say that the features which are not described in this way may not be associated with norms, as prescriptions may be normative in intention and effect, not only in expression.

In order to quantify the degree to which these terms are used for different linguistic features, I calculated the frequency of occurrence per thousand words of standardness terms and of formality terms for each of the linguistic features included in the database. This results in two rankings of linguistic features in terms of how often the two sets of terms occur in the entries. While this allows us to rank features across each dimension separately, much important detail about prescribed standards, necessary for a more meaningful comparison between prescriptions and patterns of actual use, is left out. In order to take this information into account, I select a smaller set of features for further investigation: *between you and me/I*, double negatives, *NP and I/me and NP* functioning as subject, *sneaked/snuck*, *dived/dove*, *try to/try and*, *this/these kind/sort of* and *whom/who* in object relative clauses (see Table 7.2 for an overview). For these features, I conducted a manual inspection of the statements associated with the terms selected for analysis (given in Table 7.1), considering how they are used in the entries; this is then followed by an analysis of patterns of variation of the linguistic features in the corpus data.

Table 7.2 Linguistic features analysed.

usage feature	description	examples of variants analysed; second option is considered incorrect or non-standard
all right/alright	spelling *all right* as one word	*He will be all right.* *He will be alright.*
ain't	the use of *ain't* as a variant of *be not* and *have not*	*I am not helping you.* *I ain't helping you.*
between you and me/I	the use of nominative pronoun forms in object syntactic contexts	*This is between you and me.* *This is between you and I.*
negative concord	the post-verbal use of a negative indefinite pronoun in combination with *not* on the verb	*I didn't see anybody.* *I didn't see nobody.*
NP and I/me and NP	the use of accusative pronoun forms in subject syntactic contexts	*My friends and I left early.* *Me and my friends left early.*
this/these kind/sort of	the use of *these kind/sort of* with plural nouns	*I hate this kind of decisions.* *I hate these kind of decisions.*
try to/try and	the use of *try and* as an alternative for *try to*	*I'll try to be there.* *I'll try and be there.*
sneaked/snuck *dived/dove*	the use of an irregular preterite form instead of the regular one	*We sneaked out early.* *We snuck out early.* *We dived into the pool.* *We dove into the pool.*
whom/who	the use of *who* as an alternative for *whom* in object relative clauses	*This is the person whom I saw.* *This is the person who I saw.*

Occurrences of the linguistic features were searched in three corpora: the *Corpus of Historical American English* (henceforth COHA; cf. Davies 2010, 2012), the *Corpus of Contemporary American English* (henceforth COCA; cf. Davies 2008–, 2010), as well as in a self-compiled corpus of online language use (the *Reddit Corpus*, henceforth RC).

For COHA and COCA, I used the full-text data covering the periods 1820–2000 and 1990–2012 respectively. The RC was compiled by extracting comments from the popular forum website Reddit, which can be described as an online community, to some extent similar to social networking sites, where users participate in forum discussions on specific topics. Discussions take place in so-called 'subreddits', subcommunities within Reddit where users who share similar interests, hobbies, identities, or issues, can interact with each other by initiating and participating in discussions. According to Reddit's official 2022 update, there are over fifty-seven million daily active users and over thirteen billion posts and comments, about half of which are from the United States. Reddit has been used as a source of data in a wide range of academic disciplines (cf. Proferes et al. 2021 for a recent systematic overview). For the purposes of compiling the RC, data was downloaded from the Pushshift Reddit Dataset (Baumgartner et al. 2020). Specifically, comments were extracted from seven popular subreddits, and, using information which users themselves provide about their location, only comments posted by users who state a US location as their location were selected. The resulting dataset consisted of 59,859,540 tokens spanning the period 2011–2016.

All corpus data was pre-processed with Python and annotated with the Python NLP library 'spaCy', which was used for pos-tagging and dependency parsing. The annotated data was then used to automatically extract cases of the linguistic features listed in Table 7.2, where the examples indicate the kinds of cases that were extracted. Despite the fact that spaCy's pos-tagger and parser meet the highest state-of-the-art accuracy rates, it is important to note that there is bound to be some noise when dealing with datasets of this size. The level of noise is negligible when dealing with orthographic or token-based variants, but it increases in the context of more complex syntactic constructions, though in the latter case, higher numbers of tokens can be expected to minimise the noise in the data to a degree comparable to that found in other pos-tagged or parsed corpora (cf. Kostadinova et al. in preparation). The findings are based on plotting descriptive trends in the use of the features over time, rather than testing specific hypotheses.

3. Results

3.1 Standard and non-standard variants in usage guides

The analysis of metalinguistic terms in the usage guide data reveals differences in the rate of occurrences of these terms across the various linguistic features. In order to identify the linguistic features that are more often described with terms referring to standards and norms, I calculated the frequency of occurrence of the terms across the entries for each linguistic feature included in the database and ranked the features from highest to lowest. Figures 7.1 and 7.2 show the top ten and bottom ten linguistic features in terms of frequency of occurrence of metalinguistic terms, split into terms referring to standardness (Figure 7.1), as described in Section 2, and formality (Figure 7.2). The linguistic features are listed using their labels as

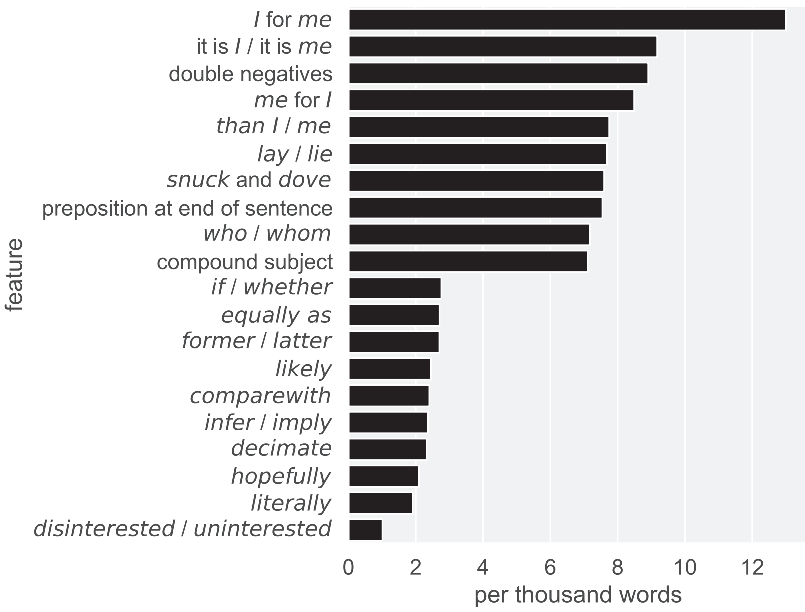


Figure 7.1 Top ten and bottom ten linguistic features in terms of rate of occurrence of STANDARDNESS terms.

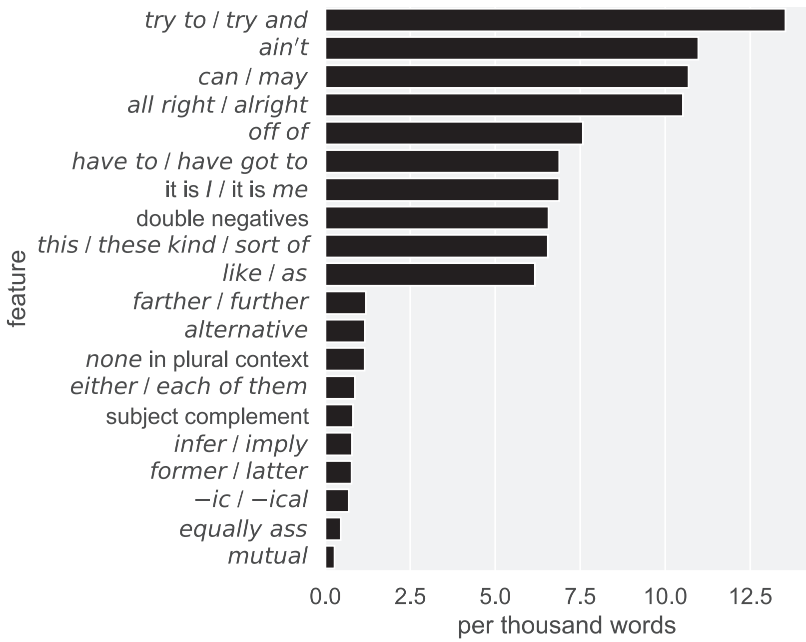


Figure 7.2 Top ten and bottom ten linguistic features in terms of rate of occurrence of FORMALITY terms.

usage problems in the database. It is important to note that some of these labels cover a broader area of usage, rather than a specific linguistic feature; in what follows, this is specified in cases where it is relevant for the analysis.

Figures 7.1 and 7.2 show that entries on the variation in pronoun use in coordinated subject and object NPs, as well as other cases in which pronominal form variation is observed, have the highest rates of terms relating to standardness. Double negatives, *lay/lie*, *snuck* and *dove*, preposition stranding and *who/whom* are also among the linguistic features whose entries have a high rate of occurrence of standardness terms. When it comes to formality terms, *try to/try* is the top feature, followed by *ain't*, *can/may*, *all right/alright* and the use of *off of*. Following this, it may be noted that while there seems to be a difference in the rankings across the two dimensions, there is also overlap: double negatives and *It is I/It is me* are high for both dimensions. Overall, it can also be observed that entries on usage problems that are morphosyntactic in nature have higher frequency of standardness and formality terms. On the bottom end of the ranking, the majority of usage problems tend to be semantic or syntactic in nature.

While these rankings provide an initial indication of linguistic features whose usage tends to be prescribed in relation to the standard, they leave out much important detail on the specific prescriptions made. As explained in Section 2, this analysis is further expanded with a close reading of the sentences in which the metalinguistic terms are used for a selection of linguistic features, given in Table 7.2 above. The closer reading of the sentences containing these terms reveals that, for some of the linguistic features, the association with the standard is stronger than for others. The use of *I* for *me* in coordinated object NPs is seen as incorrect, ungrammatical and not part of the standard in the majority of the entries, most of which cover the case of *between you and I* as a usage problem. The use of *me* for *I* in coordinated subject NPs, on the other hand, is covered less often in the entries, but when it is discussed, it is considered equally incorrect and unacceptable in Standard English. Prescriptions about both features are often accompanied by observations of their frequent use by 'careless' speakers.

Entries on double negatives cover three different constructions, each characterised in some way by the use of two negatives in the same clause. Apart from negative concord (e.g. *I haven't seen nobody*), the feature I focus on here, the term 'double negatives' also refers to litotes (e.g. *it is not unlikely*) and negation constructions with *hardly* (e.g. *I hardly have any money* vs *I hardly have no money*). Negative concord is predominantly described as a non-standard feature in present-day English. While some entries note that negative concord constructions are a means of expressing emphasis, especially in speech, others focus on the nonstandardness and informality of the feature. In other words, acceptability is restricted to specific non-standard contexts.

Entries on *snuck* and *dove* discuss the variation in preterite forms more generally but *sneaked/snuck* and *dived/dove* are two of the cases often covered (others include, for instance, *hung* and *hanged*). For *snuck* as a variant of *sneaked*, the majority of the observations agree that while it is dialectal, and tends to be typical of informal contexts, it is also spreading in use and becoming more common among educated speakers. Similar observations are made for the use of *dove* as a variant of *dived*, which is described in terms of a preference for the standard form in formal writing. For both forms, it is noted that the irregular forms are cases of changing standards, where what used to be a dialectal form has now developed into a feature acceptable as standard.

The case of *who*/*whom* is described in terms of an opposition between traditional rules of grammar and levels of formality, though, here too, the admissibility of *who* as a variant of *whom* is considered to apply primarily to spoken and informal written language. What the examined entries also note is the predominance of *who* in informal contexts, both spoken and written. The overall perspective, then, is that *whom* is restricted to formal written contexts, and that the strict grammatical distinction between the two forms is losing ground, with *who* taking over and becoming increasingly acceptable as part of the standard.

The majority of the entries discussing *try and* as a variant of *try to* characterise it as a feature of colloquial and informal style, noting only very occasionally that it is a non-standard feature. In the few cases where standards are referred to, it is specified that both *try to* and *try and* may be seen as standard, though with a stylistic difference between the two, with *try and* being associated with colloquial language. Some entries also make reference to a variety-related distinction, describing *try and* as more characteristic of the standard in British English.

Looking at the entries on *ain't* reveals a somewhat more complex picture than the one suggested by the rate of occurrence of metalinguistic terms. *Ain't* is definitely seen as non-standard and dialectal, but its use in specific contexts and for specific purposes tends to be seen as acceptable in fiction and in speech. This likely explains why terms associated with formality are also frequently used in entries on *ain't*. *Ain't* is also described as acceptable in jocular expressions, or when the speaker is trying to signal proximity or approachability.

With respect to *alright*, the majority of the entries note that the expression is never used in writing, with *all right* being the standard form. Some concessions seem to be granted to the use of *alright* in colloquial informal writing, but, on the whole, it is considered an error, and especially so in written edited prose.

Finally, the majority of the entries describe the use of these *kind/sort of with* plural nouns as incorrect and not acceptable in writing, while also noting that the construction tends to be used in informal and colloquial contexts. Its acceptability in speech is noted, but the consensus seems to be that *these kind/sort* of as a variant for *this kind/sort* of with plural nouns is ungrammatical and not part of the standard.

For all of these cases of variation, then, the general approach seems to be to describe the grammatical rule as the correct option, and to consider the non-standard variant as common or somewhat acceptable in informal colloquial speech. There are, however, some differences in how these features are treated, especially in the extent to which the grammatically 'incorrect' variant is accepted. In the case of *snuck* and *dove*, *who/whom* and *try to/try and*, the acceptability of the 'incorrect' variants seems to be fairly high. This is the strongest for *snuck* and *dove*, which seem to be growing in acceptance as part of the standard. The same goes for *who*, to some extent, though *whom* is noted as being the preferred form in writing. In the case of *try and*, its acceptance as part of the standard is seen as more characteristic of British English, but its use is not considered to be a serious error to the same extent as is the use of *between you and I* or *these kind/sort* of with plural nouns. Subjective pronoun forms in coordinated object NPs and objective pronoun forms in coordinated subject NPs, as well as *alright* and *these kind/sort of* are seen more often as errors and not considered part of the standard. Finally, *ain't* and negative concord are clear cases of non-standard features that have specific uses in speech, as well as in some kinds of writing. The overall finding here is that some of the features are described in terms of changing standards, and incoming forms are more readily accepted (*sneaked/snuck*, *dived/dove*, *whom/who* and to some extent *try to/try and*), while others are seen as non-standard and less acceptable (pronoun forms in coordinated NPs, *this/these kind/sort of*, *all right/alright, ain't* and negative concord).

3.2 Patterns in actual language use

The final step in the analysis involved investigating the patterns of actual use of the linguistic features selected for further analysis. On the basis of the summary of findings at the end of Section 3.1 above, I first cover the features that tend to be associated with changing standards (*sneaked/snuck*, *dived/dove*, *whom/who* and *try to/try and*), followed by those that are seen predominantly as non-standard (*between you and me/I*, *NP and I/me and NP* in subject positions, *all right/alright*, *this/these kind/sort of*, *ain't* and negative concord). Data are plotted across the following corpus sections: the four COHA subsections (Non-Fiction, Fiction, Magazine and Newspaper), the five COCA subsections (Academic, Fiction, Magazine, Newspaper and Spoken) and RC.

The analysis of *sneaked/snuck* and *dived/dove* (Figure 7.3 and 7.4) reveals that the occurrences of both of these verbs in the Spoken section

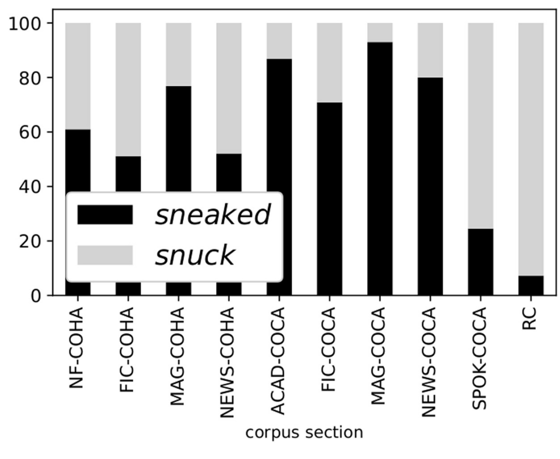


Figure 7.3 Percentage of cases of *sneaked* (n = 2,202) vs *snuck* (388) across corpus sections.

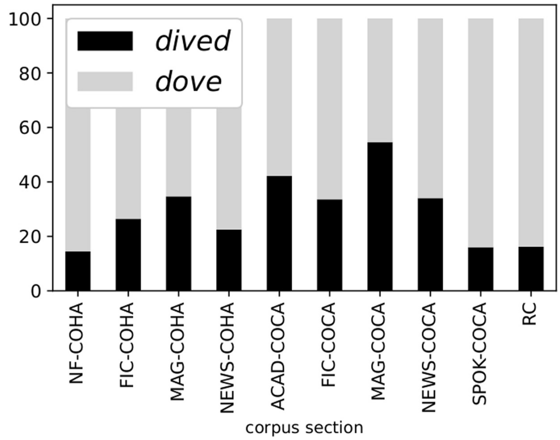


Figure 7.4 Percentage of cases of *dived* (n = 1,296) vs *dove* (n = 2,814) across corpus sections.

of COCA and the RC data are predominantly realised by the irregular forms. Looking at how this compares to the other corpus sections reveals a difference between the two verbs. In the case of *sneaked/snuck*, the written COCA sections all contain a high percentage of the regular forms, while in the COHA sections the percentages are slightly lower. The Spoken section in COCA has almost 80% of *snuck* (cf. similar results based on rate of occurrence in Anderwald 2016: 86), and the RC data contain the highest percentage of *snuck*, indicating that the change in colloquial language use may be near completion. In the case of *dived/dove*, despite some variability, and a somewhat higher percentage of use of *dived* in the written sections of COCA, it appears that the irregular form has become the more commonly used variant across most contexts of use.

In the case of *whom/who* (Figure 7.5), cases of object relative clauses found in the Spoken section of COCA and the RC data predominantly use *who* as opposed to *whom*. The opposite seems to be the case for the other corpus sections. Fiction texts in both COCA and COHA seem to have a somewhat higher percentage of *who* cases than the rest of the written corpus sections, but the percentage of *whom* cases are still high. These results seem to align with observations that *whom* is becoming

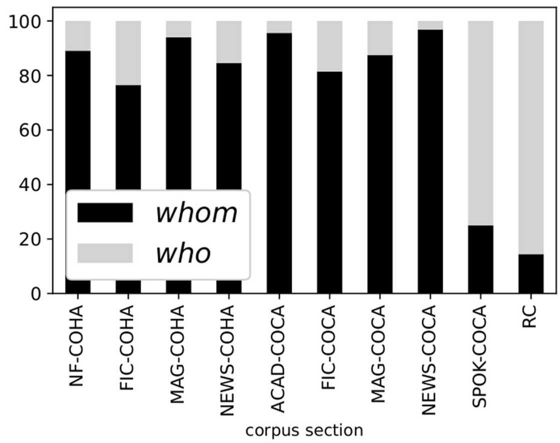


Figure 7.5 Percentage of cases of *whom* (n = 34,940) vs *who* (n = 7,273) in object relative clauses across corpus sections.

uncommon in spoken language and is increasingly being replaced by *who* (Denison 1998: 226–7), but also that it retains a certain level of vitality, because it is a useful style marker (cf. Mair 2006: 141–3). As Mair (2006: 143) further notes, the change from *whom* to *who* appears to be near complete in colloquial language, while *whom* remains robustly preferred in written contexts.

The third linguistic feature for which both variants tend to be accepted as standard is *try to*/*try and*. The distribution of percentages shown in Figure 7.6 reveals that *try to* is the predominantly used variant across all corpus sections. This may align with previous findings that *try and* is more characteristic of British, rather than American English (Biber et al. 1999: 738; Hommerber & Tottie 2007), but goes against the observation that both variants appear to be common in conversation; indeed, Hommerberg & Tottie (2007) find that *try to* is more frequent in writing, regardless of variety. The Spoken section in COCA and the RC data contain marginally higher proportions of *try and* cases, but this is a small difference. On the whole, the distribution appears stable across the various sections, indicating that there is not much difference in terms of context of use.

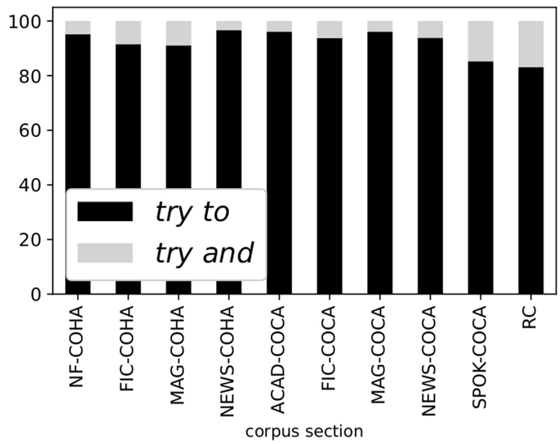


Figure 7.6 Percentage of cases of *try to* (n = 110,253) vs *try and* (n = 12,513) across corpus sections.

The patterns of use of nominative and accusative forms of pronouns in coordinated object NPs, investigated on the basis of cases of the prepositional phrase *between you and me/I*, can be seen in Figure 7.7. The figure shows the percentage of cases across the different corpus sections, illustrating that the spoken section in COCA and the data from the RC have the highest proportion of use of *between you and I* as opposed to the standard *between you and me*. The Spoken section in COHA contains just under 40% of the cases coded as *between you and I*, and this percentage increases to close to 60% for the RC data. These patterns indicate a different distribution for the use of *I* as a variant of *me* in coordinated object NPs and reveal a percentage much higher than the 4% found by Boyland (2001), while confirming to some extent Denison's (1998: 109) observation that pronoun variation in these syntactic contexts remains common in the speech of educated speakers. In contrast with *between you and me/I*, a much larger number of cases of *NP and I / me and NP* in subject positions (42,866) was extracted from the datasets; the percentage of cases of the two variants is given in Figure 7.8. Little difference can be observed in the distribution of the two variants in terms of corpus sections, with instances of *me and NP* being rare across the board, indicating little variation.

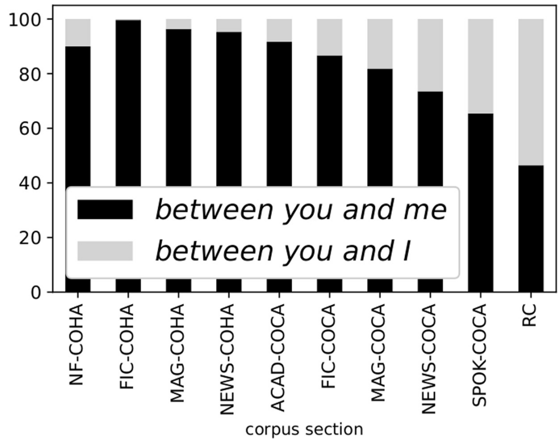


Figure 7.7 Percentage of occurrences of *between you and me* (n = 1,260) vs *between you and I* (n = 96) across corpus sections.

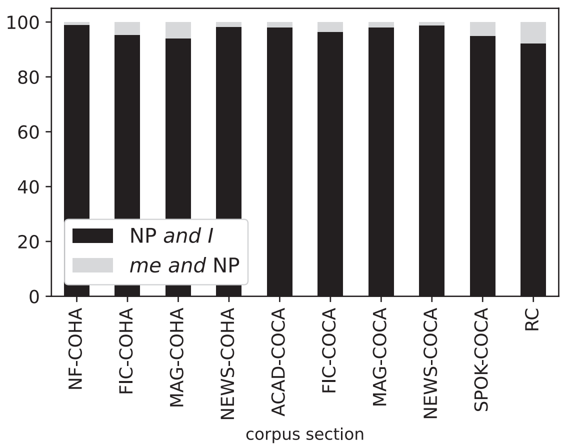


Figure 7.8 Percentage of occurrences of *NP and I* (n = 40,850) vs *me and NP* (n = 2,016) in coordinated subject NPs across corpus sections.

In the case of *all right*/*alright*, as can be seen in Figure 7.9, actual use patterns show a clear split in the use of the two variants between the RC data on the one hand, which show a clear preference for *alright*, and the rest of the corpus sections on the other hand, where *all right* is the

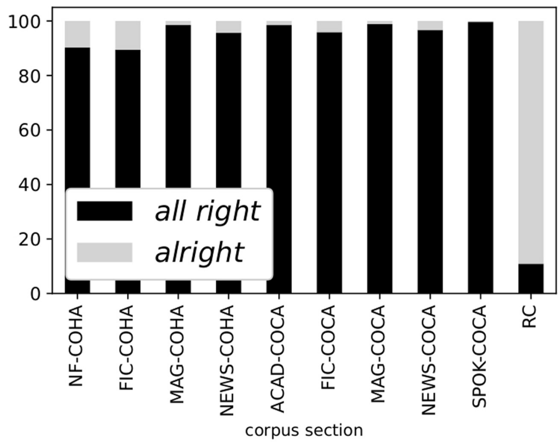


Figure 7.9 Percentage of cases of *all right* (n = 105,742) vs *alright* (n = 4,224) across corpus sections.

predominant form. It is worth noting that here the Spoken section of COCA is largely irrelevant, as the texts are transcripts of spoken conversation. In that sense, the categorical use of *all right* is not an indication of speaker choice, but an indication of standards being followed in transcription; this, however, suggests that the standard orthographic form is the form of choice.

The use of *this/these kind/sort of* with plural nouns is plotted in Figure 7.10. The overall picture suggests that *this kind/sort of* is the more commonly used variant. Corpus sections with a relatively high percentage of cases of *these kind/sort of* are the Fiction section in COHA, the Spoken section in COCA and the RC. Overall, it can be observed that while the singular forms predominate, the plural forms appear to be used as well, especially in specific contexts.

The patterns of use of negative concord as opposed to single negation are shown in Figure 7.11 and 7.12. I provide a perspective of both the percentage of use and the frequency of occurrence per million words. In terms of percentages, there is little difference across the corpus sections, with the Non-Fiction and the Fiction sections in COHA containing a slightly larger percentage of cases of negative concord. The RC data,

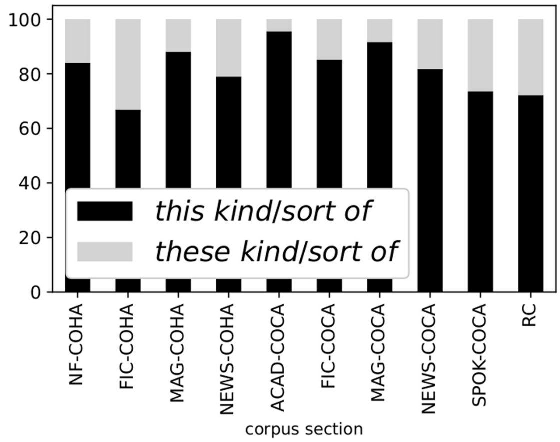


Figure 7.10 Percentage of cases of *this kind/sort of* (n = 6,022) vs *these kind/sort of* (n = 1,772) with plural nouns across corpus sections.

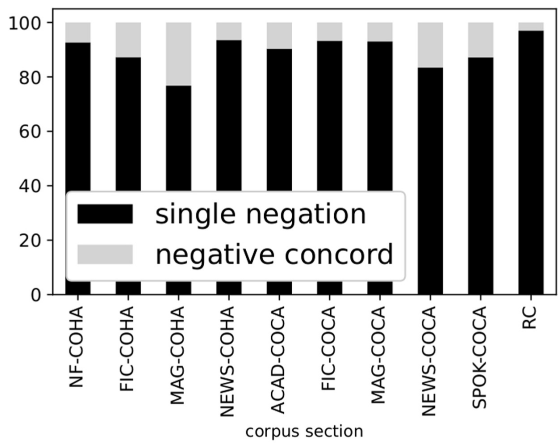


Figure 7.11 Percentage of cases of negative concord (n = 5,844) and single negation (n = 35,726).

however, does not stand out in this case. The very low percentage of cases of negative concord goes in line with its status as a non-standard feature, as well as its social salience as such (Walker 2005: 1–2; Wolfram &

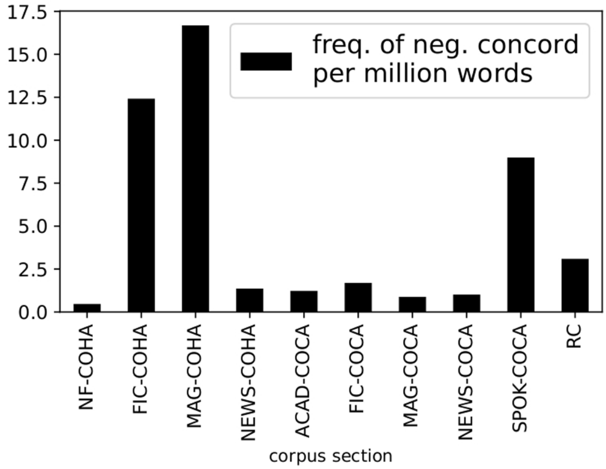


Figure 7.12 Frequency of occurrence of negative concord (n = 5,844) across corpus sections.

Schilling 2016: 162), though one might have expected a higher percentage of negative concord in the Spoken section and in the RC, given that the feature has been described as a vernacular universal (Nevalainen 2006) and 'supra-regional' and 'transnational' in terms of its nonstandardness (Gramley & Pätzold 1992, quoted in Howe 1997: 271). The rate of occurrence perspective allows us to make a more nuanced observation: here we see that Fiction texts in both COHA and COCA contain the highest rates of negative concord constructions, followed by the Spoken section in COCA and the RC, which suggests that, as a salient non-standard feature, negative concord tends to be used as a marker of nonstandardness more so than it is actually used in the context of colloquial online language use.

Figure 7.13 shows the frequency of occurrence of *ain't*, rather than the percentage of *ain't* as a variant, since *ain't* is part of a number of morphosyntactic variables (for varieties of American English, cf. e.g. Wolfram 1969, 1974; Weldon 1994; Howe & Walker 2000; Walker 2005). The rates of occurrence show that *ain't* is most frequently used in Fiction texts, but it also appears in the Spoken section of COCA, as well as in the RC data.

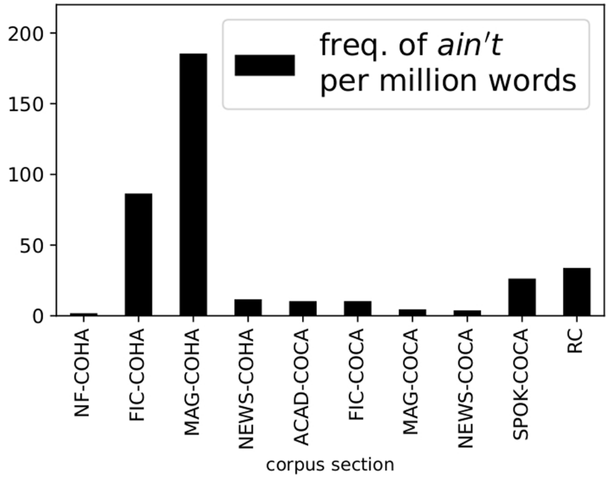


Figure 7.13 Rate of occurrence of *ain't* (n = 51,673) per million words across corpus sections.

For *ain't*, thus, we see similar patterns across corpus sections to those of negative concord, which further suggests that this linguistic feature is indeed not part of the standard, and its use is tied to specific styles or contexts of use.

To summarise, the analysis of variation patterns of these features in the corpora reveals a number of different profiles with respect to the extent to which norms can be established in actual use and can indicate standards of usage. The increasing use of *dove* can be described as a case where the irregular form may become part of the norm, though variation still exists across all contexts of use. *Snuck* for *sneaked* follows a similar pattern in spoken and online contexts, but in written texts *sneaked* seems to be used in at least half of the cases. The features *try to/try and*, *between you and me/I*, *NP and I/me and NP* in subject position, *this/these kind/sort of*, as well as negative concord and, most likely, *ain't*, are all characterised by the predominant use of the standard variant (with *between you and I* being a bit of an exception for some contexts of use). Finally, for *whom/who* and *all right/alright*, there is a clear suggestion of a register effect on the basis of the very high percentage of non-standard forms in spoken and online contexts (or only for the online contexts in the case of *all right/alright*).

4. Discussion and conclusion

The present paper set out to explore the extent to which prescriptions and patterns of language use align in American English, specifically focusing on the existence of 'double standards': the opposition between the standards which are prescribed on the one hand, and the standards observed in actual language use on the other. The comparison was based on the analysis of two sets of data: metalinguistic commentary from American usage guides and actual language use data from linguistic corpora. The results suggest a number of different relationships between prescriptions and actual use. Some of these are cases of alignment of prescriptions and use, while in other cases prescriptions and patterns of language use differ; these can be described as cases of 'double standards'.

With respect to the alignment of prescriptions and use, *dived/dove* is the only one of the examined cases that can straightforwardly be described in this way. The observations found in the metalinguistic sources that *dove* is becoming increasingly common and part of the standard are borne out by the patterns of actual use, where we see that *dove* is indeed increasingly being used across the different corpus sections, indicating acceptability across different registers and styles. *Sneaked/snuck* comes very close to this case as well, but the proportions of use of *snuck* are somewhat lower than those of *dove*, even though the prescriptions for both features are similar. This may indicate that the two are similar in salience as usage problems, which could explain why they are described in similar terms. More generally, the case of *dived/dove* shows that some prescriptions do change alongside changing standards, though this may be influenced by the extent of the change: in the case of both verbs, the incoming variant is used in at least half of the cases. In other words, as soon as an incoming variant crosses a certain threshold of frequency, it becomes salient enough as being part of the standard.

Alignment of prescriptions with patterns of language use can also be observed in the case of pronoun use in coordinated subject and object NPs, and especially in the case of the former. Here, too, the prescriptions that the non-standard variants of these features are not more generally used seem to be supported by the patterns of actual use, with *between you and I* being a slight exception in online contexts. More careful analysis taking into account a wider range of occurrences and more syntactic context of use may shed further light on the matter. For *this/these kind/sort of* with plural nouns, the more frequent variant in actual use aligns with what is considered the prescribed standard, though some of the variation may be indicative of register differences. Finally, for negative concord and *ain't*, prescriptions also align with the patterns observed in actual use.

The rest of the features exemplify cases where there are some differences between prescriptions and language use, and on the basis of these we can perhaps identify different types of 'double standards'. For *try to* vs *try and*, for instance, prescriptions are weak and generally highly permissive of this variation, especially in informal contexts. The patterns of occurrence of this linguistic feature, however, show that the standard variant in actual use appears to be *try to*. This suggests that some features may be cases of stable variation, where the non-standard variant is less frequent and is not salient as an incoming variant. Perhaps if the frequency of *try and* were to increase to the extent that it takes over *try to*, the feature will become more salient as a usage problem.

Something similar is the case with *whom/who*, to some extent, in that the predominance of *whom* in all except spoken and online contexts suggests that the standard norm is even stronger in actual use than in the metalinguistic literature. For the spoken and online contexts, *who* is the predominant variant. This is a case of double standards where a strong register distinction arises, and formal and informal contexts are clearly distinguished. A very similar case can be observed for *all right/alright*, where a clear difference between written edited texts and online written language use can be observed. Where 'double standards' arise, then, is in very specific contexts of use.

All this taken together suggests that prescriptions on standards sometimes align with actual use, which happens to be the case either for features which are salient as non-standard, or features which do not vary considerably in the first place, such as *try to/and*. Despite this general tendency, however, some cases of change seem to be accepted, suggesting that authors of prescriptions are sensitive to changing standards of use. For other features, a situation of 'double standards' arises, which can be observed not necessarily through a comparison between prescriptions and use, but through a comparison between written edited texts and spoken or online language use. This suggests that the stage of prescription appears to be characterised by the mutual influence of prescribed standards and norms observed in actual language use, with more traditional or conservative registers following prescriptions in almost all cases. Where changes take place to such an extent that the use of non-standard variants increases in written texts as well, prescriptions seem to change and align to reflect the patterns observed.

Bibliography

Albakry, Mohammed. 2007. Usage prescriptive rules in newspaper language. *Southern Journal of Linguistics* 30(2), 28–56.

Anderwald, Lieselotte. 2012. Variable past-tense forms in nineteenth-century American English: Linking normative grammars and language change. *American Speech* 87(3), 257–93.

Anderwald, Lieselotte. 2013. Natural language change or prescriptive influence? *English World-Wide* 34(2), 146–76.

Anderwald, Lieselotte. 2014. Measuring the success of prescriptivism: Quantitative grammaticography, corpus linguistics and the progressive passive. *English Language & Linguistics* 18(1), 1–21.

Anderwald, Lieselotte. 2016. *Language between description and prescription: Verbs and verb categories in nineteenth-century grammars of English*. Oxford: Oxford University Press.

Ayres-Bennett, Wendy. 2020. From Haugen's codification to Thomas's purism: Assessing the role of description and prescription, prescriptivism and purism in linguistic standardisation. *Language Policy* 19, 183–213.

Baumgartner, Jason, Savvas Zannettou, Brian Keegan, Megan Squire & Jeremy Blackburn. 2020. The Pushshift Reddit Dataset (arXiv:2001.08435). *arXiv*. <https://doi.org/10.48550/arXiv.2001.08435>

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *The Longman Grammar of Spoken and Written English*. Harlow: Longman.

Boyland, Joyce Tang. 2001. Hypercorrect pronoun case in English? Cognitive processes that account for pronoun usage. In Joan Bybee & Paul J. Hopper (eds). *Frequency and the emergence of linguistic structure*. Amsterdam, Philadelphia: Benjamins. 383–404.

Brems, Lieselotte & Kristin Davidse. 2010. The grammaticalisation of nominal type noun constructions with *kind/sort of*: chronology and paths of change. *English Studies* 91(2), 180–202.

Chapman, Don. 2017. Stalwarts, snoots and some readers: How 'Traditional Rules' are traditional. In Ingrid Tieken-Boon van Ostade & Carol Percy (eds). *Prescription and tradition in language: Establishing standards across time and space*. Bristol: Multilingual Matters. 238–52.

Chapman, Don. 2019. 'Splendidly prejudiced': Words for disapproval in English usage guides. In Birte Bös & Claudia Claridge (eds). *Norms and conventions in the history of English*. Amsterdam, Philadelphia: Benjamins. 29–47.

Curzan, Anne. 2014. *Fixing English: Prescriptivism and language history*. Cambridge: Cambridge University Press.

Dant, Doris R. 2012. Using COCA to evaluate the Chicago Manual of Style's usage prescriptions. In Joybrato Mukherjee & Magnus Huber (eds*). Corpus linguistics and variation in English: Theory and description*. Leiden: Brill. 29–39.

Davies, Mark. 2008–. *The Corpus of Contemporary American English (COCA).* <https://www.english-corpora.org/coca/> accessed 8 June 2024.

Davies, Mark. 2010. The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing* 25(4), 447–64.

Davies, Mark. 2010. *The Corpus of Historical American English (COHA).* <https://www.english-corpora.org/coha/> (Accessed 8 June 2024.)

Davies, Mark. 2012. Expanding horizons in historical linguistics with the 400-million-word Corpus of Historical American English. *Corpora* 7(2), 121–57.

Denison, David. 1998. Syntax. In Suzanne Romaine (ed.). *The Cambridge History of the English Language*. Vol. 4. Cambridge: Cambridge University Press. 92–329.

Gramley, Stephan & Kurt-Michael Pätzold. 1992. *A survey of Modern English*. London: Routledge.

Grieve, Jack. 2012. A statistical analysis of regional variation in adverb position in a corpus of written Standard American English. *Corpus Linguistics and Linguistic Theory* 8(1), 39–72.

Haugen, Einar. 1966. Dialect, language, nation. *American Anthropologist* 68(4), 922–35.

Hickey, Raymond. 2012. Standard English and standards of English. In Raymond Hickey (ed.). *Standards of English: Codified varieties around the world*. Cambridge: Cambridge University Press. 1–33.

Hinrichs, Lars, Benedikt Szmrecsanyi & Axel Bohmann. 2015. Which-hunting and the Standard English relative clause. *Language* 91(4), 806–36.

Hommerberg, Charlotte & Gunnel Tottie. 2007. *Try to* or *try and*? Verb complementation in British and American English. *ICAME Journal: Computers in English Linguistics* 31, 45–64.

Howe, Darin M. 1997. Negation and the history of African American English. *Language Variation and Change* 9(2), 267–94.

Howe, Darin M. & James A. Walker. 2000. Negation and the creole-origins hypothesis: Evidence from early African American English. In Shana Poplack (ed.). *The English History of African American English*. Malden MA: Blackwell. 109–40.

Hyper Usage Guide of English. 2014. Leiden University. <http://huge.ullet.net> accessed 8 June 2024.

Kostadinova, Viktorija. 2020. Prescriptivism as a constraint in language variation and change: Moving beyond binary oppositions. In Don Chapman & Jacob D. Rawlins (eds). *Language prescription: Values, ideologies and identity*. Bristol: Multilingual Matters. 95–120.

Kostadinova, Viktorija, Tove Larsson & Tülay Dixon. In preparation. *Prescriptivism and language use: A speaker perspective*.

Kretzschmar, Jr, William A. 2010. The development of Standard American English. In Andy Kirkpatrick (ed.). *The Routledge Handbook of World Englishes*. London, New York: Routledge. 96–112.

Kretzschmar, Jr, William A. & Charles Meyer. 2012. The idea of Standard American English. In Raymond Hickey (ed.). *Standards of English: Codified varieties around the world*. Cambridge: Cambridge University Press. 139–58.

Lippi-Green, Rosina. 2012. *English with an accent: Language, ideology and discrimination in the United States*. London, New York: Routledge.

Mair, Christian. 2006. *Twentieth-Century English: History, variation, and standardisation*. Cambridge: Cambridge University Press.

Milroy, James & Lesley Milroy. 1985. *Authority in language: Investigating language prescription and standardisation*. London, New York: Routledge

Nevalainen, Terttu. 2006. Negative concord as an English 'vernacular universal': Social history and linguistic typology. *Journal of English Linguistics* 34(3), 257–78.

Peters, Pam & Wendy Young. 1997. English grammar and the lexicography of usage. *Journal of English Linguistics* 25(4), 315–31.

Proferes, Nicholas, Naiyan Jones, Sarah Gilbert, Casey Fiesler & Michael Zimmer. 2021. Studying Reddit: A systematic overview of disciplines, approaches, methods, and ethics. *Social Media + Society* 7(2), 1–14.

Schneider, Edgar W. 2007. *Post-colonial English: Varieties around the World*. Cambridge: Cambridge University Press.

Schneider, Edgar W. 2021. Meanderings from early English to World Englishes: A complex systems perspective on morphosyntactic changes in wh-pronouns In Peter J. Grund, Megan E. Hartman (eds). *Boundaries and boundary-crossings in the history of English*. Berlin, Boston: de Gruyter. 73–105.

Sundby, Bertil, Anne Kari Bjørge & Kari E. Haugland. 1991. *A Dictionary of English normative grammar, 1700–1800*. Amsterdam, Philadelphia: Benjamins.

Tieken-Boon van Ostade, Ingrid. 2012. The codification of English in England. In Raymond Hickey (ed.). *Standards of English: Codified varieties around the World*. Cambridge: Cambridge University Press. 34–54.

Walker, James A. 2005. The *ain't* constraint: *Not*-contraction in early African American English. *Language Variation and Change* 17(1), 1–17.

Weldon, Tracey. 1994. Variability in negation in African American Vernacular English. *Language Variation and Change* 6(3), 359–97.

Wolfram, Walt. 1969. *A sociolinguistic description of Detroit Negro Speech*. Washington, D.C.: Center for Applied Linguistics.

Wolfram, Walt. 1974. *Sociolinguistic aspects of assimilation: Puerto Rican English in New York City*. Arlington, Va.: Center for Applied Linguistics.

Wolfram, Walt & Natalie Schilling. 2016. *American English: Dialects and variation*. Malden MA: Blackwell.

Yáñez-Bouza, Nuria. 2021. Methodological approaches to the study of codification, prescription, and prescriptivism. *Studia Neophilologica* 94(3), 334–64.

ROSA M. CALAFAT VILA

8 Lexicon and language ideologies: Double solutions for Catalan language use in the media[1]

ABSTRACT

This article analyses the conflict between two language solutions in Catalan media at a lexical and semantic level: one permissive, where Standard Spanish influences Catalan; and one that socialises distinct terms from solutions in Spanish. Reference style guides for the media normalise linguistic forms (words, phrases and neologisms) that come from Catalan subordination to Spanish, especially in medium- to low-level formal registers. Specialists (linguists, writers, editors) have raised the alarm on social media, reporting Hispanicisms, or forms that adopt a similar physiognomy to Spanish terms and encouraging people to use more authentic terms. Standing midway between the two approaches, the Institute for Catalan Studies has included certain Hispanicisms from the media in recent decades. This study analyses the lexical and semantic approaches of these two currents based on a qualitative sample, as well as their incorporation into language rules.

1. Introduction[2]

Media outlets (television, radio, newspapers and social media) currently balance two solutions in their approach to using a standard variety of

1 This project is partially co-financed by the Autonomous Government of the Balearic Islands, through the Direcció General de Política Universitària i Recerca, with funds from the Tourist Stay Tax Act ITS 2017–006 (PRD2018/03).

2 I have previously published an article on the Corporació Catalana de Mitjans Audiovisuals ('Catalan Broadcasting Corporation') style guide (Calafat 2020: 70–81) from the *ésadir* language portal. The lexicon from *ésadir* cited and analysed here expands on the references used in the previous article.

Catalan.[3] Certain style guides prioritise lexical choices, including neologisms, arising from Spanish language interference. Others simply opt for the most similar form to Spanish where Catalan has two alternatives for the same concept.[4] In turn, linguists and journalists critical of this more Spanish-leaning approach propose solutions that align more with customary language uses. The academic and regulatory authority represented by the Language Unit (SF) at the *Institut d'Estudis Catalans* ('Institute for Catalan Studies', IEC) seemingly attempts to balance both approaches, given that it includes forms similar to Spanish structures.[5]

This double solution seeks to reconcile two differing takes on what is, must or should be the sociolinguistic reality of the Catalan-speaking community. These contrasting positions are strongly reflected in lexicon – the most flexible aspect of language. The first approach places a modern value on words dependent on Standard Spanish, since this aligns with the one-way social bilingualism seen in legislation and by a certain sector of society, i.e. Spanish-speakers are able to live alongside Catalan-speakers in Catalan-speaking areas of Spain without knowing or speaking Catalan, although the opposite is not possible. The Spanish Constitution (1978) only specifies Spanish as the official language of the state, while other languages are recognised as co-official in the regions where they are spoken. It is through regional (lower-level) laws that other languages become jointly official. The second approach advocates for a fully functional and formal standardisation based on the language loyalty principle (Weinreich 1953),

3 Catalan is a cross-border language in Europe and spoken in four countries: Spain, France, Andorra and Italy. It is spoken across different autonomous regions in Spain (Catalonia, Valencia, the Balearic Islands, Aragon and Murcia). The statutes of autonomy in the first three regions officially recognise Catalan.

4 This article uses the term *Hispanicism* to refer to lexical solutions that are closest to Spanish and normally exist alongside another term with a more traditional Catalan physiognomy.

5 The *Institut d'Estudis Catalans* (IEC) brings together the top Catalan language experts and is organised into different sections. The Secció Filològica ('Language Unit') serves as the Catalan Academy, sets language standards and carries out language analysis.

encouraging fellow speakers to purge Catalan of interference from linguistic subordination.

The Catalan government has enacted the *Pacte Nacional per la Llengua* ('National Language Agreement'), a language policy initiative aimed at increasing the social use of Catalan. In turn, the *Qualitat de la llengua* ('Language Quality') document produced in collaboration with IEC Language Unit members highlights the current debate surrounding double solutions in the media. Some defend the acceptance of *Catanyol* (a term referring to a mix of Catalan and Spanish) and support solutions arising from interference in a strict sense: introducing elements that come from another linguistic system and used by speakers immersed in social and asymmetrical bilingualism (Payrató 1985: 58–9).[6] Meanwhile, others defend 'linguistic purism' and as a group are 'highly active on the web'.[7] These supporters of an 'authentic' standard use social media to criticise how certain media normalise language transference. They establish a dialogue between experts (linguists) and users (Kailuweit 2002: 172–4) and make them moderators of attitudes towards language loyalty (Weinreich 1953).

1.1 Study methodology and scope

This article analyses the lexical proposals (words, phraseological units and neologisms) from the two currents (interfered vs authentic Catalan) regarding the standard variation used in the media that runs the gamut of functional language registers (Lamuela 1987: 98; Bibiloni 1997: 77). The methodology uses the lexical and semantic interference classification system established by Weinreich (1953) that Payrató (1985) applied to Catalan.

6 Weinreich (1996 [1953]: 29) defines interference phenomena as deviation from the norms of either language, alternately used by a bilingual speaker as a result of language contact.

7 *Qualitat de la llengua* (p. 2), <https://llengua.gencat.cat/ca/pactenacionalperlallengua/documents/> (accessed 8 June 2024).

Lexicon is the least organised and most unstable language structure. The influence of a dominant language on a recessive one is clearly and initially reflected in vocabulary. Lexical innovations absorbed by the host language lead to a series of language phenomena that obviously impact a word's semantics: weakening lexical and semantic distinction, or strengthening lexical distinction, loss or substitution and semantic and expressive innovations (Weinreich 1953; Payrató 1985: 112–23; Argenter 2020; Pradilla 2020). The locutions, collocations (elements united by mutual attraction) and set phrases of speech are marked cultural and pragmatic elements in the lexical sphere (Levinson 1983: 22–3; Galisson 2000: 51).

The lexical component of language in audiovisual media must cover all types of diaphasic variation: medium-high (news, speciality programming) and medium-low (chat shows, fiction). The latter also includes pseudo-slang (Massanell 2009: 153–84; Calafat 2020). Certain style guides contain numerous entries at this level for lexical Hispanicisms (words and phrases). For instance, the *Llibre d'estil El Periódico de Catalunya* (2003), a non-governmental newspaper for a broad range of readers in Catalonia offering Spanish and Catalan editions; and the *Llibre d'estil* (2011) from the *ésadir* language portal and *Corporació Catalana de Mitjans Audiovisuals* (CCMA),[8] the public broadcaster owned by the Catalan government.[9]

The style guide at *El Periódico*[10] is known for adopting a broad selection of solutions not included in the standard dictionary (*Diccionari de l'Institut d'Estudis Catalans*, DIEC).[11] The guide itself sets out its transgressive approach (including *casc antic* (es: *casco antiguo*) for the standard

8 The CCMA Catalan Broadcasting Corporation guide is split into editorial guidelines (*Guía editorial*), a usage handbook (*Manual d'ús*), and the *ésadir* language portal, <https://esadir.cat/> (accessed 8 June 2024).

9 The corporation operates several television channels and radio stations (Catalunya Ràdio), as well as digital content, <https://www.ccma.cat/corporatiu/ca/el-grup/> (accessed 8 June 2024).

10 To enable quicker reading, the lexical examples taken from *El Periódico* do not indicate the page where they appear. The style guide is in alphabetical order allowing information to be easily retrieved.

11 The online version is available at <https://dlc.iec.cat/> (accessed 8 June 2024).

classical term *barri vell* ('old town'). Moreover, it justifies Hispanicisms in written pieces in the newspaper by citing the social reality of semi-speakers or imperfect speakers – individuals who do not react to mass interference, have a relaxed approach to language and sociolinguistic norms, and little grammatical or lexical skill in the subordinate language, although they are able to communicate in it (Dressler 1987: 87–96). The aim of the guide is to push the DIEC into including certain Hispanicisms:

> El model de llengua d'El Periódico es caracteritza per ser respectuós amb la normativa de l'Institut d'Estudis Catalans, sense defugir la responsabilitat que té, com a mitjà de comunicació, de reflectir els usos lingüístics habituals de la comunitat a què s'adreça. Per aquest motiu, el signe boleta (·) que precedeix una gran quantitat d'entrades avisa del caràcter no normatiu del terme, acceptat amb normalitat en els textos del diari. La mesura pretén deixar clars els límits de la transgressió de la norma en benefici de l'ús habitual, i alhora contribuir, si és el cas a la futura ampliació del diccionari normatiu.
>
> (*Llibre d'estil El Periódico de Catalunya* (2003: 9)
>
> 'The language model at El Periódico is characterised by respect for *Institut d'Estudis Catalans* standards without shirking its responsibility as a media outlet to reflect the common language uses of its community of readers. For this reason, the centred dot (·) before a large number of entries indicates the non-standard nature of the term, normally accepted in the articles of the newspaper. This measure intends to clearly mark the limits of non-standard use in favour of common usage, while contributing, where applicable, to future additions to the standard dictionary'.[12]

The intentions of *ésadir*[13] align with those of the newspaper style guide. They want innovations, neologisms and Hispanicisms on the digital, interactive language portal to be officially incorporated into the DIEC. Undoubtedly, these types of publications need to address communication requirements in a globalised world, where newly included or coined economic, social and technological expressions, terms or words are a sign

12 All translations are my own unless otherwise noted.

13 The examples taken from the *ésadir* portal may be viewed directly at <https://esadir.cat/> (last visited 1 February 2023). The information taken from the portal and provided in this article was last checked on 25 February 2023.

that a language is part of the wider world (Diki-Kidiri 1999: 69–72). Nonetheless, where inclusion seems superfluous, with Hispanicisms replacing native words, or makes it difficult for one language to use its own linguistic structures to incorporate loanwords and calques (interference in the widest sense), the dominant language interferes with the recessive language through interference in the narrowest sense. For instance, the *ésadir* portal includes a semantic expansion of the word *contemplar* 'contemplate' in the sense of *preveure* 'foresee', *considerar* 'consider', *estudiar* 'study' or *tenir en compte* 'take into account', from the Spanish homonym *contemplar*, which does have these meanings. In Catalan, the word can only be used in the sense of *posar atenció* 'look at'.

Those on the other side of the debate support a Catalan purged of Hispanicisms and this study presents a qualitative sample of lexical proposals from two influential linguists in the media in terms of followers and social impact: Dr Gabriel Bibiloni, a university lecturer, and Jordi Badia, a style guide editor at the online media outlet Vilaweb.[14] Both spotlight Spanish loanwords in media texts (posters, adverts, TV programmes, internet, etc.) while criticising, where applicable, their inclusion in the DIEC, e.g. *passi* (es:[15] *pase*), 'pass', in the sense of 'ticket, subscription, entry document'.[16] Given the scope and aim of this article, dual solution phenomena are not measured qualitatively but restricted to a qualitative analysis that aims to highlight alternative solutions to close contact between two languages.

Furthermore, the article includes an analysis of language observations made by the most representative online group *El català com cal* ('Catalan as it has to be'), whose tagline is *Parles català o calques i tradueixes literalment del castellà?* ('Do you really speak Catalan or just copy

14 Professor Bibiloni has 17,100 Twitter followers (<https://twitter.com/bibiloni>) and publishes articles on his blog (<https://bibiloni.cat>). Badia has 16,700 Twitter followers (<https://twitter.com/jbadia16>). The *El clot de les ànimes* ('Valley of the Souls') blog includes his outreach activities on standard language (<https://blocs.mesvilaweb.cat/jbadia/>). All sites were last accessed on 1 February 2023.

15 'es' = Spanish.

16 Bibiloni (12 January 2023), <https://twitter.com/bibiloni/status/1613619810155241472>

and translate literally from Spanish?'). This group lists and corrects language interference in journalism, and criticises Hispanicisms admitted by language authorities.[17] Finally, there is a cross-cutting look at *Cercle Vallcorba*.[18] Members of this group are mostly university lecturers, editors, translators and writers who align with the approach of emblematic linguist Jaume Vallcorba (Barcelona 1920–2010), and advocate for the loyalty principle in language standards and uses.[19] They use the media to adopt critical stances, and group members have written articles and books stirring controversy and conflict with academic authorities:

> Si l'aposta metodològica de la SF (Secció Filològica) per a arribar a la norma era passar per la 'descripció'. Era indefugible i inexcusable que aquesta gramàtica descriptiva inclogués una descripció de la substitució estructural el català pel castellà actualment en curs. En comptes d'això, la SF no solament silencia aquesta desfeta estructural, sinó que sovint la maquilla de 'col·loquialisme'. (Cercle Vallcorba 2022: 135)
>
> 'If the SF (Language Unit) methodology to arrive at a standard was through "description". This descriptive grammar should have inevitably and clearly included a description of the current and ongoing structural replacement of Catalan by Spanish. Instead of this, the SF not only stays silent on this structural rout but also often masks it by using the term "colloquialism"'

17 The anonymous collective has 31,200 Twitter followers (<https://twitter.com/elcatalacomcal>) and a blog (<http://elcatalacomcal.blogspot.com/>) focused on correcting Hispanicisms in Catalan in the media (last accessed 1 February 2023).

18 The group has fewer Twitter followers (<https://twitter.com/cerclevall>) than the others included here (2,205) and its blog (<https://www.cerclevallcorba.cat/>) contains members' publications from online platforms and traditional media. It published the book *El català, on va?* ('Where is Catalan Heading?') in 2022, which contains a large selection of the group's outreach work (last accessed 1 February 2023).

19 Other online groups dedicated to promoting language loyalty include *No em canviïs la llengua* ('Don't Change Language on Me') with 17,900 followers, <https://twitter.com/CanviisEm?ref_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Eauthor> (last accessed 1 February 2023).

2. Catalan-Spanish language contact: Ongoing controversies

2.1. A brief historical overview

The Catalan Royal Court and Chancellery played a key role in codifying language in the fourteenth and fifteenth centuries. However, this benchmark would disappear with the arrival of the Habsburg dynasty on the Catalan-Aragonese throne in the sixteenth century. This waning political power had a considerable impact on Catalan literature through the adoption of Spanish loanwords. Meanwhile, Spanish was entering the splendour of the Golden Age. Nevertheless, these loanwords only appeared in books (Colon 1989: 77). It was the arrival of the Bourbons in the eighteenth century that led to the imposition of Spanish in schooling and government. In the twentieth century, the military coup led by General Franco imposed a forty-year dictatorship in Spain (1936–1975) that included policies of language genocide with regard to Catalan (Benet 1995). The period saw major migration of Spanish-speakers to Catalan-speaking territories (1950s and 1960s) who altered the linguistic demography in these regions. Nonetheless, and due to a variety of factors, Catalan managed to find spaces within society. After democracy was restored, Catalan-speaking areas within Spain (represented by the regional capitals of Barcelona, Valencia and Palma) recouped a certain level of power through the statutes of autonomy (1979–1983), which granted native official status to Catalan (alongside Spanish). Regional language policies are based on the right to speak Catalan (Vernet 2004: 39–42). In this sense, the 1980s ushered in a shift towards language normalisation of Catalan in a sociolinguistic context, where an autochthonous population continued to speak Catalan in informal registers but had been denied schooling in the language. At the same time, a Spanish-speaking population coexisted in Catalan-speaking capital cities, in addition to an intellectual elite based around universities who nurtured standard, formal and literary Catalan. Broadcast services in Catalan began in 1983 on public service television (TV3) operated by the

Catalan government, which gradually reached all Catalan-speaking areas. This was followed by Canal Nou television in Valencia from 1989, and the regional broadcaster in the Balearic Islands, IB3, in 2004 (Gifreu 2011: 74–6). Nonetheless, the language model adopted by broadcasters in the twentieth century (for several reasons, the Balearic public broadcasting service is not included here) remained restricted to the local (or regional) area, and a negative dynamic in establishing a common cultural market for Catalan-speakers became entrenched (Gifreu 2011; Darder 2015). Certain style guides used words or expressions from Standard Spanish in the media to simulate colloquial language, and words that remained common outside the regional area were categorised as dialectal (and not suitable for functional media solutions). The intrusion (Aracil 1983) of Standard Spanish into Catalan dialects is even more notable in medium to low formal registers. For example, the *ésadir* blog includes the word *passo* (es: *paso*) in a figurative sense meaning to eschew doing something and fails to include the verb-adjacent clitic *me n'afluix*, 'I can't be bothered', used in other Catalan-speaking regions.[20]

2.2 *Language controversies as social ideologies*

The enactment of laws regulating the social use of Catalan in Catalonia (1983), Valencia (1983) and the Balearic Islands (1986) generated controversy regarding language solutions in the media. In Catalonia, this debate led to public confrontations between linguists who supported two different approaches. So-called defenders of a *diet* Catalan supported Catalan and Spanish moving closer together under the banner of naturalness (not intervening in the language interference process) and confined linguistic

20 Bibiloni lists Hispanicisms used by the CCMA alongside the standard form in his blog article from 8 June 2016, *Els castellanismes de la ràdio i la televisió públiques de Catalunya* ('Hispanicisms on Catalan Public Radio and Television'), <https://bibiloni.cat/blog/index.php?s=vaccinar&image.x=0&image.y=0> (last accessed 1 February 2023).

solutions to the restricted geographical area of Barcelona. As the sub-editor of the *Avui* newspaper explicitly stated:

> Així, per exemple, es tendeix a evitar *enguany* (forma dialectal que no és viva a Barcelona i la seva rodalia). [...] Les paraules que es proposen, per raons sòcio-culturals, no són neutres sinó que connoten un registre que està força allunyat de la parla col·loquial en què els normativistes les voldrien inserir. Això fa que el parlant se senti profundament incòmode quan les usa i, tret del cas que estigui molt motivat per una forta consciència ideològica, acabi abandonant-les a favor dels corresponents castellanismes, que són els termes neutres que s'adapten sense grinyolar gens al registre col·loquial. (Pla 1989: 219, 238)
>
> 'Thus, for example, one tends to avoid *enguany* 'this year' (a dialectal form not used in Barcelona and the surrounding area). [...] The words proposed for socio-cultural reasons are not neutral but rather denote a register that is far removed from the everyday speech where standard-bearers would like to see them. This means speakers feel deeply uncomfortable when they use them and, except in instances where they are highly motivated by ideology, end up abandoning their use in favour of similar Hispanicisms, which are the neutral terms adapted without any nod to the informal register'.

Linguists in this camp underpin the approaches of emblematic media language solutions: the newspapers *Avui* and *Diari de Barcelona*, and CCMA channels and stations (TV3, Catalunya Ràdio). They established the *Grup d'Estudis Catalans* ('Catalan Studies Group') with a view to lobbying academic authorities represented by the IEC (Kailuweit 2002: 171–2). In turn, linguists at universities and media sub-editors who did not agree with the *diet* version, such as *El Punt*, *Vilaweb*, etc., supported the use of a standard language stripped of Hispanicisms in line with classical standards. These are the so-called defenders of a *hardcore* Catalan.[21] Thus, a battle between *laissez faire*, *laissez passer* nature and nurture commenced to reverse the dynamics of subordination (Bibiloni 1997: 60–1).

21 The terms *diet* and *hardcore* refer to the debate in the mid-1980s between supporters of *Català light* who opted for solutions similar to Spanish, and those who supported *Català heavy* with more traditional and 'genuine' proposals (Argneter 2020: 604). Other controversies that have impacted language in the media include the representation of diatopic variations within the standard model (Calafat 2020).

Still, supporters of a more natural approach took up strategic posts at top media outlets, leading to its spread from Barcelona, with greater or lesser difficulty due to restrictive policies imposed on Catalan by the Spanish government, to the rest of the Catalan-speaking community.

In Catalonia, *diet* Catalan supporters configured a regionally limited version of the language that was socially targeted at semi-speakers – people unaware of errors and who believed any authentic standard in the recessive language held little value, since its lexicon belonged to elderly native speakers (Dressler 1987: 87–96). This is how a media solution with bilingual speakers as a model formed and spread. Ultimately, lexical choices are not neutral but adapt to belief systems within society:

> Fora dels casos de genocidi, el cessament de la reproducció d'una llengua ha d'anar acompanyada d'un canvi en la consciència dels parlants, és a dir, necessita una legitimació ideològica que compensi l'esforç de trencar amb les pràctiques conegudes i d'adoptar-ne de noves. Aquest canvi ideològic, que té a veure amb la necessitat d'integració i de creació d'identitat, és molt més suau en el cas de migracions voluntàries, sobretot en situacions d'exogàmia, i molt més costós en situacions de minorització. (Boix & Vila 1998: 186–7)
>
> 'Except for cases of genocide, blocking reproduction of a language must be accompanied by a change in speakers' mindsets, i.e. it needs ideological legitimation that compensates for the effort of breaking with known practices and adopting new ones. This ideological shift, linked to the need for integration and creation of identity, is much gentler in the case of voluntary migrations, especially in exogamic situations, and much more painful in minoritisation contexts'.

In turn, supporters of *hardcore* Catalan thought any vernacularism suitable if it contributed to stylistic differentiation and was communicationally fruitful. They based this on the fact that colloquial Catalan was a living reality that merely required transfer and planning to secondary orality designed to simulate spontaneity. In contrast, the diet group decided, in the name of modernity, to ignore traditional words from the informal register: '[…] hi ha qui considera que registre col·loquial vol dir català acastellanat […]. Si això fos així, hauríem de deduir que els nostres avis, molts dels quals eren analfabets, parlaven en registre literari, perquè no empraven català acastellanat' ('[…] some deem the colloquial register to mean Hispanicised Catalan […]. If this were the case, we would have to deduce that our grandparents, many of whom were illiterate, spoke in the literary register, since they did not use Hispanicised Catalan'; Cercle Vallcorba 2022: 207).

Hardcore and *diet* supporters held two different opposing socio-political concepts. For those who denounced the linguistic subordination of Catalan to Spanish, it was essential for the recessive language to de-normalise Spanish in Catalan-speaking areas to attain functional completeness. In 1991, certain language lecturers at the University of Girona protested against the bilingualist myth (Aracil 1983) in what would become a symbolic publication:

> Perquè desenganyem-nos: no es pot pensar en la normalització del català si no preveiem una desnormalització de l'espanyol als Països Catalans. I que no ens diguin utòpics o insolidaris; només demanem per al català el mateix tracte legal que té l'espanyol: que sigui l'única llengua oficial en el seu territori. (Prats et al. 1991: 67)
>
> 'Why delude ourselves? We cannot dream of Catalan normalisation if we do not see a de-normalisation of Spanish in Catalan-speaking areas. And pay no heed to those who call us utopian or lacking in solidarity; we are merely asking for Catalan to have the same legal consideration as Spanish: the only official language in its territory'.

The link between linguistic repertoire and socio-political organisation (Boix & Vila 1998: 121) is clear in the lexicon proposed by these solutions.

3. Lexical proposals in contrast: A qualitative analysis

The *Diccionari general de la llengua catalana* ('General Catalan Language Dictionary', DGLC) was published in 1932 and edited by the codifying standard-bearer for contemporary Catalan, Pompeu Fabra (1868–1948). The DGLC served as the academic dictionary until 1995, when the *Institut d'Estudis Catalans* released the DIEC (DIEC1), followed by a second edition in 2007 (DIEC2) in tandem with an online version. The DIEC1 incorporated certain Hispanicisms that the DGLC did not. Undoubtedly, pressure from the *Grup d'Estudis Catalans* regarding a *diet* version of Catalan influenced the academy's decisions. Linguists and media professionals criticised the academic authority for adding apparently superfluous words to the language:

> Costa molt d'entendre per què han estat admesos una bona quantitat de castellanismes absolutament superflus i que no fan cap falta a la llengua. L'única raó per a admetre'ls és la de l'ús actual que tenen en la llengua parlada pel sector de la població que encara no ha accedit al coneixement de l'estàndard. Però aquesta no és la línia que havia seguit fins ara la nostra acadèmia de la llengua, en què havien prevalgut els criteris establert per Pompeu Fabra. (Esteve et al. 1998: 35)
>
> 'It is extremely hard to understand why so many totally unnecessary Hispanicisms have been added to the language. The only possible reason is current use in spoken language by a sector of the population who has not yet gained access to the standard language. This is not the way our academy had done things to date, with criteria set by Pompeu Fabra taking precedence'.

The link between Catalan and Spanish in standardised lexicon led to the publication of language works targeted at the general public. These aimed at exposing mixed codes in Catalan and highlighting linguistic subordination enabled by the academy (Perea 1998: 115). Moreover, these texts had an educational purpose since they were aimed at a population with limited or no access to teaching in their native language: 'El cas és que el català disposa, com a llengua subordinada que és, d'un sistema de registres limitat i que els seus parlants ignoren sovint bona part dels recursos expressius' ('The truth is that Catalan, as the subordinated language that it is, has a limited system of registers, and its speakers are often unaware of many expressive resources'; Lamuela 1987: 88). This is true for the *Diccionari auxiliar* ('Auxiliary Dictionary') (1996) from grammarian Josep Ruaix, who produced sixty grammar books aimed at informal education. The publication comes in two parts: the *Repertori lexicogràfic d'interferència en el català d'avui* ('The Lexical Repertoire of Interference in Catalan Today') and *Dubtes, incorreccions, preferències i remarques lingüístiques* ('Language Doubts, Mistakes, Preferences and Remarks'). It provides an overview of lexicographic interference (words and set phrases) in the DIEC1, e.g. *forat -ada* (es: *aforado -da*) for *afranquit -ida*, 'immune'; *embotellament* (es: *embotellamiento*) for *embós*, 'traffic jam'; *empastament* (es: *empaste*) for *obturació*, 'filling'. It also highlights instances where the principle of least effort drives bilinguals towards a similar or matching form whose use leads to poorer vocabulary and semantics in the recessive language, e.g. *desbarrar* (es: *desbarrar*), instead of the compound forms *pixar fora de test*, *sortir de polleguera* 'to rant and rave' (Ruaix 1996).

3.1. Hispanicisms in style guides: Media vernacular

The media style guides of *El Periódico* and CCMA stations and channels have limited mass inclusion of Hispanicisms to medium and low formal registers. At these levels, secondary orality is central (the spontaneity the media aims to reflect) and admits vernacularism, certain slang words, semantic inaccuracies and some deeply rooted Hispanicisms (Cercle Vallcorba 2022: 207). Mass interference for the sake of naturalness, without any drive for correction, represents the perpetuation and legitimation of semi-speakers' social normality. For the editors at *ésadir*, the imperfect speaker is part of the solution. As the head of the language portal stated:

> En una societat tan bilingüitzada [...] les persones participen en converses on es barregen les dues llengües i a vegades alguna altra. [...] Com més naturalitat es vulgui donar al programa, més fàcil és que s'hi sentin [...] castellanismes habituals al carrer (*estúdia*, *raro*, *pavo*, *estupendo* ...). (Camps 2007: 5–8)
>
> 'In such a highly bilingual society [...] people have conservations where both languages, as well as others, blend. [...] The more natural they want a programme to sound, the easier it is to hear [...] normal street-level Hispanicisms being used (*estúdia* [incorrect stress], *raro*, *pavo*, *estupendo* ('studies, strange, turkey, great')'.

The *ésadir* style guide lists lexical entries by normalisation – stating whether or not they are included in the standard dictionary or other lexicographic sources – or stylistic criteria in line with their functional register. Words neither included in the standard dictionary nor accepted by the blog are qualified as *inadmissible*. The *correct/incorrect use* label indicates whether words and expressions are used in the right sense. The *overuse* label indicates there are more appropriate synonyms. It lists words from standard Catalan (*guadir* 'to enjoy') and Hispanicisms (*tardar* [es: *tardar*] for *trigar* 'to delay').

The *general use* label is reserved for vocabulary which could be used in all language registers, regardless of whether it is admitted as standard or not. *Ésadir* lists words and phrases as preferential and in use (not in the DIEC) due to universal contact with Spanish. Consider the following examples: *aguantar el tipus* (es: *aguantar el tipo*), instead of *fer el cor fort*, 'to remain composed'; *anar de* (es: *ir de/ de qué va*), for *tractar* (a book or film on a specific topic), 'to be about'; *a sac* (es: *a saco*), for *a les totes/de valent*, 'full-on'; barco (es: *barco*), for *barca o vaixell*, 'boat'; *fer mala espina* (es: *dar mala espina*), for *fer pudor de socarrim* 'to make uneasy'; *gorro* (es: *gorro*), for *gorra*, 'hat'; *papilla* (es: *papilla*), for *farinetes*, 'baby food'; *recapacitar* (es: *recapacitar*), for *reflexionar*, 'to reflect on'; *recés* (es: *receso*), for *pausa*, 'break'.

Phraseological units (PUs) are based on the free combination of words which, once linked, take on a new meaning. They are idiomatic and cultural structures. A language is a system of symbols whose decoding depends on the knowledge structure to which they ostensibly refer (Bransford & Johnson 1973 [1990]: 254). In this sense, *aguantar el tipus* (es: *aguantar el tipo*) or *fer el cor fort* are idiomatic combinations from two different cultures or knowledge structures, where the substitution of one for the other signifies a type of language subordination (Eliot 1984). *Ésadir* lists the following PUs in general use but not in the DIEC: *perdre els papers* (es: *perder los papeles*), 'to lose it' and *enfilar-se per les parets* (es: *subirse por las paredes*) 'to go up the wall'; however, it does not have a main entry for the equivalent expression with both meanings, *treure foc pels queixals* 'to be a spitfire', which is relegated to a secondary form. Position in the CCMA style guides is determined by linguistic and cognitive distance from Spanish, thus moving away from reproducing solutions in the dominant language. *Ésadirs'* acceptance of the general use phrase *donar la campanada* (es: *dar la campanada*), for *deixar bocabadat* 'to cause a stir', is another example of full or structural PU calques from Spanish. In turn, a set of PUs was incorporated into the standard dictionary in 2007: *anar a menys* (es: *ir a menos*) for *anar a mal borràs* 'to come down in the world'; *com a molt* (es: *como mucho*) for *a tot estirar* 'at most'; *en un tancar i obrir d'ulls* (es: *en un abrir y cerrar de ojos*) for *en un tres i no res* 'in the blink of an eye'; *prendre el pèl* (es: *tomar el pelo*) for *aixecar la camisa* 'to pull someone's leg'. It is important to remember that swapping PUs is a further symptom of weakening linguistic structures in the decaying language (García del Toro 2000: 228). With regard to the lexical element *gorro/gorra* (Figure 8.1), *ésadir* has over-differentiated rather than extended the semantic field held by the authentic word in the Catalan linguistic system, pointing to a dependence on Spanish: '*gorro*. Peça de vestir que cobreix i abriga el cap, generalment de tela o de punt, amb forma arrodonida i sense visera (a diferència d'una gorra)'[22] 'An item of clothing covering and sheltering the head, generally made from cotton or wool, with a rounded shape and no peak (unlike a cap)'.[23] This is one of the tools used by recessive languages to adjust their own system to innovations from language contact.

Finally, lexical (and semantic) interference is highly represented under the informal use label ('paraules i expressions pròpies de registres informals, contextos col·loquials i estils pròxims a la llengua espontània' ('words and expressions from informal registers, vernacular contexts and styles similar to spontaneous language').[24] The following is a qualitative sample of Hispanicisms incorporated by *ésadir* as the main term,[25] while relegating the authentic equivalent still used in the Catalan-speaking community to a secondary position of diatopic and diaphasic variation: *alfombra* (es: *alfombra*) for *estora/catifa* 'carpet, rug'; *apalancar-se* (es: *apalancarse*) for *escarxofar-se* 'sprawl, spread out'; *aplatanat*

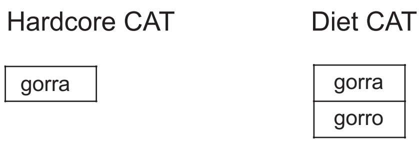


Figure 8.1 Semantic over-differentiation: *gorra/gorro*.

22 The distinction between *gorra/gorro* is an instance of over-differentiation that would likely not have occurred without interference from Spanish.
23 <https://esadir.cat/entrades/fitxa/node/gorro> (accessed on 3 January 2023).
24 <https://esadir.cat/lexic/usgenrest> (last accessed 10 December 2022).
25 These words may be viewed in the informal uses section: <https://esadir.cat/lexic/usrestringit> (last accessed 10 December 2022).

(es: *aplatanado*) for *aixafat* 'lethargic'; *atontat* (es: *atontado*) for *aturat/encantat* 'stunned'; *barco* (es: *barco*) for *barca/vaixell* 'boat'; *bolso* (es: *bolso*) for *bossa* 'handbag'; *calbo* (es: *calvo*) for *calb* 'bald'; *canguelo* (es: *canguelo*) for *cagalló/cagarines* 'jitters'; *currar* (es: *currar*) for *pencar* 'to graft'; *cutre* (es: *cutre*) for *deixat/lleig/brut* 'shabby, tacky'; *disfrutar* (es: *disfrutar*) for *xalar/fruir/gaudir* 'to enjoy'; *endollar/enxufar* (es: *enchufar*) for *tenir bo* 'to pull strings'; *esconyar-se* (es: *escoñarse*) for *fotre's* 'to smash, hurt oneself'; *flipar* (es: *flipar*) for *xalar/badar els ulls com a taronges* 'to freak, flip out'; *rotllo* (es: *rollo*) for *murga/punyeta* 'bore'; *tonto* (es: *tonto*) for *pallús/ruc/ximple/beneit* 'silly, foolish, dumb'.

Thus, *ésadir* uses several mechanisms for words mimicking the vernacular in medium to low formality, in particular replacing authentic words with Hispanicisms: *apalancar-se, calbo* … ; or over-differentiation, *bolso* ('handbag'), *bossa* (any other bag). These phenomena embody interference from bilingual speech and the principle of least effort and normally lead to lexical and semantic impoverishment in the recessive language. This involves the use of omnibus words encompassing two or more words in the host language (Payrato 1985: 113; Dalmau 1935: 80), e.g. *tonto* and *cutre* (Figures 8.2 and 8.3), which *ésadir* limits to the informal register and notes they are not listed in the DIEC.

The *El Periódico* style guide also lists Hispanicisms not in the DIEC as standard or general use, whereas *ésadir* restricts or limits their usage, e.g. *vivenda* (es: *vivienda*) instead of *habitatge*, which for *El Periódico*

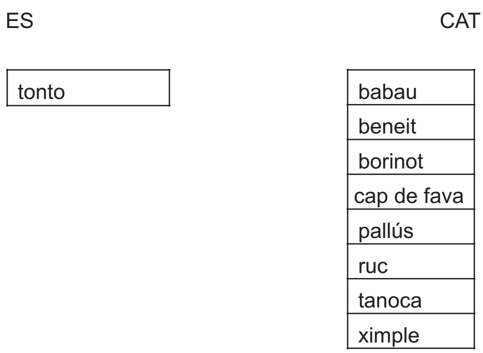


Figure 8.2 Lexical substitution (semantic impoverishment): *tonto*.

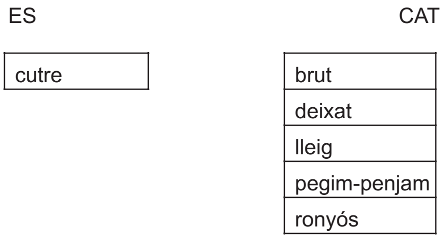


Figure 8.3 Lexical substitution (semantic impoverishment): *cutre*.

is the preferred term; in contrast *ésadir*, while listing *vivenda*, leans towards *habitatge* 'housing'. Furthermore, *El Periódico* adds a range of Hispanicisms not listed on the CCMA language portal: *birlar* (es: *birlar*) for *pispar* 'to nick, pinch, steal'; *bírria* (es: *birria*) for *nyap/bunyol/esguerro/rave* 'mess, botch'; *bronca* (es: *bronca*) for *brega* 'row, fight, quarrel'; *florero* (es: *florero*) for *fer bonic* 'to embellish, beautify in a metaphorical sense'; *posavasos* (es: *posavasos*) for *rodal* 'coaster'. It is important to note the term *posavasos* that *El Periódico* structurally calques directly from Spanish. *Ésadir* chooses a semantic calque, *sotagot*, and reproduces the Spanish model with lexical elements from Catalan (es: *posa*, cat: *sota*; es: *vaso*, cat: *got*). The *ésadir* search engine excludes the term *rodal*, which is listed as a secondary option in the main search for *sotagot*. A search for *rodal* takes you to the definition for a corona around a star, alongside the word *halo*.

In terms of adopting and adapting neologisms, the style guide has different criteria that fall either into the *diet* or *hardcore* camp. One should be aware that 'els neologismes es formen per alteració semàntica de mots ja existents, que adquireixen noves accepcions, o bé per generació de mots nous mitjançant diversos procediments' ('neologisms are formed by semantic alteration of existing words and acquire new acceptations, or by creating new words through different procedures'; Bruguera 1985: 97–8). *El Periódico* prefers to either accent loanwords (*dòping, dríbling, estànding*) or use loan translation or calques – taking another language model (English-Spanish-Catalan) and applying elements from the host language such as root translation or morphological adaptation, e.g. *barranquisme* (es: *barranquismo*) 'canyoning'; *parapent* (es: *parapente*) 'paragliding'. In this sense, the dominant language stands between the wider world and the recessive language (Aracil 1983). *Ésadir* lists accented loanwords, e.g. *estànding, doping*, and also uses calques in line with the morphology of the target language, e.g. *driblatge* (listed in the DIEC).

The style guide from *El Punt* newspaper (an unpublished work from linguist Pere Drou in the early 2000s) adapts neologisms by applying loan translations that align with the target language structure, without using Spanish as a relay language.[26] *El Punt's* approach is mentioned since its guide appeared at the same time as the *El Periódico* style guide and highlights the papers' contrasting solutions. *El Punt* notes that its scope is not limited to one region but covers the entire Catalan-speaking community. It proposes a prescriptive standard media solution for users, stripping out certain foreign words (Anglicisms) and calques (Hispanicisms): 'En una situació de subordinació lingüística com la del català, l'actitud lingüística dels mitjans de comunicació és més transcendental. Ens hi juguem no solament la llengua del futur, sinó el futur de la llengua' ('In a context of language subordination such as with Catalan, the linguistic attitudes of the media are of greater importance. Both the language of the future and the future of the language are at risk'; *El Punt*: 14). In this sense, it proposes *descens de barrancs* for *barranquisme* 'canyoning'; *ala de pendent* for *parparent* 'paragliding'; *dopatge* for *dòping* 'doping'; *driblatge* for *dríbling* 'dribbling'; and *habitatge d'alt nivell* for *estànding* 'luxury home' (Calafat 2010: 47).

3.2 Linguists in favour of the authentic word

This section looks at two linguists who promote a standard language through their standout media presence on social networks. Jordi Badia's

26 The examples from this newspaper were collected in 2009 and published in Calafat (2010). *El Punt* (established in 1979) merged with *Avui* (established in 1979) in 2011. This led to the creation of *El Punt Avui*.

Twitter account, under the tagline *salvemelsmots* ('Let's Rescue Our Words'), alerts followers to authentic Catalan lexicon being affected and pressured by Hispanicisms, which media channels support by defending a hybrid Catalan, especially for registers simulating orality:[27]

> Així mateix, els mitjans de comunicació han contribuït, dissortadament, a esborrar la diferència de registres [...], observem una barrija-barreja de formalitat i informalitat. En el cas català, els registres més malmesos són els informals: el col·loquial i el vulgar. (Badia 2022: 11)
>
> 'The media have had an outsized impact on erasing differences in register [...], formality and informality are swept aside. For Catalan, informal registers are the most affected: colloquial speech and slang'.

For instance, *El Periódico* includes the loanword *gilipolles* (es: *gilipollas*).[28] In a tweet from 21 January 2023, Badia showcased his expert corrective approach to followers by writing about the Catalan terms covering the semantic field as the omnibus word: *carallot, gamarús, curt de gambals, malparit,* 'pillock, nitwit, dimb-witted, bastard'.[29] Badia's other aim is to raise awareness about pressure from Hispanicisms that impact language choices, e.g. where Catalan has two equivalent words, the one most resembling the Spanish lexeme wins out socially. In this sense, where Catalan has two words (mostly partial synonyms) for the same idea, the principle of least effort prevails amongst bilingual speakers who constantly work in two languages, e.g. *estrangular/escanyar* (es: *estrangular*) 'to strangle'. Although both are authentic traditional terms first documented in the thirteenth century, *escanyar* has seen lower use and does not normally appear in the media, leading to a subsequent loss of semantic distinction for Catalan. *Escanyar* has other meanings too. In addition to covering the semantic field of its sibling term (*estrangular*), its other nuances cover to over-constrict the neck, make it difficult to speak due to emotion and

27 Badia published the book *Salvem els mots,* ('Let's Rescue Our Words') in 2022, a collection of his lexicological tweets from the two years prior to publication.

28 *El Periódico* puts *gilipolles* in italics and reserves it for vernacular language. It provides *estúpid* as an equivalent term in the entry, whose formality level is higher (2002: 170).

29 Tweet (21 January 2023), <https://twitter.com/cerclevall>

the figurative sense of *escanyapobres* (*usurer*) 'loan shark'; *escanyabarbes* (*barber dolent*) 'bad barber' (Badia 2022: 86–7).

Badia looks at and promotes certain words and PUs. He questions phrases – whether Hispanicisms or traditionally Catalan PUs with a similar structure to Hispanicisms – used to the detriment of an equivalent term that is more structurally distant from Spanish, such as *a cegues* (es: *a ciegas*) instead of a *ulls clucs* 'in the dark';[30] *de cop/de sobte* (es: *de golpe*) 'suddenly' (Badia 2022: 55); *fet pols/fet una coca* (es: *hecho polvo*) 'shattered' (Badia 2022: 94). *Ésadir* includes the phrase *estar fet pols* in the informal use register, while dictionaries do not list it. Other PUs that Badia mentions include *no dir ni piu/no badar boca* (es: *no decir ni pio*) 'to not breathe a word';[31] *no importa/tant se val* (es: *no importa*) 'no matter';[32] *prendre's s'ho a pit/ prendre's ho a la valenta* (es: *tomárselo a pecho)* 'to take to heart' (Badia 2022: 166–7); *sense fonament/sense solta ni volta* (es: *sin fundamento*) 'baseless', (Badia 2022: 185–6);[33] *posar-se guapo/endiumenjar-se o empolainar-se* (es: *ponerse guapo*) 'to get dressed up' (Badia 2022: 160–1); *posar en marxa/engegar* (es: *poner en marcha*) 'to set in motion', (Badia 2022: 160–1). No dictionary admits the phrase *posar-se guapo*. As for *posar en marxa*, it is included as preferential by *El Periódico*, but not explicitly admitted by the DIEC, although the dictionary does use it in the explanatory text for certain definitions: in the entry for *accionar* 'to operate', it states '*posar en marxa un mecanisme*' ('to start up a mechanism'; Badia 2022: 161). This is a good example of mass media power and social normality.

With regard to lexemes whose resemblance to Hispanicisms covers two or more lexical and semantic terms in Catalan, as well as contributing to impoverishment in the recessive language, Badia mentions pairs that pre-date the similar Spanish term, e.g. *captivar/corprendre* (es: *cautivar*) 'to captivate' (Badia 2022: 46); *firmar/signar* (es: *firmar*)

30 Tweet (28 September 2018), <https://blocs.mesvilaweb.cat/2018/10/?author_name=jbadia>
31 Tweet (29 October 2020), <https://blocs.mesvilaweb.cat/2020/12/?author_name=jbadia>
32 Badia (2022: 138–9).
33 Badia (2022: 185–6).

'to sign' (Badia 2022: 95); *fracassar/fer figa* (es: *fracasar*) 'to fail' (Badia 2022: 96–7); *lamentar/saber greu* (es: *lamentar*) 'to regret' (Badia 2022: 119); *lamentarse/plànyer-se* (es: *lamentarse*) 'to complain', (Badia 2022: 119). The *desar/guardar* pair (Figure 8.4) is a clear example of the loss of semantic distinction (*desar*: to leave something somewhere when not in use; *guardar*: preserve, maintain, save), which gets lost when merging both meanings into *guardar* (es: *guardar*) 'to save' (Badia 2022: 103–4). The CCMA language portal lists *guardar* under the 'overuse' label, stating that depending on the context, *desar* is the preferred term, although it does not come up in the search engine.

The term *lidiar* (es: *lidiar*) 'to grapple'[34] is a good example of interference in the strictest sense, which feeds terms with a Spanish structure into the recessive language. Catalan has several words covering the semantic field of this neologism, i.e. *encarar-se, lluitar, oposar-se, trampejar*. The *ésadir* style guide, however, admits it in general use as suitable for covering all functional registers. In contrast, the DIEC does not admit it.

Finally, it is important to mention an instance of semantic and lexical reduction applied to both Hispanicisms and traditional Catalan terms. The phrase *portar-se bé* matches two Spanish collocations (*llevarse bien* and *portarse bien*). Catalan has *avenir-se* 'to get along' for the former and *fer bondat* 'to be kind' for the latter. The phrase *portar-se bé* is listed in the DIEC under both meanings (Badia 2022: 158).

The IEC released 640 changes in late 2020 in the standard dictionary. These changes added lexical and semantic interferences in the strictest sense that affected phrases and collocations. In his article, *Encerts, dubtes i mancances en les modificacions del diccionari del IEC* ('Sucesses, Doubts

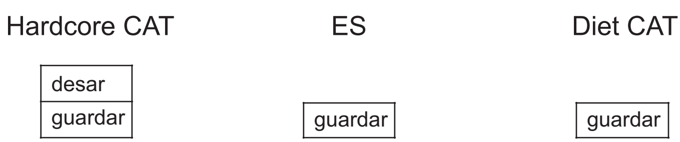


Figure 8.4 Blurring a distinction: *guardar*.

34 Tweet (28 November 2020), <https://blocs.mesvilaweb.cat/2020/12/?author_name=jbadia>

and Shortcomings in Changes to the IEC Dictionary') published in the online newspaper Vilaweb,[35] Badia mentioned the new PUs, as well as other examples, in the dictionary: *a (un) tir de pedra* (es: *a tiro de piedra*) 'a stone's throw'; *a la curta o a la llarga* (es: *a la corta o a la larga*) 'in the short/long'; *en entredit* (ES: *en entredicho*) 'into question'; *fet un fàstic* (es: *hecho un asco*) 'filthy'; *posar en antecedents* (es: *poner en antecedentes*) 'to set the scene'; *importar-li un rave* (es: *importarle un rábano*) 'to not care'. The authentic Catalan terms for each of these additions are a *quatre passes*, *tard o d'hora*, *en dubte*, *fet una pelleringa*, *advertir per endavant* and *donar-se-li'n un clau*.

Gabriel Bibiloni, a lecturer at the University of the Balearic Islands, publishes articles in the press and on his blog (*El blog de Gabriel Bibiloni*) with linguistic explanations for the lexicographical and grammatical proposals he releases on Twitter. In the blog piece *El castellanismes de la ràdio i la televisió públiques de Catalunya* ('Hispanicisms on Catalan Public Television and Radio') from 8 June 2016,[36] he sets out neologistic Hispanicisms affecting the grammatical structure of Catalan collocations which have been admitted as general use in the *ésadir* style guide: *anar de compres* (es: *ir de compres*) for *anar a comprar* 'to go shopping'; *anar de rebaixes* (es: *ir de rebajas*) for *anar a les rebaixes* 'to hit the sales'; *anar per lliure* (es: *ir por libre*) for *compte propi* 'to go it alone'.

It is interesting to highlight the contrast between the neologisms he proposes and those terms which are accepted by media channels. The present article contains a representative sample from the linguist which stands in opposition to general use entries from *ésadir*. Bibiloni uses neologistic solutions akin to those from other Romance languages, while respecting the Catalan linguistic system. The proposals avoid systemic calques from Spanish and align with interference in the widest sense. As

35 Blog (03 December 2022), <https://www.vilaweb.cat/noticies/modificacions-diccionari-iec/>

36 Blog (8 June 2016), <https://bibiloni.cat/blog/els-castellanismes-de-la-radio-i-la-televisio-publiques-de-catalunya/#comments> (last accessed 2 February 2023).

initial examples, let us first look at *pastera*[37] 'small boat', *botelló*[38] 'street drinking' and *cribratge*[39] 'screening'. Bibiloni's solutions are *barca* so as to avoid the semantic Hispanicism of *pastera* (es: *patera*), a term designating a small boat carrying migrants from Africa to Europe.[40] Standard Spanish has extended the meaning of *patera* which, like the Catalan term *pastera*, generally means a small keelless boat with a flat bottom. The neologism applies to a very different type of boat with a highly specific use (this meaning does not appear in the DIEC). The *ésadir* style guide opts for a general use listing for the phonetically adapted Hispanicism *patera* and the semantic calque *pastera*. In turn, *botelló*, meaning '(young people) drinking alcohol on the street', is a loanword from the Spanish *botellón*. *Ésadir* lists it as general use, alongside the adaptation *botellot* and the structural calque *botellada*. Bibiloni proposes the latter as an alternative since the lexeme *botella* 'bottle' is a Hispanicism rooted in Catalan and included in the DIEC. The third element is a semantic calque from Spanish, disguised as a Latinism: *cribatge* meaning a medical screening of a population to detect those with a disease or abnormality. This neologism was included in the DIEC in 2021. Bibiloni makes no alternative proposal and merely explains its linguistic interference in the strictest sense on his blog post from 27 August 2020, *'Cribratge', la dependència permanent de l'espanyol* ('Screening, the Permanent Dependence on Spanish'):

> Cribratge [...]. No existeix en català el mot *cribrar* – ja tenim *porgar* i *garbellar* – i, per tant, no pot fer un derivat. Lògicament no és un mot evolutiu. I tampoc és un neologisme pres d'una altra llengua dels que són acceptats pel seu caràcter internacional. Ni és un llatinisme, perquè en llatí no hi ha cap mot que es pugui identificar amb cribratge. És el mot espanyol cribado amb maquillatges catalans i

37 Tweet (26 December 2019), <https://twitter.com/bibiloni/status/1210215197668974592>

38 Tweet (13 August 2021), <https://twitter.com/bibiloni/status/1426146081310711812>

39 Tweet (28 August 2020), <https://twitter.com/bibiloni/status/1299409422138191877>

40 Blog (4 July 2009) *Pastera, sense rumb ni timó*, 'The Rudderless Pastera', <https://bibiloni.cat/blog/index.php?s=pastera&image.x=0&image.y=0> (last accessed 20 January 2023).

> llatins. Una mostra més de la nostra pertinaç dependència de l'espanyol, que ens desbarata sistemàticament la llengua.
>
> '*Cribratge* [...] the verb *cribrar* does not exist in Catalan – we already have *porgar* and *garbellar* – and, therefore, no derivative is possible. Logically, it is not an evolutionary word. Neither is it a neologism from those admitted due to their international nature and taken from another language. Nor is it a Latinism, since there is no word one can link it to in Latin. It is the Spanish term *cribado* with a Catalan and Latin make up. A further example of our persistent dependence on Spanish, a language that systematically disrupts our own'.[41]

Bibiloni's online critique of the loanword (*patera*) and calque (*pastera*) had an impact on some journalists. The linguist himself explained this in a tweet from 26 December 2019,[42] stating that the Catalan edition of the Balearic news bulletin on the national public broadcaster TVE (local programming) used the word *barca* and not *pastera*. Further examples of the linguist's media impact include the semantic calque *rescatar* (es: *rescatar*) for *salvar* 'to rescue' (cf. Figure 8.5). Bibiloni published the following tweet on 2 July 2018:

> Aquests dies es parla molt del rescat de gent perduda enmig de la mar. Doncs no: rescatar és obtenir l'alliberament d'un ostatge pagant diners. Res més. Si no hi ha paga no hi ha rescat. Els nàufrags i els excursionistes perduts són salvats [...]
>
> 'Recently we've heard a lot about freeing [*rescatar*] castaways and travellers lost at sea. No: *rescatar* means freeing a hostage by paying a ransom. Nothing else. If you don't pay, there's no *rescat*'.[43]

The term was then adopted by a programme on TV3. Jordi Badia tweeted on 14 September 2019 that thanks to Bibiloni's work, some journalists noticed the calque: 'En Gabriel Bibiloni ens va fer veure que l'Open Arms no rescatava immigrants, sinó que els salvava. I avui al FAQS [...] hem sentit gairebé tothora salvar' ('Gabriel Bibiloni made us see that the Open

41 Blog (27 August 2020), <https://bibiloni.cat/blog/cribratge-la-dependencia-perenne-de-lespanyol/ (last accessed on 3 April 2022)>

42 Tweet (26 December 2019), <https://twitter.com/bibiloni/status/1210215197668974592>

43 Tweet (2 July 2018), <https://twitter.com/bibiloni/status/1013812633328390145>

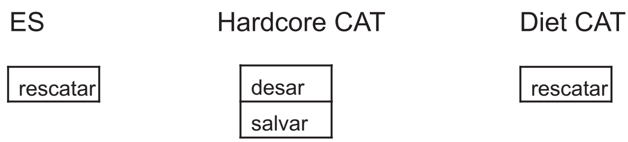


Figure 8.5 Blurring a distinction: *rescatar*.

Arms wasn't freeing [*rescatar*] migrants but rescuing [*salvar*] them. And today in the FAQs [...] we've just seen *salvar*').[44]

Finally, we would like to provide a small sample of the neologisms proposed by Bibiloni, based on a comparison of the term with two or more Romance languages, and contrasted with those in *ésadir* with similar features to the Spanish term:

- *màscara* (*ésadir*: *mascareta*): *mascarilla* (es), *masque* (fr), *máscara* (pt), *mascherina* (it), *mask* (en).[45]
- *lletra morta* (*ésadir*: *paper mullat*): *papel mojado* (es), *lettre morte* (fr), *lettera morta* (it) 'dead letter'.[46]
- *repressalles* (old Catalan), (*ésadir*: *represaliar*, *prendre represàlies*): *represàlies* (es), *représailles* (fr), *rappresaglie* (it), *reprasalhas* (oc) 'reprisals'. The term *represàlies* is included in the DIEC yet the dictionary does not include the verb *represaliar* 'retaliate'.[47]

There are also online groups of linguists who dedicate their time to teaching authentic words and expressions to the public to replace Hispanicisms. This is the mission of the group *El català com cal* that has over 31,300 Twitter followers. They call on speakers' conscience (loyalty in linguistic

44 Tweet (14 September 2019), <https://twitter.com/jbadia16/status/1172983779687829504?ref_src=twsrc%5Etfw>
45 Blog (22 August 2020), *Sobre màscares i* mascarats ('On Masks and Traitors'), <https://bibiloni.cat/blog/sobre-mascares-i-mascarats/>
46 Tweet (7 January 2020), <https://twitter.com/bibiloni/status/1214639552293412871>
47 Tweet (20 November 2020), <https://twitter.com/bibiloni/status/1329855448837664770>

use and form) to eradicate the unilateral dependence on Spanish in everyday speech, particularly focusing on PUs and clitics – elements most commonly replaced by omnibus terms, e.g. *menjar-te el tarro* (es: *comerte el tarro*) for *capficar-s'hi*, *barrinar-n'hi, encaparrar-s'hi* 'to think hard',[48] or *liar-la* (es: *liarla*) for *fer merder, embolicar la troca* 'to complicate things'.[49] They also run a blog under the same name where a large selection of content deals with vernacular Hispanicisms in the media, e.g. *els castellanismes de la setmana, els castellanismes més freqüents, els més recents i els castellanismes a la televisió i ràdios catalanes* ('Hispanicisms of the week, the most common and most recent Hispanicisms on TV and radio'). It reports and corrects language interference in journalism. Some terms are also included in *ésadir* under restricted use, e.g. *pringar* (es: *pringar*) for *fotre's*, *llepar*, *tocar el rebre* 'to be screwed'; *pringat* (es: *pringado*) for *llefiscós*, *empastifat* 'fool/idiot';[50] *carinyo* (es: *cariño*) for *estimació*, *afecte* 'affection'.[51] Moreover, the group also points out Hispanicisms in the DIEC, e.g. *entorpir* (es: *entorpecer*) for *torbar*, *entrebancar* 'to hinder'.[52]

4. Conclusions

Style guides and language professionals have two different visions of Catalan in the media. On the one hand, some want language to reflect contact between Catalan and Spanish, especially in medium to low formal registers. They argue the standard should reflect the language spoken

48 Tweet (26 June 2021), <https://twitter.com/elcatalacomcal/status/1408755747530358785>

49 Tweet (11 October 2019), <https://twitter.com/elcatalacomcal/status/1182610525231964160>

50 Tweet (20 February 2020), <https://twitter.com/elcatalacomcal/status/1230461126279860229?s=11>

51 Tweet (8 February 2023), <https://twitter.com/elcatalacomcal?lang=ca>

52 Tweet (14 February 2020), <https://twitter.com/elcatalacomcal/status/1228312262840061952?s=11>

on the streets by bilinguals. The style guides on the Catalan Broadcasting Corporation language portal (*ésadir*) and from *El Periódico* align with this approach, albeit at differing levels. *El Periódico* particularly stands out for its massive inclusion of Hispanicisms. On the other hand, and in the name of language loyalty in form and use, language professionals (university lecturers, linguists, editors from other media outlets) are committed to purging the media of Hispanicisms and use social media for language education. Experts address users and encourage them to speak correctly. Both visions acknowledge academic authority, although they have a different relationship to it. The first would like Hispanicisms and neologisms to be included in the DIEC which, across its different editions (1995, 2007, 2022), has indeed added terms from direct contact with Spanish. The second critiques the inclusion of these words and PUs in the standard dictionary.

The qualitative sample in this article contains lexical and semantic innovations whose inclusion in style guides illustrate the incorporation of language phenomena due to a unilateral dependence on another language (Catalan subordination to Spanish). In turn, it highlights lexical elements that have been admitted into the standard dictionary. Highly cultural PUs and words proposed as replacements for Catalan terms (*cutre* for *deixat/lleig/brut*) often involve blurring the lexical and semantic distinction of Catalan; this mostly affects medium to low formality although it can also impact medium to high formal registers, albeit to a lesser extent. *Ésadir*, for example, includes non-standard terms for general use: *gorro*, *papilla*, *aguantar el tipus*.

Proposed neologisms in the form of loanwords or calques also define the ideological linguistic position of media outlets: *El Periódico* leans towards neologistic loanwords with adapted accents, e.g. *estànding*, which contrasts with other style guides (*El Punt*) that lean towards structural calques, i.e. loan translation adapted to the linguistic elements of the host language, such as *habitatge d'alt nivell*. *Ésadir* has a preference for general use loanwords such as *estànding*, although it does include double entries in the form of loanwords and adapted calques, e.g. *dòping/dopatge*; *dríbling/driblatge*.

Linguists who take part in online discussions focus on three main points: (a) promoting common lexical and PUs in informal speech within the Catalan language community; (b) reporting Hispanicisms in journalism and interference in everyday speech; and (c) criticising the inclusion of words and expressions in the DIEC that originate from interference in the strictest sense. Nonetheless, they also cover other areas. Badia, for example, works on authentic Catalan word pairs, highlighting how lexemes similar to Spanish terms triumph in the media and often lead to the blurring of semantic contrasts. Bibiloni, a university lecturer, stands out for his neologistic proposals where he highlights the dominant language standing between the wider world and the recessive language through neologisms originating from Spanish interference yet incorporated into Catalan.

In short, we look at double linguistic solutions for standard Catalan in the media. This situation emerges from two distinct ideologies, i.e. on the one hand, Catalan's dependence on Spanish and, on the other, the independence of Catalan and its speakers with regard to the cultural and linguistic universe of the dominant language.

Bibliography

Aracil, Lluís V. 1983. *Dir la Realitat*. Barcelona: Països Catalans.

Argenter, Joan. 2020. Languages in Contact: A Socio-cultural Approach. In Joan Argenter, Joan & Jens Lüdtke (eds). *Manual of Catalan Linguistics*. Berlin, Boston: de Gruyter. 597–628.

Calafat, Rosa. 2010. *Torcebraç entre dues cultures. De l'ecosistema de les llengües: de discursos i de percepcions*. Barcelona: IEC.

Calafat, Rosa. 2020. El centre i la perifèria de la llengua catalana als llibres d'estil: Catalunya i les Illes Balears. *Revista de Llengua i Dret* 73, 69–81.

Badia, Jordi. 2022. *Salvem els mots*. Barcelona: Rosa dels Vents.

Benet, Josep. 1995. *L'intent franquista de genocidi cultural contra Catalunya*. Barcelona: Publicacions Abadia de Montserrat.

Bibiloni, Gabriel. 1997. *Llengua estàndard i variació lingüística*. València: 3i4 edicions.

Boix, Emili & Francesc X. Vila. 1998. *Sociolingüística catalana*. Barcelona: Ariel.
Bransford, John D. & Marcia K. Jonhson. 1990. Consideraciones sobre algunos problemas de la comprensión. In Francisco Valle, Susana del Viso, José M. Igoa & Fernando Cuetos (eds). *Lecturas de psicolingüística. Comprensión y producción del lenguaje*. Madrid: Alianza. 229–63.
Bruguera, Jordi. 1985. *Història del lèxic català*. Barcelona: Enciclopèdia Catalana.
Camps, Oriol. 2007. Sobre la qualitat de la llengua dels mitjans de la CCRTV. *Quaderns del CAC* 28, 3–11. <https://www.cac.cat/documentacio/la-qualitat-la-llengua-els-mitjans-audiovisuals> (accessed 30 November 2024).
Colon, Germà. 1989. *El español y el catalán, juntos y en contraste*. Barcelona: Ariel.
Dalmau. 1935. Poliglotisme passiu. *La Revista XXI.* 74–81. <https://filcat.uab.cat/assaigrepublica/project/poliglotisme-passiu/> (accessed 7 June 2024).
Darder, Laia. 2015. La variació lingüística en els mitjans audiovisuals com a estratègia conscient. *Treballs de Sociolingüística Catalana* 25, 303–16.
Diki-Kidiri, M. 1999. Terminologie pour le développement. In Teresa Cabré (ed.), *Terminología y modelos culturales*. Barcelona: Institut Universitari de Lingüística Aplicada de la Universitat Pompeu Fabra. 67–74.
Dressler, Wolfgang. 1987. La mort de les llengües. *Límits* 3, 87–97.
Drou, Pere. 2000. *Llibre d'estil d'El Punt: model de llengua*. Unpublished manuscript.
Eliot, Thomas S. 1984. *Notas para la definición de la cultura*. Barcelona: Bruguera.
Esteve, Francesc, Josep Ferrer, Lluís Marquet & Juli Moll. 1998. El diccionari de l'Institut. Una aproximació sistemàtica. *Els Marges* 60, 5–96.
Galisson, Robert. 2000. La pragmatique lexiculturelle pour accéder autrement, à une autre culture, par un autre lexique. *Mélanges Crapel* 25, 47–73.
García del Toro, C. (2002). Traduir els fraseologismes entre català i espanyol. In Vicent Salvador & Adolfo Piquer (eds). *El discurs prefabricat. Estudis de fraseologia teòrica i aplicada,* Castelló de la Plana: Universitat Jaume I. 227–40.
Gifreu, Josep. 2003. Els mitjans audiovisuals, laboratoris de segona oralitat. In Joan Martí & Josep M. Mestres (eds). *L'oralitat i els mitjans de comunicació: actes del seminari del CUIMPB-CEL*. Barcelona: Institut d'Estudis Catalans. 137–47.
Kailuweit, Rolf. 2002. Català heavy-català light: una problemàtica de la 'lingüística de profans'. *Zeitschrift für Katalanistik* 15, 169–82.
Lamuela, Xavier. 1987. *Català, occità, friülà: Llengües subordinades i planificació lingüística*. Barcelona: Quaderns Crema.
Levinson Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
Llibre d'estil de la Corporació Catalana de Mitjans Audiovisuals. 2011. <https://www.ccma.cat/llibredestil/> (accessed 7 June 2024).
Llibre d'estil El Periódico de Catalunya. 2003. Barcelona: Edicions Primera Plana.

Massanell, Mar. 2009. El col·loquial mediatitzat en les sèries de ficció i el català nord-occidental: un encaix difícil. In Margarida Bassols & Mila Segarra (eds), *El col·loquial dels mitjans de comunicació*. Vic: Eumo. 159–84.

Ninyoles, Rafael. 2017. *Conflicte lingüístic valencià. El País Valencià a l'eix mediterrani*. Alacant: Publicacions Universitat d'Alacant.

Payrató, Lluís. 1985. *La interferència lingüística. Comentaris i exemples català-castellà*. Barcelona: Curial-Publicacions de l'Abadia de Montserrat.

Perea, M. Pilar. 1998. De re lexicografia. *Els Marges* 60, 96–120.

Pla, Albert. 1989. Problemes dels correctors i assessors dels mitjans de comunicació. In Joan Martí, Lídia Pons & Joan Solà (eds). *Actes de les Terceres Jornades d'Estudi de la Llengua Normativa*. Barcelona: Publicacions de l'Abadia de Montserrat. 214–59.

Pradilla, Miquel A. 2020. Social and Functional Variation in Catalan. In Joan Argenter & Jens Lüdtke (eds), *Manual of Catalan Linguistics*. Berlin, Boston: de Gruyter, 397–420.

Prats, Modest, August Rafanell & Albert Rossich. 1990. *El futur de la llengua catalana*. Barcelona, Empúries.

Ruaix, Josep. 1996. *Diccionari auxiliar*. Moià: Ruaix editor.

Vallcorba, Cercle. 2022. *El català, on va?* Barcelona: Nexum edicions.

Vernet, Jaume. 2004. El marc constitucional i estatutari de la llengua a debat: balanç i perspectives. *Llengua, societat i comunicació* (LSC) 2, 39–42.

Weinreich, Uriel. 1953. *Languages in Contact. Findings and Problems*. New York: Publications of the Linguistic Circle of New York. [Catalan edition, 1996. Valencia: Bromera].

SIMON PRÖLL

9 The two faces of Standard German: Structural consequences of nativisation

ABSTRACT
This paper shows that, after centuries without a single native speaker, Standard German is transitioning to a fully natural language, undergoing undirected language acquisition processes by a critical number of speakers. This nativisation process leads to an increasing rift between the grammatical systems of written and spoken Standard German. Evidence for this comes from an ongoing shift from a non-standard suprasegmental system to the Standard German system in south German speakers: While the non-standard system allows accented syllables in bisyllabic words to be light, Standard German requires them to be heavy. Furthermore, there is clear evidence for emerging differences in the morphological systems of written and spoken Standard German that cannot be explained by mere performance phenomena: German TV newscasters display nominative/accusative syncretisms that are highly stigmatised in written codification. The catalyst for this can be traced to the nativisation of Standard German and its resulting reintegration into general trajectories of language change all over the Germanic language area.

1. Introduction: Stages of standardisation in German (and elsewhere)

For this article, I will focus on what could be called the final stages of standardisation (or, depending on the underlying definition of standardisation, events that happen post-standardisation). I will concentrate on the case of German, where these stages can be illustrated with data from the current timeframe, showing that the process of gaining native speakers leads to a grammatical bifurcation of Standard German into a written and a spoken variety.

The standardisation process of German is by and large a relatively typical one. It follows the same basic process that has been well researched for many contemporary languages and that can reliably be observed in many others, despite some small individual deviations or characteristics of individual cases (cf. e.g. Durrell 1999; Auer 2005; Vandenbussche 2007): Variants or even whole varieties are selected for codification and functional elaboration. Variation is minimised, and a certain status is assigned to the resulting variety (cf. e.g. Haugen 1966: 933; Durrell 1999; Inoue 2006). Standardisation is usually initially based upon written language, and it is also oriented towards written language (cf. Haugen 1966: 929); spoken language only becomes an object of standardisation at advanced stages (if at all). Helmut Weiß (2005) outlines four consecutive 'ages' of standard languages (albeit with a special focus on German) to model this process (cf. Table 9.1):

In the beginning, the emerging variety in question is constructed purely for written use (phase 1). Then, to improve upon its functionality, it will begin to feature structures that are in line with spoken language (phase 2). This is the basis for functioning as an actual spoken communication system, meaning that the variety can now be spoken as a second language (phase 3, cf. also Auer 2005: 15–8). Finally, this spoken language system is acquired by speakers in a natural process as their first language (phase 4). This last step, from phase 3 to phase 4, is the focus of this paper. I propose to call this step the nativisation of Standard German, because it marks the transition from a language that is learned (as a second

Table 9.1 'Ages' of a standard language, following Weiß (2005).

phase	mediality	condition
1	written language	learned secondarily
2	written language	learned secondarily, oriented on spoken language
3	spoken language	learned secondarily
4	spoken language	acquired primarily

language) by older children and adults to a language that is acquired in first language acquisition.

This can be regarded as a return to fitness from a psycholinguistic angle: As a part of the standardisation process, the emerging standard – as a purely written phenomenon – was never subject to first language acquisition. This allows for the introduction of structures that might not occur or might not even be possible in first language acquisition contexts (cf. Weiß 2001, 2004, 2005). As Chomsky (1995: 51) states, standard languages usually are 'partially invented' and could structurally even 'violate the principles of language' (see also Weiß 2001: 88; Paul 1920/1995: 48, 404–22).

The process of nativisation should thus lead to structural changes: In the same way that structural change can be observed in pidgins as they creolise, the acquisition of a standard language as a L1 has the potential to trigger structural changes as they nativise, simply because L1 acquisition might require – or rather, enforce – features that contradict the 'invented' structure of a written standard language. Standard German as a written language is almost conclusively codified in the late twentieth and early twenty-first century, meaning that these changes could hardly affect the written variety, but only the spoken Standard – it is thus probable that nativisation might lead to a bifurcation of Standard German, into a written standard variety and a structurally distinct spoken standard variety. This raises several questions:

1. Is it possible to detect whether German is in phase 3 or phase 4 at any given time?
2. Can we find structural consequences of this phase transition?

Concerning question number 1, we would have to somehow answer the question whether there are native speakers of Standard German (Pröll 2021b) – or rather, how we can find out reliably. Section 2 will elaborate on this.

Regarding question number 2, it is important to clarify that we need not limit ourselves to differences that result directly from differences in mediality. Of course, a spoken standard will differ from a written one, simply because being spoken involves phonology and phonetics. The simplest case of a relation between written and spoken standard varieties is thus: written standard + phonological rules + phonetic phenomena = spoken standard.

This would mean that the spoken standard could be completely deducible from the written one (assuming we know the respective phonological and phonetic rules). By and large, it would be a waste of time researching spoken standard on its own if it were reducible to merely being a predictable realisation of the written standard. Thus, in Section 3, we shall be dealing with morphological differences, i.e. differences in grammar that are not simply due to modality, but rather show the emergence of two distinct grammars of Standard German.

2. The current state of nativisation

2.1 Background: Standard German syllable structure

How could we possibly detect whether speakers of German speak Standard German? The most promising approach would probably be to not look for

- single features bound to specific lexemes (that could be idiosyncratic) or
- features that are apparent to the speakers themselves (meaning that they could have the knowledge and the means to control them according to the situation).

Rather, we should be looking for a typological feature that is hard to control willingly or appears to be largely invisible to laymen. For this task, I chose syllable structure, as it probably satisfies both conditions at once.

The phonology and phonetics of spoken Standard German have in large parts been shaped by spoken Low German, a process which in itself is part of German standardisation history:[1] During phase 3, written High German (supplemented with only a few Low German loanwords) acquired distinct reading pronunciations by native speakers of both High and Low German dialects. The respective phonologies served as the basis. But whereas the High German speakers tended to continue to speak their base dialects up into the twentieth century, the Low German speakers more readily – for various social reasons beyond the scope of this paper (see e.g. Stellmacher 1981: 94; Polenz 1994: 143–44; König 2004) – adopted their reading pronunciation of written High German as a vernacular, over time passing it on to their children as an input in their L1-acquisition. The modern German orthoepy that is codified in (most) dictionaries on pronunciation[2] and realised by professional speakers thus shows features present in Low German base dialects, but not in High German. Most prominent probably is the contrast between aspirated vs. not aspirated that serves as a distinction for the sets of plosives (instead of a distinction via an intensity contrast in the High German varieties, cf. Schmidt & Vennemann 1985).

On a less prominent note, there is a distinction concerning what types of syllables are possible and what types are not. In multisyllabic words of Standard German, the syllable carrying stress has to be heavy (i.e. it must carry two morae). There must not be light stressed syllables – this is a constraint that cannot be violated. In cases like German *Mutter* [mʊtɐ] 'mother', this leads to ambisyllabic consonants. The consonant following a short vowel nucleus in the first syllable serves both as the coda for the first syllable as well as the onset for the second syllable (see Figure 9.1). Through weight by position (cf. Hayes 1989: 258), it gets assigned a mora, and thus the first syllable is heavy. In German (and

1 This is also supported by synchronic microtypological evidence that shows how modern Standard German is a product of Upper German and Low German-language contact with respect to morphosyntactic and phonological variables (cf. Pröll & Weber 2023).

2 The two recent editions of the Duden Pronunciation Dictionary (e.g. Kleiner & Knöbl 2023) move away from this by adopting a far more descriptive perspective on 'standardness'. However, the point of the Low German influence remains largely untouched by this.

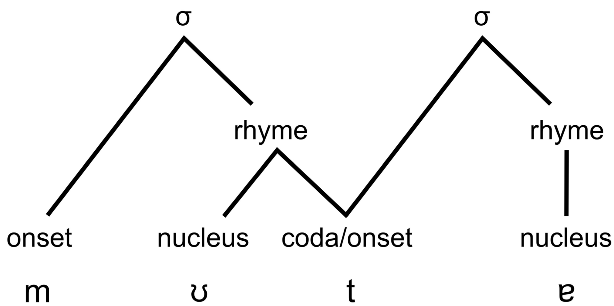


Figure 9.1 Standard German syllable structure of *Mutter* 'mother', featuring syllable joint.

Dutch) phonology, this is known as *Silbengelenk* 'syllable joint' (cf. van der Hulst 1985; Ramers 1992).

In other German varieties, for example Alemannic base dialects (such as Zurich High Alemannic) or regiolects (e.g. the East Swabian varieties that form the basis of our data used below), there is no such restriction concerning stressed syllables (a trait inherited from Old High German). A stressed syllable may be light, its vowel short and lax, the coda empty (cf. Auer 1989; Caro Reina 2019); thus, there is simply no need for a syllable joint.[3] (However, the onset following a stressed light syllable usually undergoes lenition, as can be seen in Figure 9.2.)

This means that in the Alemannic region, we can use the existence of an ambisyllabic consonant as a litmus test of whether we are dealing with a Standard German suprasegmental system or not.

3 Some High Alemannic dialects feature geminates, but this is another issue and is to be understood as a property of a segment, not as a suprasegmental – conspicuously, having geminates blocks the occurrence of syllable joints in languages, cf. van der Hulst (1985).

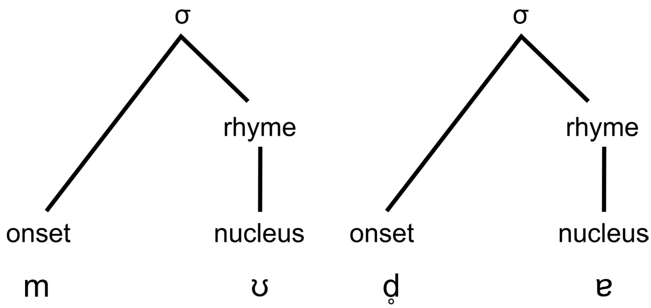


Figure 9.2 Alemannic syllable structure of *Mutter* 'mother' featuring lenition.

2.2 Evidence: Kindergarten

Pröll, Freienstein & Ernst (2016) conducted a study on ambisyllabicity in an East Swabian (i.e. Alemannic) kindergarten, with thirty-five children from three to six years as informants. The children individually had to guess, shout out and clap to bisyllabic lexemes that were presented to them in the form of pictures. The intention was to create a breaking point, to bring them to articulate both syllables with an unnaturally long transition phase between them (thus the clapping, in order to encourage syllabification through a rhythmic structure), to see (a) whether they would realise some form of a syllable joint at all and (b) how they would actually bridge the gap between the two syllables. Let us concentrate on the results for the continuants /m n ŋ l s f ʃ / (featured in the lexemes *Hammer* 'hammer', *Sonne* 'sun', *Finger* 'finger', *Teller* 'plate', *Messer* 'knife', *Affe* 'ape', *Flasche* 'bottle', *Knochen* 'bone').

Unsurprisingly, age matters: There are considerable developmental differences in the children's ability to access sub-lexical information, in this case to segment lexemes into syllables. This hints at a change from a holistic approach in lexical retrieval to one that is supported by (supra-) segmental information (see Pröll, Freienstein & Ernst 2016: 164–8). Still, if we ignore these individual trajectories and look at the variation concerning syllable structure as a whole, there is a clear pattern:

- 30.4 % of the instances show a short vowel in a light, open syllable. This is impossible in Standard German, so these children clearly use an Alemannic system.
- 6.1 % of the instances feature a long vowel in a heavy, open syllable, meaning that the children had to apply compensatory lengthening (cf. Hayes 1989) to the vowel. Lengthening makes the stressed syllable heavy. This would not be necessary in an Alemannic system, because it would be unproblematic to have a light syllable. The fact that these children feel the need to lengthen the vowel indicates that they follow the Standard German weight restriction that requires stressed syllables to always be heavy. These children natively use a Standard German syllable system.
- 52.9 % show a syllable joint. In these instances, it is definitively clear that the children adopt the Standard German syllable structure, too.

As a preliminary result we can thus establish that there are indeed children who have acquired a Standard German syllable system (i.e. a phonological feature that originated in Low German) in a region that previously had a typologically quite distinct system.

2.3 Evidence: Apparent-time analysis

After establishing the existence of the nativisation of Standard German, we can concentrate on tracking its spread. In an apparent-time study (Pröll & Kleiner 2016; Pröll 2021a), realisations from 142 adults born and raised in the same East Swabian (Alemannic) region were gathered, this time capturing reading pronunciations of intended Standard use in the form of word lists. Four lexemes that vary with respect to the realisations of the vowels were chosen for analysis: *Dusche* 'shower [noun]', *Liter* 'litre', *rächen* 'to avenge' and *rösten* 'to roast', all of which may feature short (lax) or long (tense) vowels in their stressed syllables. Again, analyses show that two systems are competing with each other, and that the underlying issue is a suprasegmental one.

For *Dusche* and *Liter*, the typical Alemannic form is with a light (short) open syllable (circled in light grey in Figure 9.3, left). Under an apparent-time perspective, however, these realisations are predominantly found in older speakers (Figure 9.3, right), whereas younger speakers shift towards heavy (long) open syllables. As stressed above, the Alemannic system would allow both types, but the Standard German system (dark grey) only allows heavy stressed syllables.

The verbs *rächen* and *rösten* already have heavy long vowels in the autochthonous Alemannic system; however, the apparent-time analysis shows that these forms are being given up (Figure 9.4, right). Instead, people are moving towards a syllable structure that has a light short vowel, but still retains its weight, by making the following consonant ambisyllabic.

Thus, the pattern seen within the kindergarten data (in Section 2.2) surfaces again. If syllables need to be heavy, there are two ways to achieve this heaviness. Either (a) the vowel is lengthened or (b) a syllable joint is formed by means of an ambisyllabic consonant. The factors that decide which strategy is applied when are not yet clear.

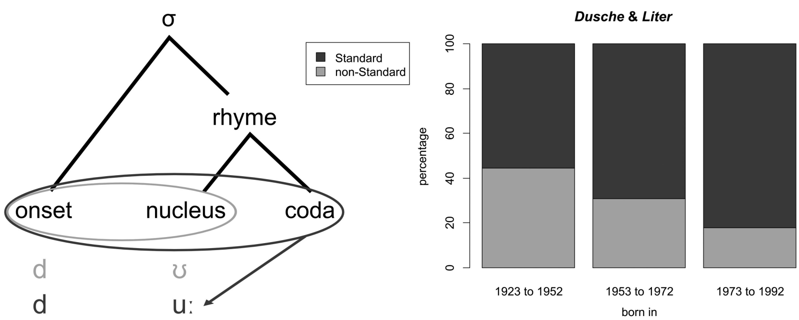


Figure 9.3 Change in syllable structure from light to heavy.

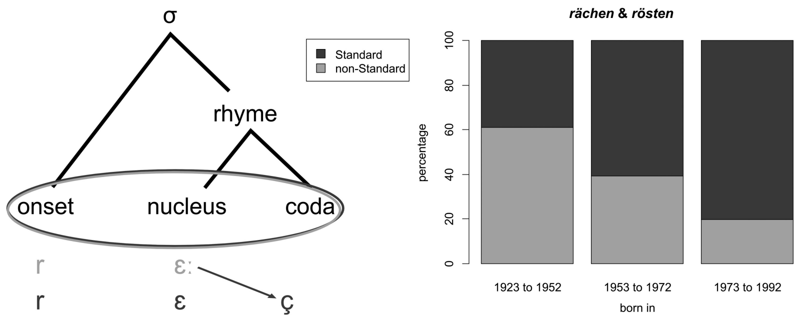


Figure 9.4 Change in syllable structure from heavy vowel to light vowel and syllable joint.

But whatever the underlying reason, this kind of measurement helps us to estimate the number of native[4] speakers of Standard German. Not all of the informants in the studies can be said to be native speakers of Standard German, but some are, and their number is increasing. The consequence is that Standard German has become a fully natural language, completing Weiß' (2005) stages outlined in Section 1. In fact, one could argue that this implies that the standardisation process of German has come to an end. Every change happening in the wake of nativisation would thus no longer be part of standardisation, but post-standardisation.

3. The morphological bifurcation of Standard German

As a next step, let us focus on the structural consequences that this process of nativisation has on the grammar of Standard German, or rather the grammars of both emerging and diverging varieties of Standard German. For this, we will take a closer look at the system of the indefinite article.

4 This might include the problematic category 'native-like', as there do not seem to be practical differences or consequences to distinguish between them, cf. Pröll (2021b).

3.1 *Background: Standard German indefinite article system (written paradigm)*

The system of the Standard German indefinite article is consistently – and without variation or doubt – codified and documented for the written form in the way it is displayed in Table 9.2 (see e.g. Zifonun et al. 1997: 34–5; Helbig & Buscha 2001: 324).

Nominative and accusative in neuter and feminine are syncretic, dative and genitive forms in the feminine as well. The feminine forms are relatively distinct from the other genders, but show advanced syncretism, whereas masculine and neuter are almost identical: They only differ from each other in one specific form, the accusative masculine *einen*. This form shall be the main focus of our following analyses.

3.2 *Evidence: TV news data*

Table 9.2 Written Standard German paradigm of *ein*.

<table>
<tr><th></th><th>masculine</th><th>neuter</th><th>feminine</th></tr>
<tr><td>nominative</td><td>ein</td><td rowspan="2">ein</td><td rowspan="2">eine</td></tr>
<tr><td>accusative</td><td>einen</td></tr>
<tr><td>dative</td><td>einem</td><td>einem</td><td rowspan="2">einer</td></tr>
<tr><td>genitive</td><td>eines</td><td>eines</td></tr>
</table>

To gather data for analysing the indefinite article system, professional newscasters on nationwide German television programmes were recorded. The point of using professional speakers is to minimise doubt that the forms obtained are in fact Standard German, as the threshold of 'standardness' remains debatable (cf. Pröll 2021b) – professional newscasters should satisfy more of the possible criteria than any other group of speakers.

The corpus consists of approximately seventy hours of recordings of fourteen news programmes on ten different TV stations. This yields a total of 1,006 instances where the written standard form would be masculine accusative *einen*. 625 of them (62.13 %) have a second syllable, either a schwa [ə] or a nasal with two intensity peaks or an increased relative length, thus they can be understood as a spoken representation of the written Standard German system seen above. In the other 381 instances (37.87 %), however, the spoken form is realised monosyllabic, most of the time as [ʔaɪn]. This results in the system shown in Table 9.3, where there is no distinct form for the accusative masculine anymore.

Table 9.3 Syncretic spoken Standard German paradigm of *ein*.

<table>
<tr><th></th><th>masculine</th><th>neuter</th><th>feminine</th></tr>
<tr><td>nominative</td><td>ʔaɪn</td><td rowspan="2">ʔaɪn</td><td rowspan="2">ʔaɪnə</td></tr>
<tr><td>accusative</td><td>ʔaɪn</td></tr>
<tr><td>dative</td><td>ʔaɪnəm</td><td>ʔaɪnəm</td><td rowspan="2">ʔaɪnɐ</td></tr>
<tr><td>genitive</td><td>ʔaɪnəs</td><td>ʔaɪnəs</td></tr>
</table>

But this single formal change has rather drastic consequences for the functional side of the paradigm as a whole: First of all, there is no differentiation between nominative and accusative left in any gender, but secondly, there is no formal difference between masculine and neuter left either. Thus, it is much more appropriate to assume the underlying paradigm as in Table 9.4.

So, while the written standard variety shows a system with four cases and three distinct genera, the spoken standard variety has speakers who

Table 9.4 Syncretic spoken Standard German paradigm of *ein*.

<table>
<tr><th></th><th>masc = neut</th><th>feminine</th></tr>
<tr><td>nom = acc</td><td>ʔaɪn</td><td>ʔaɪnə</td></tr>
<tr><td>dative</td><td>ʔaɪnəm</td><td rowspan="2">ʔaɪnɐ</td></tr>
<tr><td>genitive</td><td>ʔaɪnəs</td></tr>
</table>

conflate the nominative and accusative case as well as the distinction between masculine and neuter.

3.3 Phonologically or morphologically driven change?

Of course, such syncretisms are not unheard of. We do find similar scenarios in standardised languages, especially the ones with a deeper orthographic system (i.e. languages where the correspondence between phonemes and graphemes has deviated from being one-to-one), such as the syncretisms in the verbal system of Standard French that are more advanced (and more regular) in the spoken than in the written form (see Table 9.5). Often this deeper system is due to relatively early codification, simply meaning that the orthography represents a phonological system long since overwritten by language change.[5]

5 One notable counterexample would be Faroese, where the rift between written and spoken form (we should be careful to call spoken Faroese standardised) is rather large, although the orthographic system is relatively young – the reason for this is that the modern standardised orthography has been modelled to be archaic by choice, in an attempt to (a) show the etymological links to its neighbour Icelandic and (b) capture the wealth of variation in spoken Faroese (cf. Hagström 2005). A historicising orthography has the advantage of not favouring any recent variety over the other; in a way, it is equally distant to all modern varieties it represents; thus it does not favour any single one over the others (cf. Braunmüller 1999: 267).

Table 9.5 Syncretisms in two written vs. spoken French verb classes.

		-er verbs		*-ir* verbs	
		written	spoken	written	spoken
singular	1st	*parle*	[paʁl]	*choisis*	[ʃwazi]
	2nd	*parles*	[paʁl]	*choisis*	[ʃwazi]
	3rd	*parle*	[paʁl]	*choisit*	[ʃwazi]
plural	1st	*parlons*	[paʁlɔ̃]	*choisissons*	[ʃwazisɔ̃]
	2nd	*parlez*	[paʁle]	*choisissez*	[ʃwazise]
	3rd	*parlent*	[paʁl]	*choisissent*	[ʃwazis]

At first glance, we might now assume that these changes actually support the overly simple model we criticised earlier on. Unstressed written *-en* and *-em* have a full orthoepic realisation of [ən] and [əm], but the [ə] is regularly omitted (see Hahn & Siebenhaar 2019), leaving just a syllabic nasal [n̩] or [m̩] – and this nasal might get assimilated. An explanation that is purely based on articulatory reduction seems obvious here, meaning that the spoken standard in fact was nothing more than the written standard plus phonology and phonetics.

A closer look at the morphological systems of German non-standard varieties can shed further light on this. Spoken colloquial German (north of the so-called 'Mainlinie', cf. Durrell 1989 and the corresponding maps[6] of the *Atlas zur deutschen Alltagssprache*, Elspaß & Möller 2003) has a reduced system with no formal distinction between masculine and neuter nor nominative and accusative, as can be seen in Table 9.6. (There is no genitive, as it practically does not exist in the reduced paradigm.)

In written colloquial German (e.g. online chats), there are signs for the emergence of a system with semi-reduced forms (see e.g. Burri 2003; Vogel 2006). These forms retain a vocalic element (see Table 9.7).

This system is highly stigmatised (see e.g. Ziegler 2012: 310–11) but salient among laypeople. The issue is that the form *nen* originates as a

6 Online: <https://www.atlas-alltagssprache.de/runde-5/f18a-b/> (accessed 30.11.2024).

Table 9.6 Paradigm of reduced (spoken) colloquial forms.

	masc = neut	feminine
nom = acc	*n*	*ne*
dative	*m*	*ner*

Table 9.7 Paradigm of semi-reduced (written) colloquial forms.

	masc = neut	feminine
nom = acc	*nen*	*ne*
dative	*nem*	*ner*

shortened version of the very distinct accusative masculine form *einen*. Thus, users who are unfamiliar with this kind of usage tend to parse *nen* as an accusative masculine, even in nominative masculine or masculine / accusative neuter positions. However, *nen* is not to be understood as the short form of accusative masculine *einen*, but a more explicit written form of the syncretic spoken *n* that we saw earlier in Table 9.6 (with nom=acc masc=neut). There seems to be a lower threshold of what constitutes a well-formed lexeme in written colloquial German, and *n* does not seem to have enough graphematic substance for the job.

Despite different forms and apparently different trajectories of analogical expansion – in most systems, the nominative form expands, while in the last part it was the accusative –, all these changes result in the same formal system (see Table 9.8), and this even happens in written German, where there cannot possibly be a phonological process in action. Three times the identical system, albeit in different varieties or registers, across different modalities and 'repeated across multiple exponents' (Baerman 2008: 221) strongly suggests that the pattern is neither coincidental nor purely phonological in nature.

Table 9.8 Abstraction: German paradigm of *ein*, reduced (spoken Standard, spoken and written non-Standard).

	masculine	neuter	feminine
nominative	A		B
accusative			
dative	C		D
(genitive)	(E)		(D)

If we take these findings into consideration, it seems that there is a morphomic[7] difference between the written and the spoken standard, with the spoken standard following a system that is present in other varieties of German as well. Remember that the Standard speakers are trained professionals; although the consequences for the morphological system are grave, laypeople do not seem to notice at all. The written non-standard systems, however, are indeed noticed and highly stigmatised, despite being morphologically identical to the speech of newscasters. Now, obviously, the emergence of this syncretic spoken system can be interpreted as evidence for the nativisation of Standard German.

4. Consequences and prognoses

This paper argued that there are two kinds of Standard German that differ in more respects than just their modality, and also discussed the likely reasons for this: First, we could see that Standard German is at a stage of its development where it is evolving into a fully natural language, meaning that it is subject to undirected first language acquisition processes with all of its social and psychological implications. Second, we got

7 The term 'morphomic' reflects the notion that a morphological phenomenon cannot be explained solely through processes that take place outside of the scope of morphology, see Enger (2019).

a first glance at the concrete consequences of this, namely morphological differences that are of a morphomic kind. Spoken Standard German is more than an aberration of an 'original' written standard, and it cannot be derived exhaustively from the written form via cognitive processing, performance phenomena or phonetic realisation on parts of the speakers. It is evolving its own grammar that cannot be explained just by differences in mediality.

The written standard – not that surprisingly – conserves (or rather freezes) an older state of morphological complexity in German that has been overwritten in other varieties, including the spoken standard. In languages with a deep orthographic system (such as English, for example), this process has of course been fruitful for historical phonology, but the findings for morphological processes in German have so far been extremely sparse. A rare exception to this is a current study on semantic change in the tense systems of German varieties (Seiler & Weber 2022), which can show how the change in the German tense systems is progressing – with the exception of written Standard German, where the Early New High German status is frozen by codification.

The apparent-time analysis shows that spoken Standard German is spreading in southern Germany. Again, this is not surprising at all, but it is important that it actually is the percentage of fully native speakers that is increasing. This fits in with more recent attempts to periodise German-language history, such as by Elspaß (2008) (cf. Figure 9.5), in which a relevant turning point for the language history of German is proposed for the middle of the twentieth century; of course, this can be motivated by the large socio-historical changes following the end of the Second World War, including the resettlement of a large part of the German-speaking population in Europe. But this also corresponds to the time frame during which nativisation gains momentum, as German seems to reach a state

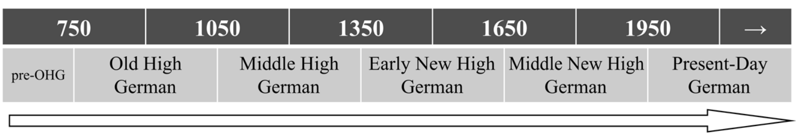


Figure 9.5 Periodisation of German (following Elspaß 2008).

where a critical number of speakers grows up with Standard (or near-Standard) German as an input for L1-acquisition.[8] I would argue for considering this as a crucial factor in both the structural and socio-historical transition to present (or Present-Day) German as well.

From a typological point of view, we can see that while the modern written standard variety clearly differs from most of its donor varieties, the spoken variety does so to a lesser extent. The habitual trajectory of change processes in the Germanic languages was put on hold for Standard German due to its codification and thus its conservation. Now, however, nativisation allows for the structural (re-)integration of the spoken standard variety into the large group of naturally spoken varieties of German. This also implies that documenting and understanding the structural consequences of nativisation is relevant because the current structure of the standard – or rather the two different faces of the standard – could not be explained without understanding the process behind its bifurcation.

Bibliography

Auer, Peter. 1989. Zur Dehnung im Alemannischen (am Beispiel des Konstanzer Stadtdialekts). *Zeitschrift für Dialektologie und Linguistik* 61(1), 32–57.

Auer, Peter. 2005. Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In Nicole Delbecque, Johan van der Auwera & Dirk Geeraerts (eds). *Perspectives on Variation*. Berlin: Mouton de Gruyter. 7–42.

Baerman, Matthew. 2008. Case syncretism. In Andrej L. Malchukov & Andrew Spencer (eds). *The Oxford Handbook of Case*. Oxford: Oxford University Press. 219–30.

Burri, Gabriela. 2003. Spontanschreibung im Chat. *Linguistik online* 15(3), 3–31.

8 It could be argued that neither the principle nor the time frame is limited to German: The sociolinguistic development of Danish (as sketched by Pedersen 2003), for example, lends itself to an analogous interpretation.

Caro Reina, Javier. 2019. Wortsprachliche Merkmale im Alemannischen. *Linguistik Online* 98(5), 235–71.

Chomsky, Noam. 1995. Language and nature. *Mind* 104(413), 1–61.

Durrell, Martin. 1989. Die 'Mainlinie' als sprachliche Grenze. In Wolfgang Putschke, Werner Veith & Peter Wiesinger (eds). *Dialektgeographie und Dialektologie. Günter Bellmann zum 60. Geburtstag von seinen Schülern und Freunden*. Marburg: Elwert. 89–109.

Durrell, Martin. 1999. Standardsprache in England und Deutschland. *Zeitschrift für Germanistische Linguistik* 27(3), 285–308.

Enger, Hans-Olav. 2019. Thoughts on morphomes, on a Scandinavian background. In Antje Dammel & Oliver Schallert (eds). *Morphological Variation. Theoretical and Empirical Perspectives*. Amsterdam: Benjamins. 159–96.

Elspaß, Stephan. 2008. Vom Mittelneuhochdeutschen (bis ca. 1950) zum Gegenwartsdeutsch. *Zeitschrift für Dialektologie und Linguistik* 75(1), 1–20.

Elspaß, Stephan & Robert Möller. 2003–. *Atlas zur deutschen Alltagssprache* (*AdA*). Online <http://www.atlas-alltagssprache.de> accessed 8 June 2024.

Hagström, Björn. 2005. Written language and forms of speech in Faroese in the 20th century. In Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karker, Hans-Peter Naumann & Ulf Teleman (eds), *The Nordic Languages. An International Handbook of the History of the North Germanic Languages*. Vol. 2. Berlin, New York: de Gruyter. 1750–58.

Hahn, Matthias & Beat Siebenhaar. 2019. Schwa unbreakable – Reduktion von Schwa im Gebrauchsstandard und die Sonderposition des ostoberdeutschen Sprachraums. In Sebastian Kürschner, Mechthild Habermann & Peter O. Müller (eds). *Methodik moderner Dialektforschung: Erhebung, Aufbereitung und Auswertung von Daten am Beispiel des Oberdeutschen*. Hildesheim: Olms. 215–36.

Haugen, Einar. 1966. Dialect, language, nation. *American Anthropologist* 68, 922–35.

Hayes, Bruce. 1989. Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry* 20(2), 253–306.

Helbig, Gerhard & Joachim Buscha. 2001. *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht*. Berlin, München, Wien, Zürich & New York: Langenscheidt.

Hulst, Harry van der. 1985. Ambisyllabicity in Dutch. In Hans Bennis & Frits Beukema (eds). *Linguistics in the Netherlands 1985*. Dordrecht: Floris. 57–66.

Inoue, Miyako. 2006. Standardization. In Edward Keith Brown, Ronald E. Asher & James M. Y. Simpson (eds). *Encyclopedia of Language & Linguistics. Second edition*. Amsterdam: Elsevier. 121–27.

Kleiner, Stefan & Ralf Knöbl. 2023. *Duden – Das Aussprachewörterbuch*. 8., completely revised edition. Berlin: Dudenverlag.

König, Werner. 2004. Die Aussprache des Standarddeutschen als Sprachkontaktphänomen. In Horst Haider Munske (ed.). *Deutsch im Kontakt mit germanischen Sprachen*. Tübingen: Niemeyer. 175–201.

Paul, Hermann. 1920/1995. *Prinzipien der Sprachgeschichte*. Tübingen: Niemeyer.

Pedersen, Inge Lise. 2003. Traditional dialects of Danish and the de-dialectization 1900–2000. *International Journal of the Sociology of Language* 159, 9–29.

Polenz, Peter von. 1994. *Deutsche Sprachgeschichte vom Spätmittelalter bis zur Gegenwart*. Vol. II. *17. und 18. Jahrhundert*. Berlin, New York: de Gruyter.

Pröll, Simon. 2021a. Die Nativierung des Standarddeutschen. Mikrotypologische Evidenz für suprasegmentalen Wandel. *Zeitschrift für Angewandte Linguistik* 75(1), 305–29.

Pröll, Simon. 2021b. How many natives with how many systems? Intra-individual variation and the threshold of multilingualism in Standard German speakers. In Alexander Werth, Lars Bülow, Markus Schiegg & Simone Pfenninger (eds). *Intra-individual Variation in Language*. Berlin, New York: de Gruyter Mouton. 379–405.

Pröll, Simon, Jan-Claas Freienstein & Oliver Ernst. 2016. Exemplarbasierte Annäherungen an das Silbengelenk. *Zeitschrift für germanistische Linguistik* 44(2), 149–71.

Pröll, Simon & Stefan Kleiner. 2016. Silbengrenzen im Gebrauchsstandard. Empirie, Theorie und Typologie. *Deutsche Sprache* 44(3), 193–213.

Pröll, Simon & Thilo Weber. 2023. Afrikaans in a quantitative typology of Germanic standard and non-standard varieties. In Adri Breed (ed.). *Proceedings of the 4th International Afrikaans Grammar Workshop*. Potchefstroom: North-West University. 39–68.

Ramers, Karl Heinz. 1992. Ambisilbische Konsonanten im Deutschen. In Peter Eisenberg, Karl Heinz Ramers & Heinz Vater (eds). *Silbenphonologie des Deutschen*. Tübingen: Narr. 246–83.

Røyneland, Unn. 2010. Vertical convergence of linguistic varieties in a language space. In Peter Auer & Jürgen Erich Schmidt. (eds). *Language and Space*. Vol. 1: *Theories and Methods*. Berlin, New York: de Gruyter. 259–74.

Salmons, Joseph. 2012. *A History of German*. Oxford: Oxford University Press.

Schmidt, Oskar & Theo Vennemann. 1985. Die niederdeutschen Grundlagen des neuhochdeutschen Lautsystems. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 107, 1–20, 157–73.

Seiler, Guido & Thilo Weber. 2022. Reichenbach meets underspecification. A novel approach to the perfect-past-cycle in German (and elsewhere). *Journal of Historical Linguistics* 12(1), 108–66.

Stellmacher, Dieter. 1981. *Niederdeutsch. Formen und Forschungen*. Tübingen: Niemeyer.

Vandenbussche, Wim. 2007. Shared Standardization Factors in the History of Sixteen Germanic Languages. In Christian Fandrych (ed.). *Standard, Variation und Sprachwandel in germanischen Sprachen*. Tübingen: Narr. 25–36.

Vogel, Petra M. 2006. 'Ich hab da nen kleines Problem!' Zur neuen Kurzform des indefiniten Artikels im Deutschen. *Zeitschrift für Dialektologie und Linguistik* 73(2), 176–93.

Weiß, Helmut. 2001. On two types of natural languages. Some consequences for linguistics. *Theoretical Linguistics* 27(1), 87–103.

Weiß, Helmut. 2004. A question of relevance. Some remarks on standard languages. *Studies in Language* 28(3), 648–74.

Weiß, Helmut. 2005. Von den vier Lebensaltern einer Standardsprache. Zur Rolle von Spracherwerb und Medialität. *Deutsche Sprache* 33(4), 289–307.

Ziegler, Evelyn. 2012. *Merkel hat 'nen Neuen!* Die Kurzform *nen* – ein Mündlichkeitsmarker par excellence. In Britt-Marie Schuster & Doris Tophinke (eds). *Andersschreiben. Formen, Funktionen, Traditionen*. Berlin: Erich Schmidt. 295–316.

Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker, 1997. *Grammatik der deutschen Sprache*. Berlin, New York: de Gruyter.

PART II

Standards in pluricentric settings

FELIX TACKE

10 A corpus-driven approach to emerging national standard varieties in Spanish: The *Corpus Radiofónico Panhispánico* (CO.RA.PAN)

ABSTRACT

Spanish is considered one of the most advanced among Romance languages in terms of pluricentric development, and there is no doubt that regional standard languages have been emerging in Latin America for some time now. While the pluricentric nature of Spanish has been acknowledged both by the research community and relevant actors in the field of codification, there is still a lack of systematic empirical studies. To address this gap, the upcoming *Corpus Radiofónico Panhispánico* (CO.RA.PAN) aims to gather and make available authentic oral speech data produced by professional speakers of major radio broadcasters in Spanish-speaking countries. Unlike traditional dialectological and general sociolinguistic research, the project focuses on formal speech situations reflecting standard and near-standard speech. This contribution will provide a general presentation of the CO.RA.PAN project, including its overall conception, technical goals, and limitations.

1. Introduction[1]

There is no doubt that the Spanish language has made significant progress in terms of pluricentric development, as recognised by both the research community and relevant actors in codification. However, compared to other pluricentric languages such as English, there is still little consensus on the exact nature of Spanish pluricentricity. While there is a general consensus that 'múltiples normas ejemplares americanas' ('multiple exemplary American norms'; Stepanov 1971: 1167) exist alongside the traditional normative centre in Madrid, the question of the precise number and configuration of American usage norms that have emerged through bottom-up standardisation processes remains largely unresolved due to a gap of systematic and comparable empirical data. This paper describes the upcoming *Corpus Radiofónico Panhispánico* (CO.RA.PAN), a corpus project designed to address this gap by gathering data on the emergence of national usage norms in Spanish-speaking countries. Additionally, the CO.RA.PAN project seeks to gain a deeper understanding of double-standard situations that arise due to the competition between newly emerging endogenous norms and the traditional exogenous standard language. To accomplish this, the corpus will comprise authentic speech data

1 This paper is based on a presentation held in 2022 and was written in early 2023. It outlines the conceptual design of the CO.RA.PAN project as originally planned. Since then, the project has been resumed at Marburg University (2023) and fully implemented using AI-based automatic transcription. As of July 2025, the corpus comprises 1.5 million transcribed words, which were subsequently reviewed and manually classified by speaker type and communicative setting. The data are encoded in a transparent, well-documented JSON format and are accessible through a sophisticated web app at https://corapan.online.uni-marburg.de/. For technical documentation, including versioning information and regularly updated specifications, see https://hispanistica.online.uni-marburg.de/. We thank Carolin Grote and Yannick Bächler for their work on transcription review and classification. An earlier pilot study at Bonn University (2021–2022) informed the initial design; thanks are due to Ana Goás Pérez, Melanie Martín Rincón, and Hanna-Marie Senne, and especially to Lisa Thomas for technical support. Due to editorial constraints, the technical updates and implementation details could not be reflected in the final version of this paper.

gathered from the major radio broadcasters in each Spanish-speaking country. Through this method, we aim to gain insight into the *national oralisation norms* of each country.[2]

Studying the emergence of new American standard languages within Spanish involves focusing on what Stewart (1968 [1962]: 534) referred to as *informal standardisation*, in contrast to codification processes, i.e. explicit or *formal standardisation* (cf. Amorós-Negre 2014: 208). Rather than a top-down imposition of norms by institutional actors, informal standardisation involves a bottom-up emergence of norms through usage (cf. Lebsanft & Tacke 2020: 14).[3] In recent years, research on standard languages and linguistic normativity has become increasingly sophisticated and developed a complex methodology. To get a complete picture of standardisation, which encompasses the nature of informal standard languages and their status in their respective communities, it is crucial to consider different viewpoints. While informal standard languages may provide the groundwork for codification, they are not yet codified. As such, describing them requires first studying the *descriptive norms of production* to determine how a language is currently used by the norm-setting elite in a specific community. However, recent research has shown that the 'objective' description of usage norms does not always align with speech communities' attitudes and beliefs about how their language ought to be used. As a result, the status of an informal standard language within a community and the level of identification its speakers have with it can vary significantly from case to case (which is, at the same time, an indicator for pluricentric development as discussed in Bierbach 2000 and Muhr 2016). This is due to the fact that the 'old' prescriptive norm may still affect speakers' attitudes and perceptions. Therefore, it is important

2 Regarding the concept of 'national oralisation norms' in German dialectology and standardology, see, e.g. Schmidt (2004) and Schmidt & Herrgen (2011), as well as Auer (2017).

3 In this context, *bottom-up* should not be equated with the standardisation processes 'from below' as described by Elspaß (2021), which may involve the contribution of the lower strata of society to standardisation processes. Instead, *bottom-up* simply implies that specific norms are not explicitly imposed by an authority. However, it is possible that informal standardisation still relies predominantly on social elites.

to consider not only the objective study of usage norms, but also the representations, attitudes, and perceptions of the speakers. Tests that examine the speakers' perception of specific features of both 'old' prescriptive norms and emerging standard varieties, for example, can reveal significant discrepancies and contradictions.[4] In emerging centres of pluricentric languages, such inconsistencies are often indicators of the persistence of double-standard settings.[5] In this sense, Pustka et al. (2017: 113) are right to suggest that, '[m]éthodologiquement, la question de la norme ne peut être résolue que par une approche combinant étude de représentations, de production, et de perception' ('methodologically, the question of the norm can only be solved by an approach that combines the study of representations, productions and perceptions').[6] Given the complexity of the issue, it is important to acknowledge the limitations of the CO.RA.PAN project, as it cannot fully encompass all viewpoints. Nevertheless, this corpus-based project provides a valuable contribution, particularly in the study of the descriptive norms of oral language usage in Spanish-speaking countries, thereby laying the foundation for future research. In addition, the corpus data can also facilitate new approaches to perception tests. However, it is important to keep in mind that a comprehensive understanding of language usage norms and pluricentric development requires supplementary studies on attitudes and representations (see Section 4).

The paper is structured as follows: Section 2 will give a general overview of the study of pluricentricity within the field of Spanish linguistics and the problems that arise from them when it comes to empirical research on emerging standard languages. Section 3 will present the upcoming CO.RA.PAN project and describe its conception as well as its limitations by defining its basic tenets. Section 4 will discuss further research perspectives and possibilities emanating from a corpus of

4 For research on the perception of language varieties, including standard speech, cf. Krefeld & Pustka (2010) and Postlep (2010).

5 See, e.g. Chalier's (2021) study on French pronunciation norms (cf. also Tacke 2023).

6 All translations are my own unless otherwise noted.

authentic oral data that represents standard and near-standard speech. Finally, Section 5 offers some basic conclusions.

2. Researching the emergence of new standard languages in Spanish-speaking countries

The pluricentric nature of the Spanish language as it is spoken today in Spain and across the Americas is a subject of interest on various fronts. Firstly, pluricentricity is a verifiable reality that stems from the regional and national differences observed in language usage, particularly in written language and even more so in oral speech settings, where the use of a standard variety is to be expected (i.e. in formal speech by educated speakers). Secondly, pluricentricity has evolved into a political concept that entails recognising and embracing linguistic diversity – in terms of different varieties of Spanish – within the domain of language codification and modernisation. The *Real Academia Española*'s remarkable policy shift towards pluricentricity in the cultivation of the standard language since the early 1990s has led to numerous studies and significant insights into the incorporation of factual pluricentricity into the official codification instruments of Spanish language academies (see Lebsanft 2004, 2007; Amorós-Negre 2012; Greußlich 2015: 66–74; Pöll 2021): The implementation of the slogan *unidad en la diversidad* ('unity in diversity') has brought about a new perspective and reflects a new mindset that embraces variation in the Spanish language. As a result, the *Real Academia Española* has attempted to codify a *norma panhispánica* ('Panhispanic norm'), meant to be the sum of all prestigious variants in the various *hablas cultas*, while also transitioning towards a more descriptive approach to codification. This approach dedicates significant space to the description of national, regional and local variants within its most important reference works, including the *Nueva gramática de la lengua española* (NGLE; cf. Tacke 2011; Greußlich 2015: 74–83). Therefore, it may appear that the subject of Spanish pluricentricity has been extensively researched and is backed by a shared understanding among both the scientific community and the political institutions responsible for cultivating the standard language. However, upon closer examination of our understanding of the geographical variation of the Spanish standard language, including the features of individual national standard varieties and their overall structure, it becomes evident that our knowledge of factual pluricentricity lags behind the study of political pluricentricity and its implications for the continued codification and modernisation efforts of language academies. Hence, the question of how to approach the empirical study of Spanish pluricentricity is key. In this context, Prifti (2021: 189) rightly points out that '[l]a investigación de las dinámicas lingüísticas pluricéntricas debe basarse en un sólido fundamento empírico. La calidad de la elicitación de datos empíricos resulta particularmente relevante y requiere, por eso, una atención especial' ('The study of pluricentric dynamics of language must be based on a solid empirical foundation. The quality of the elicited empirical data is particularly relevant and therefore requires special attention'). However, despite an increasing number of empirical studies that aim to describe pluricentric tendencies in Spanish, particularly through analyses of its usage in daily newspapers and audiovisual mass media (cf. the contributions in Lebsanft et al. 2012 and Greußlich & Lebsanft 2020a), we still lack the means to study the use of national varieties of Spanish in a systematic and comparable fashion.[7]

There has been, however, some debate on the best ways to investigate the development of standard varieties of Spanish. Particularly noteworthy in this regard are the contributions of Wulf Oesterreicher (2000) and Mechtild Bierbach (2000) at the turn of the millennium. While Bierbach's focus is on a comprehensive understanding of pluricentricity from both linguistic and cultural perspectives, Oesterreicher is primarily

7 Unfortunately, an ambitious and promising effort to provide high-quality data in this field, the *Proyecto de estudio del habla culta de las principales ciudades de Hispanoamérica* (1964–), in which teams of sociolinguists collected speech data from educated speakers of all major Spanish-speaking cities in Latin America, appears to have been completely stalled for a long time, and the collected data remains inaccessible to the wider scientific community (cf. Bierbach 2000: 167–8). For information on other projects exploring variation among educated speakers in Spanish-speaking countries, see the overview in Greußlich (2015: 87–90).

concerned with providing a detailed description of linguistic phenomena and their variation, particularly their diasystematic status across the Americas. He notes that the configuration of Spanish pluricentricity is more complex than that of English or Portuguese due to the fact that standard variants in Spanish tend to cross national boundaries instead of being limited by them. Hence, Oesterreicher argues that it is crucial to gather precise data on each linguistic variant and the geographical extent of its 'standardness', as some variants may be considered standard in one country but substandard in another and vice versa. He rejects the notion of national standard varieties, as the features of these varieties may overlap across neighbouring countries. Instead, he proposes the concept of supranational regional standard languages, which he calls 'gestaffelte[n] Teilstandards mit einem jeweils hohen Überschneidungsanteil' ('staggered partial standard languages with a high proportion of overlapping'; Oesterreicher 2000: 308). Bierbach (2000), on the other hand, takes a broader perspective by exploring the cultural implications of the concept of pluricentricity and the role of the nation-state as a framework and a point of reference for its development. In other words, Bierbach is interested not only in the dynamics of language use in speech production but also in the dynamics of how it is perceived by national communities, while her perspective also acknowledges the potential of emerging national standard varieties to shape the linguistic identity of their speakers. Consequently, Bierbach (2000: 167–8) suggests an alternative approach to studying Spanish pluricentricity. Instead of focusing on preselected linguistic features and their variants, she advocates for a country-by-country approach – in line with Canfield (1981) and Lipski (1994) – that takes the nation-state as a starting point and describes each country's national variety as a 'subsystem' of Spanish.

It is noteworthy that despite the divergent research interests of both scholars, Oesterreicher (2000: 308) acknowledges that the country-by-country approach is the only method capable of resolving the methodological challenges posed by his focus on describing phenomena on a case-by-case basis and identifying higher-level regional standard languages: 'Als methodisches Prinzip hat zu gelten: Man muß durchaus von kleinräumig-nationalen Gebieten *ausgehen* und den in diesen jeweils gültigen Standard beschreiben' ('The following methodological principle should be respected: It is necessary to take the small-scale national space as a *starting point* and to describe the respective standard language from there'). From our point of view, the country-by-country approach is not only a necessary methodological concession but also an aim in and of itself when it comes to understanding the dynamics of standard language usage in the Americas. Only the national framework offers the possibility to gather reliable data on informal standardisation processes and the emergence of double-standard situations.[8]

The type of empirical data that is most suitable for analysing the emergence of national standard varieties is another crucial aspect to consider. Pluricentric development is primarily a process closely tied to the novel communicative conditions that have arisen in Late Modernity and are typical of it. Spanish language culture, like other Romance language cultures such as Italian and Catalan, has undergone a paradigm shift in recent decades. This shift has brought about a new understanding and configuration of *standardness*, including the conception of standard languages and standard language usage (cf. Tacke 2024). While standardness traditionally pertains to formal and ideally written speech, the current understanding of language usage is no longer limited to literary language. Today, the concept of standardness encompasses oral speech situations as well and caters to the needs of the more complex communicative settings of Late Modernity. Here, linguistic means are evaluated not only based on their correctness but also on their appropriateness in different speech settings. This reconfiguration of the notion of standard is increasingly being referred to as *destandardisation* and *restandardisation* in the context of Spanish (cf. Amorós-Negre 2014, 2020; Greußlich & Lebsanft 2020b; Tacke 2024). Restandardisation as a notion is conceived of as a possible result of destandardisation (cf. Mattheier 1997), which refers to the phenomenon of a codified (written) standard language losing its

8 This does not imply that the study of supranational regional standards and their overlapping features is invalid; rather, it suggests that such a description is of secondary importance and mainly pertains to the investigation of Modern Spanish as 'a totality of subsystems' (Stepanov & Švejcer 1981: 223; cf. Bierbach 2000: 167).

relevance for communicative purposes in late modern societies, 'whereby the established standard language loses its position as the one and only "best language"' (Ayres-Bennett 2021: 49; cf. also Lebsanft & Tacke 2020: 21–2 and Kristiansen 2021). Restandardisation then refers to the emergence of new standard languages, or *neostandards*, that are better suited for modern communicative purposes, which typically include oral usage in speech settings of varying degrees of formality.[9] As a result, double-standard situations arise. Since restandardisation generally emerges from spoken language, it usually 'consists in the formation of regional standards with minor differences between one another, each containing features which were formerly considered to be substandard and censured as "bad usage", "only spoken", "to be avoided", "incorrect", etc'., as Berruto (2017: 37) states regarding Italian. Concerning Spanish, the notion of destandardisation describes the ever-decreasing relevance of the traditional European standard language in everyday (formal) communication in American countries.[10] In contrast, restandardisation describes the emergence of regional and national models of good usage, which usually originate from oral communication and result in new American neostandard languages. In American countries, both processes go hand in hand and translate into a 'de-europeanisation' (Bell 2011: 179) of the language. Against this backdrop, Greußlich (2015: 85–6) is absolutely right when he points out pluricentric development in Spanish is best studied based on the analysis of spoken language rather than written communication: 'es lícita la hipótesis de que las características más relevantes en el marco de la normatividad pluricéntrica del español deben pertenecer al ámbito de lo hablado' ('it is legitimate to assume that the most relevant characteristics in the domain of the pluricentric normativity of Spanish

9 For a comprehensive discussion of the notions of destandardisation and restandardisation, see Tacke (2024).

10 Schiffman (1998: 363; quoted according to Kristiansen 2021: 679) argues that a double-standard situation resulting from restandardisation may endure for an extended period. He points out that 'restandardization will never totally replace the older standard language; the older norm will simply be elevated to a "classical" status that it will continue to inhabit' (cf. already Joseph 1987: 173–4).

would pertain to the realm of spoken language'). Based on these considerations, the most effective approach to analysing the processes of informal (re)standardisation is to build a corpus consisting of oral speech data gathered from standard or near-standard speech settings, rather than relying solely on written corpora such as daily newspapers. The data will show, for each country, what is being broadcast as exemplary Spanish. However, whether the news anchors' Spanish is in fact representative of the respective country's *norma culta* is another question. In the next section, we will discuss how the CO.RA.PAN project plans to address these issues.

3. The CO.RA.PAN project: Conception and limitations

The CO.RA.PAN project employs a country-by-country analysis as both a necessary methodological starting point and a conceptual premise, providing a framework for deeper understanding of the dynamics of Spanish as a pluricentric language. The corpus will exclusively comprise authentic oral speech data gathered from the major radio broadcaster(s) in each Spanish-speaking country. The goal is to gather approximately equal amounts of speech data from each country. Additionally, the corpus will include various types of radio programmes, classified according to their communicative setting, to better capture the current processes of informal standardisation and the use of linguistic features in standard and near-standard speech situations. The overall goal of this project is to compile a corpus of standard and near-standard speech from major radio broadcasters in Spanish-speaking countries, which will be searchable and amenable to quantitative analysis.[11] Accordingly, the project will

11 A similar approach to the study of usage norms has been applied to German as part of the Regionalsprache.de or REDE project (see Ganswindt et al. 2015; for information on other dialectological and standardological projects, cf. the contributions in Kehrein et al. 2015). However, the data used in the REDE project

concentrate on language use that reflects 'descriptive norms of production' in the sense of 'national oralisation norms' (cf. Section 1), which are most apparent in the speech behaviour of professional speakers. As we will argue below, this particular social group is believed to have an unrivalled status and influence when it comes to norm-setting language use. In the subsequent sections, we will examine these basic tenets in greater depth.

3.1 The model speaker approach: Radio presenters

The first problem that arises when it comes to the study of descriptive norms of production is the question of whose speech represents best the so-called 'good use' (Sp. *buen uso* < Fr. *bon usage*; cf. Lebsanft & Tacke 2020: 31–2) of specific language communities. Hence, a key challenge is determining who is the 'best speaker'? Broadly speaking, the traditional approach regards the regularities observed in the speech or writing of educated speakers as representative of 'good use' and constitutive of the corresponding standard language. This age-old assumption, already formulated by Quintilian, is based on the principles of *consuetudo* and *consensus eruditorum*, which refer to the consensus among educated speakers on what constitutes the good use of language. In the Spanish tradition, these principles are reflected in the notions of *hablantes cultos* ('educated speakers'), whose speech regularities constitute the so-called *habla culta* ('educated speech variety'), from which the *norma culta* ('learned norm') can be derived (cf. Lebsanft & Tacke 2020: 37–40). The underlying reasoning behind these assumptions is twofold: first, as a social elite, educated speakers are often seen as role models for less educated individuals; second, their speech, particularly in formal settings, exhibits less variation compared to other social groups, as it aligns more closely with the traditional

does not rely on broadcasts; instead, it is derived from a variety of sources such as interviews and other forms of authentic speech data.

literary standard language. Historically, the group with the most impact on the overall society in linguistic terms has been the men and women at and around the royal court (e.g. in the UK, in France and in Spain). This is, of course, no longer the case. However, in Late Modernity, a new social elite has emerged due to compulsory schooling, consisting of the middle and higher middle socioeconomic class. In this sense, Chalier (2021: 16), who studies pronunciation norms in French communities, affirms that '[l]a notion d'élites sociales porteuses de la norme est restée cependant bien présente dans la conscience des locuteurs jusqu'à aujourd'hui, aussi bien dans la communauté francophone que dans d'autres communautés linguistiques' ('the notion of social elites that carry the norm is, however, still present in the speakers' consciousness until today, not only in the Francophone community but also in other speech communities'). Accordingly, the Spanish language academies, in 2005, explained that the Spanish standard language is largely based on the speech of educated speakers in formal settings: it is 'esta expresión culta formal la que constituye el español estándar: la lengua que todos empleamos o aspiramos a emplear, cuando sentimos la necesidad de expresarnos con corrección' (DPD 2005, XIV) ('this learned and formal speech that constitutes the Spanish standard language: the language we all use or aspire to use when we feel the need to express ourselves correctly').

This 'Spanish standard language', based on the speech of educated speakers in formal situations, might be quite homogeneous and is often evoked to symbolise the unity of the Spanish language across countries and continents, which is a political goal. Nevertheless, the Spanish language academies acknowledge that educated speech varieties, despite being based on a common literary language, are only similar and not identical: 'Por su carácter de lengua supranacional, hablada en más de veinte países, el español constituye, en realidad, un conjunto de normas diversas' (DPD 2005, XIV) ('Due to being a supranational language spoken in more than twenty countries, the Spanish language actually constitutes an assembly of diverse norms'). Hence, even the speech of educated speakers in formal situations displays 'variaciones mínimas entre las diferentes zonas' (ibid.) ('minimal variations between different regions'). In light of this assumption, the key challenge in understanding the emergence of national standard varieties will be examining the *actual extent of variation* between the educated speech varieties of different regions and countries.

The study of differences between *hablas cultas* has traditionally been the domain of Labovian sociolinguistics, which involves conducting interviews with carefully selected informants – typically educated speakers. However, recent standardological research has taken a different approach, shifting its focus from educated speakers as a broadly defined elite group to a more specific group: professional speakers, also known as 'model speakers'. Given the ubiquitous presence of audiovisual mass media in modern speech communities, it is widely accepted that TV and radio presenters serve as linguistic models for the community, that is, as a reference for how the language ought to be used. As a result, certain broadcasting stations and news programmes have gained model status in different language communities, particularly in regard to reflecting pronunciation norms. Examples of this include 'BBC English' in Great Britain, 'Tagesschau-Deutsch' in Germany and the use of Italian by RAI in Italy, as noted by Chalier (2021: 16). Although it is challenging to provide concrete evidence for the influence of mass media speech norms on the speech patterns of the broader community, most scholars agree that such an impact is highly plausible (cf. Reinke 2004: 14; Chalier 2021: 16).[12]

The CO.RA.PAN project is founded on this very assumption that news anchors and other professional speakers employed by radio broadcasting stations are the best representatives of each country's unique national standard variety. It is reasonable to assume that journalists who

12 Numerous studies examining linguistic norms and their representations in Canadian French have demonstrated that the language used by the presenters of Radio-Canada's televised news broadcasts serves as a reference model for a vast majority of Quebec's French-speaking population (for references, see Chalier 2021: 16). Chalier (2021: 16) also highlights the research conducted by Stuart-Smith (2011, 2012) on English, which suggests that specific popular broadcasts have had a significant impact on language use in Glasgow. Older studies by Levy (1979) indicate that viewers may even develop emotional and social attachments to news anchors, despite the one-way communication between televised news broadcasts and their audience.

are hired by broadcasting stations to present and discuss news on air have either undergone specialised linguistic training or were selected for their exceptional eloquence. Furthermore, professional speakers in the broadcasting industry are significantly more accustomed to speaking in public and expressing themselves spontaneously yet eloquently than non-professional educated speakers such as, for instance, university professors or politicians who, while also engaging in public speaking, may not always aim to speak in the most neutral and exemplary manner. While news anchors of the most significant programmes in countries like the United Kingdom (BBC) or Germany (ARD/Tagesschau) typically receive specialised training and ongoing linguistic supervision, it cannot be assumed that every news presenter working for public and private broadcasting stations across the twenty American countries receive similar training. The emphasis on and financial resources allocated to professional linguistic education and training may differ considerably. Nonetheless, the project is based on the premise that by practicing their profession and being broadcasted to large audiences, news anchors and other professional speakers receive a model status for national audiences that is unmatched by any other type of educated speaker.[13] Therefore, they are deemed to be the most suitable linguistic sources for studying the emergence of national standard varieties in the Spanish-speaking world.[14]

13 See, e.g. Ávila (2003, 2016), who studies pronunciation norms in Spanish-speaking America by relying on the model speaker approach.

14 Of course, one could argue that in the twenty-first century, with the rise of social media, traditional audiovisual mass media such as TV and radio may have lost some of their impact. Younger generations of speakers may be more influenced linguistically by the informal language used by social media influencers, who sometimes have tens of millions of followers, than by newsreaders. In this sense, the German philosopher Habermas (2021) has recently argued that the public sphere and the public as we know it has been shifting in recent years towards what he calls *fragmentierte Öffentlichkeiten* 'fragmented public spheres'. Bierbach (2000: 167–8) already suggested two decades ago that the national standard norm 'läßt sich sicherlich nicht mehr allein aufgrund des Sprachhabitus einer bestimmten sozialen Schicht gewinnen' ('can certainly no longer be obtained solely on the basis of the linguistic habitus of a particular social class'). However, as argued

3.2 Authentic oral speech: Different communicative settings in radio broadcasts

Unlike most online corpora that focus on written speech data, the CO.RA.PAN project will gather authentic oral speech data. At least one radio broadcaster with national coverage is selected from each country. The corpus will record the radio station's news and news-related programmes, which are typically broadcast during the early morning and morning hours, when audience numbers are highest (usually between six and ten o'clock local time). The rationale behind this choice is to focus on programmes with the highest number of listeners in each country, as they are most likely to have a significant influence on audiences in terms of linguistic norm diffusion.[15] This approach offers several advantages, including the fact that it provides authentic speech data (as opposed to elicited data in traditional sociolinguistic research) and the ability to compare empirical data across different countries, since radio broadcasts generally have a similar format and structure. Moreover, news broadcasts are an ideal object of study due to their suitability for analysing nationally specific discourses in terms of vocabulary and cultural references, while still providing comparable data on current political and social issues across different countries. This makes them a promising text type for research purposes, as Greußlich (2015: 87) confirms, considering not radio but TV news. According to him,

> noticieros televisivos contienen discursos apropiados en términos de sus características concepcionales: Comparten las ventajas de los textos de prensa diaria,

above, the study of professional speakers' language use remains the most promising – and feasible – approach to investigate emerging national standard varieties.

15 While TV news programmes have the potential to have an even greater impact on national speaker communities, creating a corpus of TV shows, as suggested by Greußlich (see below), would significantly complicate the overall project. The transcription of audiovisual material, which involves the integration of oral speech, texts, and images, would greatly complicate the technical processing of the data.

> ofreciendo una gran ventaja adicional en tanto que facilitan datos fónicos; [...]. Además, tomando en cuenta el alcance geográfico divergente de las diversas cadenas televisivas, que va desde programas provinciales instalados en y para una sola ciudad hasta cadenas globales, un corpus de noticieros televisivos se presta a dar cuenta de la compleja jerarquía de normas con la que se debe contar en el caso de la hispanofonía (Pöll 2012 entre otros).
>
> 'TV news consist of discourses suited in terms of their conceptual characteristics: they share the advantages of texts from the daily press and offer the additional advantage of facilitating phonic data; (...) Furthermore, considering the divergent geographic reach of different TV stations, ranging from provincial programmes located in and broadcasting for individual cities to global broadcasting stations, a corpus of TV news could account for the complex hierarchy of norms to be expected in the case of the Spanish-speaking world (Pöll 2012 among others)'.

The two types of radio programmes that will be included in the corpus can be broadly categorised as follows: 'News broadcasts' refer to programmes that feature daily news segments, which are typically pre-written and read live on air by professional presenters. These shows usually last between one and ten minutes and may include quotes, such as previously recorded speech from individuals related to the news segment, as well as verbal contributions from correspondents working for the broadcaster. The second type of programme recorded during the morning hours, referred to as 'news-related shows' in the corpus, is also presented by professional speakers, sometimes famous journalists. However, these shows consist of less scripted segments, in which one or more speakers discuss current events, often picking up topics from the news programme and engaging in conversations with colleagues and/or guests. While these programmes may include utterances read from a script, they tend to be rather spontaneous speech about previously fixed political or relevant societal topics of the day. Compared to the read news programmes, these shows are much longer and may last several hours.

Although both news broadcasts and news-related programmes provide 'oral speech data', their speech settings and degree of orality are quite distinct from a conceptual standpoint. In this context, it is useful to apply Koch and Oesterreicher's theoretical framework to the concepts of *language of distance* ('Sprache der Distanz') and *language of immediacy* ('Sprache der Nähe') (Koch & Oesterreicher ²2011, 2012; cf. also

Kehrein & Fischer 2016) to describe the speech settings of these broadcasts and evaluate their respective *degree of proximity to standard speech*. Koch and Oesterreicher aimed to overcome traditional dichotomous notions of orality and literacy when analysing speech. To achieve this goal, they adopted a distinction first made by Söll (1974) between conceptual orality and literacy, on the one hand, and medial orality and literacy on the other hand. This approach enables the analysis of all kinds of communicative situations, as a text may be conceptually very elaborated but realised in the oral medium, while a spontaneously produced chat message or WhatsApp message may be written but correspond conceptually to barely elaborated oral speech. This continuum of possibilities underscores the need to consider both conceptual and medial dimensions in the analysis of speech. Koch and Oesterreicher's concepts of 'immediacy' and 'distance' describe the relationship and affinity between communicative factors and the coding strategies employed by speakers in specific speech settings. Since we are applying these terms to different forms of radio broadcasts, it is evident that we are dealing exclusively with speech realised in the oral medium. However, the degree of 'distance' and 'immediacy' varies according to the specific setting, even though it may be assumed that it leans more towards the 'distance' pole, i.e. to conceptually written speech. For instance, news broadcasts rely on elaborated written texts that are read out on air. The interaction is characterised not by dialogic speech between familiar partners in a private face-to-face situation but by monologic speech between a presenter and an anonymous audience of thousands (or even millions) who are spatially distant from the presenter in a clearly public speech setting. The subject matter of news broadcasts is predetermined and maintained throughout the programme, and the presenters are typically emotionally detached from the topics discussed. This is in contrast to private conversations between familiar partners, where a high degree of emotional involvement and affectivity can be expected. The coding strategies employed in news broadcasts (or their preparation) and the resulting speech production correspond perfectly to the ideal of orally transmitted written language (or 'factual standard speech', cf. note 2).

The communicative factors that characterise news-related programmes are similar, but not identical, to those of news broadcasts and offer more potential for variation. Depending on the individual show, these programmes may include monologic segments but are typically more dialogic, with more than one person discussing the news of the day and other topics. The participants in these programmes are often familiar with each other, but they are aware that they are speaking in public in front of a large audience. Although the subjects are mostly predetermined, as in news broadcasts, there is often room for spontaneity. It is noteworthy that the presenters of news-related programmes are sometimes emotionally involved in the topics being discussed, especially since they are often allowed to express their opinions. In sum, while news-related programmes may initially seem similar to private conversations between familiar partners, they are actually more scripted and planned than they appear. Although they are not as elaborated as the news segments read on air, the subjects of conversation are usually predetermined. Importantly, the anchors of these programmes are professional speakers, just like the news presenters (sometimes the same person presents both programmes). Therefore, it is reasonable to assume that, despite a higher degree of spontaneity, news-related programmes provide 'near-standard speech' and are thus a suitable object of study when it comes to describing the emergence of national standard varieties. Nevertheless, even though the corpus gathers authentic radio speech, it is important to consider that the language used by radio presenters may not perfectly reflect the *habla culta* of a country. Instead, it might merely reflect a broadcast norm based on it. Similarly, the corpus can only partially represent the (oral) standard varieties in general, while it will be most representative of the verbal tradition of radiophonic news speech at the present time.

In addition to news broadcasts and news-related programmes, which represent standard and near-standard speech settings respectively, there is a third type of speech setting that appears in both types of programmes: pre-recorded announcements by the radio broadcasters and commercials. Pre-recorded announcements by the radio broadcasters consist of voices introducing radio shows or announcing the upcoming programme schedule for the day and upcoming days. Commercial breaks, on the other hand, include advertising from a variety of actors, ranging from public authorities promoting their services and recommendations (e.g. vaccination campaigns) to local companies.

The fact that all of these formats are pre-recorded is of utmost importance, as they represent highly elaborated speech that is not only based on preconceived written texts or scripts but also on a completely 'unnatural' oral realisation of these texts. This is because the recording is usually done by professionals and repeated until the final product is considered perfect. By contrast, commercial breaks may contain more linguistic variation (or its mimicry), departing from formal standard speech in order to create a sense of closeness to potential customers. Given this context, we can reasonably conclude that pre-recorded announcement segments by radio broadcasters provide valuable insights into the producers' perception of perfect standard speech – either in accordance with the national standard variety or with the traditional (exonormative) standard language. Although pre-recorded announcement segments by radio broadcasters may offer less authentic oral speech than news and news-related broadcasts, they provide valuable insight into how professionals believe language should be used. These segments represent, in other words, the respective nation's 'idealised standard speech'. Therefore, pre-recorded announcements may be an interesting object of study, but it should be noted that they provide much less linguistic corpus material than on-air programmes. Their analysis is most valuable for the study of national norms of pronunciation, but less so for lexical, morphological and syntactic variation.

3.3 Comparing 'idealised standard speech', 'standard speech' and 'near-standard speech': Analytical goals

The CO.RA.PAN corpus aims to offer authentic oral speech data that accurately represents standard language use in comparable speech settings by professional speakers across all Spanish-speaking countries. While one of the main objectives of the corpus is to facilitate country-by-country analysis, its potential analytical applications go beyond that. In order to provide a comprehensive description of pluricentric tendencies in Spanish, it is essential to closely examine each country's internal dynamics. The gathering of data from three different yet 'neighbouring' communicative standard speech settings offers a new approach to the study of internal dynamics since the comparison of specific linguistic phenomena allows us to see how 'advanced' certain features are in becoming standard variants. The reasoning behind this approach is similar to what we know about how language change in general functions. In most cases, language change is the result of innovations (often 'errors' from a normative standpoint) that originate not in formal speech settings but in spontaneous speech from where these innovations might spread to more formal registers and generalise in all kinds of communicative domains. The CO.RA.PAN project starts from the same premise. The emergence of national standard varieties can be conceptualised in terms of 'restandardisation' and specifically 'regionalisation' processes (see Section 2). This means that regional variants formerly considered as substandard in relation to the traditional model of (written) standard speech are re-evaluated by the speech community (i.e. by educated speakers) as standard means of expression and therefore increasingly considered as adequate in formal speech registers and written speech. In this process, the formerly exogenous literary standard language is gradually substituted through historical phases that Schneider (2003), in his study of the *Dynamics of New Englishes*, has dubbed 'nativisation' and 'endonormative stabilisation'. By 'nativisation' Schneider (2003: 247) refers to 'the transition from the acceptance of a distant mother country as the source of both political power and linguistic and cultural guidance to gradual independence – or at least a phase of striving towards it'; the subsequent phase, 'endonormative stabilisation', on the other hand, 'is marked by the gradual adoption and acceptance of an indigenous linguistic norm, supported by a new, locally rooted linguistic self-confidence' (Schneider 2003: 249–50). Metaphorically speaking, the endonormative stabilisation (or regionalisation of the standard language) encompasses thus the 'rise' in status of linguistic variants from substandard to standard. Of course, this gradual acceptance of endonormative variants, that is, the overall shift from a traditional to a new standard model of speech, does not occur without friction among (educated) speakers and leads inevitably to double-standard situations, i.e. to the coexistence of the old (exogeneous) codified standard language based on literary traditions and the upcoming new endonormative standard more rooted in oral use. Applied to the study of emerging national standard varieties in Spanish America by means of the CO.RA.PAN corpus, the basic idea is to describe this process of endonormative stabilisation and the existence of double-standard situations by observing the linguistic habits of professional speakers in the above-mentioned different settings. In practical terms, this entails the possibility, provided by the corpus design (cf. Section 3.4), to compare systematically the use of specific variants and their respective relative frequencies in near-standard speech (news-related programmes) and standard speech (news broadcasts), possibly looking at the 'idealised standard speech' of pre-recorded announcements as well.[16] As an illustrative example, the weakening of /s/ in syllable-final position, a well-known trait of many American varieties, provides intriguing insights. A comparison of the standard speech observed in news anchors in Argentina, Uruguay and Chile shows that the weakening (manifesting as aspiration [h]

Table 10.1 Weakening of syllable-finale /s/ before consonant in news broadcasts.

	Argentina	Uruguay	Chile
[s]	24.2 %	5.9 %	13.8 %
[h]	75.8 %	68.6 %	82.2 %
[ ]	0 %	25.5 %	3.4 %

16 However, the comparatively small amount of data provided by pre-recorded announcements limits the possibility and will at best allow for the analysis of pronunciation norms since the study of lexical and morphosyntactic phenomena requires much more extensive corpus material.

Table 10.2 Weakening of syllable-finale /s/ before consonant in pre-recorded announcements.

	Argentina	Uruguay	Chile
[s]	31 %	100 %	100 %
[h]	69 %	0 %	0 %
[ ]	0 %	0 %	0 %

or elision [ ]) can be regarded as a characteristic feature of all three national standard varieties, with Uruguayan speech being the most progressive.

However, while the weakening is, as expected, even more prominent in the rather spontaneous speech of news-related programmes (accounting for over 90% in each country), the 'idealised standard speech' of pre-recorded announcements shows another picture: Only in Argentina is the difference between these situations minor, indicating that the weakening of syllable-final /s/ unmistakably belongs to the national standard variety. In contrast, in neighbouring countries, the conservation of /s/ as [s] is still idealised, even though it does not accurately reflect actual usage (see Table 10.2). As a result, a dual-standard situation persists.

Another potential research perspective resides in the possibility of comparing the utterances of professional speakers' near-standard speech with those of non-professional (yet mostly educated) speakers that contribute as guests to the broadcasts.

In sum, by providing speech data gathered from different speech settings for each country, the CO.RA.PAN corpus will provide the means to study the dynamics of emerging standard varieties. The possibility to compare the use of specific variants in (a) idealised standard speech, (b) standard speech and (c) near-standard speech in a systematic fashion will offer new research perspectives and hopefully reflect the current dynamics of 'endonormative stabilisation' and the realities of double-standard situations in the Spanish-speaking world. Research perspectives comprise thus both the possibility to study the dynamics of individual countries and to analyse pluricentric tendencies across countries.

3.4 Technical goals: Transcriptions and basic tagging

The technical objectives of the CO.RA.PAN project must align with the outlined conception. However, it is important to note that like any other corpus project centred on oral speech data, there are inherent constraints, such as time and financial limitations, which may affect the extent to which these goals can be achieved. In this section, we will discuss the fundamental technical principles that guide the ongoing work on the CO.RA.PAN corpus, including (a) acquiring and transcribing data, (b) encoding document metadata and (c) encoding speaker types and communication settings based on the overall conception and research perspectives outlined above.

The first technical aspect is data acquisition and transcription. Given that radio stations and their broadcasts are now accessible through the internet and can be recorded while airing or downloaded later as podcasts, acquiring oral data has become relatively easy and potentially limitless, unlike sociolinguistic interviews with informants. However, the subsequent step presents a more significant challenge. As with any corpus of oral speech data, it is crucial to transcribe the oral speech data appropriately in order to facilitate systematic linguistic analysis.

Our initial approach for the CO.RA.PAN project involves providing basic orthographic transcriptions using a set of simple conventions, similar to the *transcripción ortográfica enriquecida* 'enriched orthographic transcription' used in the PRESEEA corpus. This approach enables us to gather as much speech data as possible before incorporating more complex features such as part-of-speech tagging and linguistic annotations at a later stage.[17] Our ultimate objective is to accumulate enough data to enable meaningful quantitative studies on lexical, morphological, syntactic, semantic and pragmatic phenomena. However, transcribing oral

17 The question of how much data is 'enough' – let alone 'representative' – is difficult to answer. The impossibility of achieving the goal of creating a 'representative' corpus is well known (cf., e.g. Kabatek 2013). This applies to both current and past usage.

data, even with a basic orthographic transcription, can be time and cost-intensive, particularly when aiming to create a corpus consisting of millions of words.

To estimate the technical, financial and temporal requirements of the CO.RA.PAN project accurately, a sample study was conducted between November 2021 and July 2022. During this period, we transcribed a sample of approximately thirty-one hours of previously recorded audio documents, gathered from radio stations in four different countries, resulting in a corpus of about 260,000 words. We found that transcribing one minute of this type of speech required between eight and ten minutes for a qualified native speaker with basic training. Consequently, gathering enough data through manual transcription would be too expensive and time consuming even with external funding. The upcoming CO.RA.PAN project will therefore turn to another solution, that is, to automatic transcriptions through fee-based third-party software (such as Amberscript or Transcribe[18]). A small sample transcription of this kind of standard and near-standard speech has shown almost perfect results, although a subsequent review and revision will still be necessary. The cost, however, only amounts to a fraction of what manual transcription would incur.

Another technical goal of the CO.RA.PAN project is to develop a sustainable corpus of transcribed oral speech data that adheres to current research data standards. However, it is important to note that despite the existence of coding conventions such as the highly standardised TEI (Text Encoding Initiative), implementing them to fit the unique requirements of a corpus project like CO.RA.PAN has proven challenging. There is no one-size-fits-all template, and adapting existing templates to suit the individual needs of the CO.RA.PAN corpus is a complex undertaking that has proven quite impractical. Given these challenges, we have decided to utilise TEI conventions solely for the coding of the 'macrostructure' of the corpus, that is, the metadata information for each audio document,

18 The final decision will depend on transcription quality and final pricing. Currently, Transcript is the cheapest option, while Amberscript offers the best transcription results.

including country, radio station, radio programme, date, hour, anchors and length. However, we will not be applying these conventions to the coding of the transcripts themselves.

As for the transcripts of the audio documents, its 'microstructure', they must be encoded according to the basic conception and research perspective discussed in the previous section. This involves systematically distinguishing between (i) different speaker types and (ii) communicative settings during the data processing stage. Speaker types will be categorised into professional speakers, which include anchors and correspondents, non-professional educated speakers such as politicians and scientists who appear on the programmes, and non-professional, non-educated or lesser-educated speakers. Meanwhile, communicative settings will be categorised into different text types, such as news broadcasts (standard speech), news-related programmes (near-standard speech) and pre-recorded announcements of the broadcaster as well as commercials. To ensure the data can be easily accessible and analysed through open-source tools like TXM or CQP, basic coding will rely on specifically adapted XML tags.

The ultimate objective is to create a systematically searchable online corpus that is publicly accessible, similar to recent corpus projects such as the *Corpus Diacrónico y Diatópico del Español de América* (Academia Mexicana de la Lengua). This will be achieved by systematically tagging the text based on the variables mentioned above, as well as through document metadata. As a result, the search engine will be enabled to filter results by country, radio station, speaker type (professional vs. non-professional) and communicative settings (i.e. text types reflecting standard and near-standard speech). It will also be possible to quantify the data according to these categories. The final scale of the corpus and its long-term accessibility, in terms of data storage and web hosting, will of course depend on funding. Nonetheless, our technical provisions, as outlined above, will include a simple TEI (document metadata) and general XML coding (transcripts), which can be easily adapted to future standards, ensuring the sustainability of the corpus.

4. Further research potential: Language representations, attitudes and perceptions

As we have shown in the preceding section, the main purpose of the CO.RA.PAN corpus is to provide systematic and comparable linguistic data that reflect descriptive norms of production and therefore allow us to gain detailed insight into the emergence of new American standard varieties. To get a full picture of pluricentric development in all its aspects, however, it will be necessary to contrast the insight into norms of production with data on both the speakers' representations and attitudes towards how their language ought to be used. Especially relevant for evaluating the degree of 'endonormative stabilisation' of national standard varieties in each country will be the question of how the double-standard situation is evaluated by speakers: while appreciation in terms of correctness for the national standard variety (when compared to the European standard) indicates its consolidation inside the community and the emergence of a national linguistic identity, the opposite shows the prevalence of older conceptions of standardness. To this regard, the data gathered by the CO.RA.PAN corpus may complement existing research and indicate further perspectives for future attitudinal studies (see, for instance, Moreno Fernández & Moreno Fernández 2004; Chiquito & Quesada Pacheco 2014; Cestero & Paredes García 2018). The third important aspect to be considered is the actual perception of national standard features by speech communities. While most studies work with stimuli that were fabricated specifically for such tests in order to have all studied linguistic features perfectly displayed (cf. Tacke 2023: 213), the quality (and amount) of the speech data provided by the CO.RA.PAN corpus should facilitate the possibility to extract authentic yet comparable stimuli across all countries.

5. Conclusion

The CO.RA.PAN project aims to gather authentic speech data from radio stations across all Spanish-speaking countries. The project is based on the assumption that the language used by professional speakers, such as news anchors, best reflects the emergence of national standard varieties. The future online corpus offers two potential research possibilities: Firstly, the corpus will enable systematic and comparative empirical studies of specific linguistic features across countries. Secondly, it will provide the means to study the dynamics of emerging standard varieties within each country by classifying the speech data into different speech settings that reflect (a) 'idealised standard speech', (b) 'standard speech' and (c) 'near-standard speech'. This second perspective also provides an opportunity to gain deeper insights into the realities of double-standard situations across different Spanish-speaking countries, which in turn will contribute to a more comprehensive understanding of the current dynamics of 'endonormative stabilisation'. Consequently, the project offers research perspectives that encompass both the analysis of individual country dynamics and the exploration of pluricentric tendencies across countries.

The outlined technical objectives of the project must align with this conception. To this end, the speech data will be transcribed, and the resulting transcripts will be XML-tagged based on a classification of speaker types and communicative settings. It should be acknowledged, however, that there are inherent constraints, such as time and financial limitations, which may impact the final extension of the corpus.

Finally, it is worth noting that the CO.RA.PAN project is primarily designed to facilitate the analysis of descriptive norms of speech production. Only at a later stage, when the data on speech production is combined with results from studies of language representations, attitudes and perception tests, will it be possible to obtain a comprehensive picture of the development and future of Spanish pluricentricity. As of now, the primary challenge is to collect and process as much speech data as possible.

Bibliography

Amorós-Negre, Carla. 2012. El pluricentrismo de la lengua española ¿un nuevo ideologema en el discurso institucional? El desafío de la glosodidáctica. *Revista internacional de lingüística iberoamericana* 19(1), 127–47.

Amorós-Negre, Carla. 2014. *Las lenguas en la sociedad*. Madrid: Síntesis.

Amorós-Negre, Carla. 2020. Los procesos de restandarización lingüística en la hispanofonía: prescripción y norma mediática de la CNN en Español. In Sebastian Greußlich & Franz Lebsanft (eds). *El español, lengua pluricéntrica. Discurso, gramática, léxico y medios de comunicación masiva*. Göttingen: Vandenhoeck & Ruprecht, Bonn University Press. 271–96.

Auer, Peter. 2017. The Neo-Standard of Italy and Elsewhere in Europe. In Massimo Cerruti, Claudia Crocco & Stefania Marzo (eds). *Towards a New Standard. Theoretical and Empirical Studies on the Restandardization of Italian*. Berlin, Boston: de Gruyter. 365–74.

Ávila, Raul. 2003. La pronunciación del español. Medios de difusión masiva y norma culta. *Nueva Revista de Filología Hispánica* 51(1), 57–79.

Ávila, Raul. 2016. La pronunciación del español en la radio de veinte países hispánicos: comentarios deportivos espontáneos. *Nueva Revista de Filología Hispánica* 64(2), 495–509.

Ayres-Bennett, Wendy. 2021. Modelling Language Standardization. In Wendy Ayres-Bennett & John Bellamy (eds). *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 27–64.

Bell, Allan. 2011. Leaving home: de-Europeanization in a post-colonial variety of broadcast news. In Tore Kristiansen & Nikolas Coupland (eds). *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus. 177–98.

Berruto, Gaetano. 2017. What is changing in Italian today? Phenomena of restandardization in syntax and morphology: an overview. In Massimo Cerruti, Claudia Crocco & Stefania Marzo (eds). *Towards a New Standard. Theoretical and Empirical Studies on the Restandardization of Italian*. Berlin, Boston: de Gruyter. 31–60.

Bierbach, Mechtild. 2000. Spanisch – eine plurizentrische Sprache? Zum Problem von norma culta und Varietät in der hispanophonen Welt. *Vox Romanica* 59, 143–70.

Canfield, D. L. 1981. *Spanish Pronunciation in the Americas*. Chicago: University of Chicago Press.

Cestero, Ana M. & Florentino Paredes García. 2018. Creencias y actitudes hacia las variedades cultas del español actual: el proyecto PRECAVES XXI. *Boletín de Filología* 53(2), 11–43.

Chalier, Marc. 2021. *Les normes de prononciation du français. Une étude perceptive panfrancophone*. Berlin, Boston: de Gruyter.

Chiquito, Ana B. & Miguel Á. Quesada Pacheco (eds). 2014. *Actitudes lingüísticas de los hispanohablantes hacia el idioma español y sus variantes*. Bergen: University of Bergen.

CORDIAM = *Academia Mexicana de la Lengua. Corpus Diacrónico y Diatópico del Español de América*, <www.cordiam.org.> (accessed 08 August 2024).

Elspaß, Stephan. 2021. Language Standardization in a View 'from Below'. In Wendy Ayres-Bennett & John Bellamy (eds). *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 93–114.

Ganswindt, Brigitte, Roland Kehrein & Alfred Lameli. 2015. Regionalsprache.de (REDE). In Roland Kehrein, Alfred Lameli & Stefan Rabanus (eds), *Regionale Variation des Deutschen. Projekte und Perspektiven*. Berlin, Boston: de Gruyter. 425–57.

Greußlich, Sebastian. 2015. El pluricentrismo de la cultura lingüística hispánica: política lingüística, los estándares regionales y la cuestión de su codificación. *Lexis* 39(1), 57–99.

Greußlich, Sebastian & Franz Lebsanft (eds). 2020a. *El español, lengua pluricéntrica. Discurso, gramática, léxico y medios de comunicación masiva*. Göttingen: Vandenhoeck & Ruprecht, Bonn University Press.

Greußlich, Sebastian & Franz Lebsanft. 2020b. Introducción: Pluricentrismo, estándares regionales, normas implícitas y medios de comunicación masiva. In Sebastian Greußlich & Franz Lebsanft (eds). *El español, lengua pluricéntrica. Discurso, gramática, léxico y medios de comunicación masiva*. Göttingen: Vandenhoeck & Ruprecht, Bonn University Press. 11–36.

Habermas, Jürgen. 2021. Überlegungen und Hypothesen zu einem erneuten Strukturwandel der politischen Öffentlichkeit. In *Leviathan. Berliner Zeitschrift für Sozialwissenschaft* (37), 470–500. [Special issue *Ein neuer Struktur der Öffentlichkeit?*, ed. by Martin Seeliger & Sebastian Sevignani].

Joseph, John E. 1987. *Eloquence and Power. The Rise of Language Standards and Standard Languages*. London: Frances Pinter.

Kabatek, Johannes. 2013. ¿Es posible una lingüística histórica basada en un corpus representativo? *Iberoromania* 77, 8–28.

Kehrein, Roland, Alfred Lameli & Stefan Rabanus (eds). 2015. *Regionale Variation des Deutschen. Projekte und Perspektiven*. Berlin, Boston: de Gruyter.

Kehrein, Roland & Hanna Fischer. 2016. Nähe, Distanz und Regionalsprache. In Mathilde Hennig & Helmuth Feilke (eds). *Zur Karriere von Nähe und Distanz*. Berlin, Boston: de Gruyter. 213–57.

Koch, Peter & Wulf Oesterreicher. 2011. *Gesprochene Sprache in der Romania*, 2nd ed. Berlin, New York: de Gruyter.

Koch, Peter & Wulf Oesterreicher. 2012. Language of Immediacy – Language of Distance: Orality and literacy from the perspective of language theory and linguistic history. In Claudia Lange, Beatrix Weber & Göran Wolf, (eds.). *Communicative Spaces. Variation, Contact, and Change. Papers in Honour of Ursula Schaefer*. Frankfurt am Main: Lang. 445–7.

Krefeld, Thomas & Elissa Pustka. 2010. Für eine perzeptive Varietätenlinguistik. In Thomas Krefeld & Elissa Pustka (eds). *Perzeptive Varietätenlinguistik*. Frankfurt am Main: Lang. 9–30.

Kristiansen, Tore. 2021. Destandardization. In Wendy Ayres-Bennett & John Bellamy (eds). *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 667–90.

Lebsanft, Franz. 2004. Plurizentrische Sprachkultur in der spanischsprachigen Welt. In Alberto Gil & Osthus, Dietmar/Polzin-Haumann, Claudia (eds). *Romanische Sprachwissenschaft. Zeugnisse für Vielfalt und Profil eines Faches. Festschrift für Christian Schmitt zum 60. Geburtstag*. Frankfurt a.M.: Lang. 205–20.

Lebsanft, Franz. 2007. Norma pluricéntrica del español y Academias de la Lengua. In Christopher F. Laferl & Bernhard Pöll (eds). *Amerika und die Norm. Literatur sprache als Modell?* Tübingen: Niemeyer. 227–46.

Lebsanft, Franz & Felix Tacke. 2020. Romance standardology: Roots and traditions. In Franz Lebsanft & Felix Tacke (eds). *Manual of Standardization in the Romance Languages*. Berlin, Boston: de Gruyter. 3–59.

Lebsanft, Franz, Wiltrud Mihatsch & Claudia Polzin-Haumann (eds). 2012. *El español, ¿desde las variedades a la lengua pluricéntrica?* Madrid: Iberoamericana.

Levy, Mark R. 1979. Watching TV news as para-social interaction. *Journal of Broadcasting* 23, 69–80.

Lipski, John M. 1994. *Latin American Spanish*. New York: Longman.

Mattheier, Klaus J. 1997. Über Destandardisierung, Umstandardisierung und Standardisierung in modernen europäischen Standardsprachen. In Klaus J. Mattheier & Edgar Radtke (eds). *Standardisierung und Destandardisierung europäischer Nationalsprachen* Frankfurt a.M.: Lang, 1–9.

Moreno Fernández, Francisco & Juliana Moreno Fernández. 2004. Percepción de las variedades lingüísticas de español por parte de hablantes de Madrid. *Lingüística Española Actual* 26(1), 5–38.

Muhr, Rudolf. 2016. The state of the art of research on pluricentric languages: Where we were and where we are now. In Rudolf Muhr (ed.). *Pluricentric Languages and Non-Dominant Varieties Worldwide. Part I: Pluricentric Languages across Continents. Features and Usage*. Frankfurt a.M., Wien: Lang. 13–37.

NGLE = Real Academia Española & Asociación de Academias de la Lengua Española. 2009/2011. *Nueva gramática de la lengua española*. Madrid: Espasa.

Oesterreicher, Wulf. 2000. Plurizentrische Sprachkultur – der Varietätenraum des Spanischen. *Romanistisches Jahrbuch* 51, 287–318.

Pöll, Bernhard. 2012. Situaciones pluricéntricas en comparación: el español frente a otras lenguas pluricéntricas. In Franz Lebsanft, Wiltrud Mihatsch & Claudia Polzin-Haumann (eds). *El español, ¿desde las variedades a la lengua pluricéntrica?* Madrid: Iberoamericana. 29–46.

Pöll, Bernhard. 2021. Spanish today. Pluricentricity and codification. In Danae Perez, Marianne Hundt, Johannes Kabatek & Daniel Schreier (eds). *English and Spanish. World Languages in Interaction*. Cambridge: Cambridge University Press. 163–83.

Postlep, Sebastian. 2010. *Zwischen Huesca und Lérida. Perzeptive Profilierung eines diatopischen Kontinuums*. Frankfurt a.M.: Lang.

PRESEEA = Moreno Fernández, Francisco. 2014–. *Proyecto para el estudio sociolingüístico del español de España y de América*. Alcalá de Henares: Universidad de Alcalá. <https://preseea.uah.es/> (accessed 30.12.2024).

Prifti, Elton. 2021. Hacia el pluricentrismo: moldeamientos modernos del español en América. In Eva M. Eckkrammer (ed.). *Manual del español en América*. Berlin, Boston: de Gruyter. 183–206.

Pustka, Elissa, Marc Chalier & Luise Jansen. 2017. À la recherche d'une norme de prononciation: le modèle des présentateurs de télévision. *Journal of French Language Studies* 27(1), 101–15.

Reinke, Kristin. 2004. *Sprachnorm und Sprachqualität im frankophonen Fernsehen von Québec. Untersuchung anhand phonologischer und morphologischer Variablen*. Tübingen: Niemeyer.

Schiffman, Harold F. 1998. Standardization or restandardization: The case for 'Standard' Spoken Tamil. *Language in Society* 27(3), 359–85.

Schmidt, Jürgen E. 2004. Die deutsche Standardsprache: Eine Varietät – drei Oralisierungsnormen. In Ludwig M. Eichinger & Werner Kallmeyer (eds). *Standardvariation – Wie viel Variation verträgt die deutsche Sprache?* Berlin, New York: de Gruyter. 278–305.

Schmidt, Jürgen E. & Joachim Herrgen. 2011. *Sprachdynamik. Eine Einführung in die moderne Regionalsprachenforschung*. Berlin: Erich Schmidt.

Schneider, Edgar W. 2003. The Dynamics of New Englishes: From Identity Construction to Dialect Birth. *Language* 79(2), 233–81.

Söll, Ludwig. 1974. *Gesprochenes und geschriebenes Französisch*. Berlin: Schmidt.

Stepanov, Georgij V. 1971. Algunas cuestiones metodológicas del español americano. In Alexandru Rosetti & Sandra Reinheimer-Rîpeanu (eds). *Actele celui de-al XII-lea Congres International de Lingvistică şi Filologie Romanică*. Bukarest: Editura Academiei. 1165–7.

Stepanov, Georgij V. & Aleksandr D. Švejcer. 1981. Toward a Study of Transplanted Languages. In Brigitte Schlieben-Lange (ed.). *Logos semantikos. Studia linguistica in honorem Eugenio Coseriu 1921–1981*. Vol. 5: *Geschichte und Architektur der Sprachen*. Berlin, New York: de Gruyter. 219–25.

Stewart, William. 1968 [1962]. A sociolinguistic typology for describing national multilingualism. In Joshua A. Fishman (ed.). *Readings in the Sociology of Language*. Den Haag: Mouton. 530–45.

Stuart-Smith, Jane. 2011. The view from the couch. Changing perspectives on the role of the television in changing language ideologies and use. In Tore Kristiansen & Nikolas Coupland (eds). *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus. 223–39.

Stuart-Smith, Jane. 2012. English and the media. Television. In Anton Berg & Laurel J. Brinton (eds). *English Historical Linguistics*. Vol. 1. Berlin, New York: Mouton de Gruyter. 1075–88.

Tacke, Felix. 2011. Plurizentrik und normativer Diskurs in der *Nueva gramática de la lengua española*. *Romanische Forschungen* 123(2), 145–66.

Tacke, Felix. 2023. Review of: Chalier (2021). *Romanische Forschungen* 135(2), 210–5.

Tacke, Felix. 2024. Die Rekonfiguration und Modernisierung romanischer Standardsprachen: Zu einem doppelten Paradigmenwechsel im Italienischen, Spanischen und Katalanischen. In Lidia Becker, Julia Kuhn, Christina Ossenkop, Claudia Polzin-Haumann & Elton Prifti (eds). *Zwischen Pluralität und Präskription. Sprachnormen in der Romania in Geschichte und Gegenwart*. Tübingen: Narr. 27–65.

KONSTANTIN NIEHAUS

11 Double standards in Austria? Enregistering 'Austrian German' and 'German German' in the late twentieth century[1]

ABSTRACT

This article addresses a sociolinguistic dilemma that the national codification of 'Austrian (Standard) German' has created: Although the national codex in Austria, the Austrian Dictionary, successfully counters monocentrism (which considers only one variant to be the 'correct' standard), it also applies top-down language planning dominated by Vienna and supports calls for a purist separation of 'Austrian German' and 'German German'. On the other hand, critiques of the dictionary often involve an elitist, conservative perspective. Additionally, both supporters and critics can also follow implicit political stances. This article utilises the theoretical framework of enregisterment to analyse how 'Austrian German' has come to be recognised as a socially meaningful semiotic register which indexes national identity. A diachronic critical discourse analysis of newspaper and magazine articles focuses on the constitutive phase of national codification in Austria during the late twentieth century and presents a qualitative analysis of the discourse on 'Austrian German', its socio-political background, agents and their motivations.

1 I am indebted to Erik Schleef (Salzburg), Mason Wirtz (Zurich) and two anonymous reviewers for their helpful comments.

1. Motivation: Who created double standards in Austria and why?

'Austrian (Standard) German' presents a particularly interesting case for research on Standard German: Austria is the only German-speaking nation in which a language codex focused on national variants, *austriacisms* (variants specific to Austria), has been established by the government: the *Österreichisches Wörterbuch* 'Austrian Dictionary' (ÖWB). Its establishment was directed against the monocentric idea that there should only be one correct standard variety dominated by Germany's usage (Ammon 1995: 218–19). Until today, national identity[1] in Austria is often construed via perceived national standard variants (Auer 2013) and the existence of a national dictionary plays a key part in this. Specific linguistic forms and specific language use (e.g. the use of a certain variety or language) evoke constructions of (national) identity. The latter can be understood as a part of the social meaning of language, i.e. 'a concept that encompasses matters such as register (in the narrower sense of situational appropriateness), stance (certainty, authority, etc.) and social identity (class, ethnicity, interactional role, etc.)' (Johnstone et al. 2006: 81).[2] In our case, speakers summarise the linguistic forms imbued with the social meaning of 'Austrian' under the term 'Austrian German' (AG), and the selection of forms is generally dominated by East Austrian features, not least in the Austrian Dictionary (cf. Ammon 1995: 211). In contrast, the actual use of Standard German, i.e. *Gebrauchsstandards* 'standards of usage' in model contexts (Ammon 1995: 88) varies considerably across nations and regions: Austria shares many features of standard lexis,

1 Identity is seen here as a flexible construction which marks social belonging in a specific context, based on culturally conventionalised practices (based on Kroskrity 2001).

2 If the term '(social) identity' appears too vague or problematic to the reader, we can alternatively utilise the distinction between 'social types' associated with demographic categories, 'personae' in a specific local context, and 'stances' during interactions (cf. Moore & Podesva 2009). Then, we can classify 'Austrian' as a social type.

phonology and grammar with southern Germany and German-speaking Switzerland. In addition, there is a clear inner-Austrian division between the western provinces of Vorarlberg and Tyrol and the eastern areas (Elspaß & Kleiner 2019). Thus, discrepancies between the codification of a national standard and the often non-national standard of use in Austria can be seen as double standards, created by national symbolism.

From a critical perspective, the case of AG presents a dilemma to normative authorities: to conceptualise a national standard for reasons of 'linguistic justice' amongst national standard varieties (Oakes 2021) creates the injustice of ignoring or downplaying inter- and intra-national standard variation for reasons of national linguistic unity (Auer 2013: 41). This means that *certain* variants – those which index 'Austrian' and evoke a respective social identity – more easily make their way into a national codex than others. Based on that, we should understand AG as an indexical, socio-semiotic 'register' rather than a variety (Auer 2013).

This article argues that social meaning is crucial in understanding the interaction between language use and codification, more specifically in the case of Austria. The focus of this article is therefore to highlight issues of historical language policy which were involved in the national codification of AG and how social meaning was assigned to a selection of linguistic variants. To do so, we need to look at the (historical) discourse, understood here broadly as a semiotic ensemble of human practices including texts and dialogues which feature an overarching cultural, political, or social topic meaningful to the construction of identities (based on Reisigl & Wodak 2009: 89). Central research questions are: who created AG, i.e. which speakers or institutions have held authority and have shaped the view of AG in historical (public) discourse? What motivated speakers to take part in the discourse on AG? How is the codification via the Austrian Dictionary reflected in the discourse and *vice versa*?

There is no in-depth empirical analysis of how the 'register' AG emerged during the late twentieth century (more general accounts, of course, exist, e.g. Ammon 1995: 137–42, 181–96; Wiesinger 2014: 231–58). This article will therefore analyse the major ideological issues in the discourse on AG and its codification via the Austrian Dictionary, and identify and critically reflect on agents, topics and motivations in a broader socio-cultural context. I will sketch out the codification in Austria with a special focus on the ÖWB. Afterwards, I will elaborate on the idea of AG as a register in a theoretical section on *enregisterment* (Agha 2007), i.e. the process of how registers are construed and function in discourse. I will then lay out the approach of a Critical Discourse Analysis (CDA) and how it can be applied in our case, i.e. as a qualitative analysis of the historical metalinguistic discourse in newspaper and magazine articles of two crucial phases in the history of Austria as a nation-state after World War II:

- the debate surrounding the ÖWB's emergence in 1946–51
- the controversy of the completely reworked edition of the ÖWB in 1979–80

This analysis should provide new evidence on how the social meaning of AG emerged and developed. An outlook will point out further discursive processes in Austria's recent history which would be well worth analysing.

2. Codification in Austria and the ÖWB

An extensive recapitulation of codified Standard German in Austria and the ÖWB is beyond the scope of this chapter. Hence, I will focus on the aspects most relevant to the present study.

Codification in Austria is partly exonormative (based on foreign usage) and partly endonormative (based on native usage; for the distinction: Ayres-Bennett 2021: 33). The Germany-based Duden volumes and, since 2011, the free services of the publisher's website are often considered 'supra-regional' authorities even outside of Germany, in places such as Austria. Despite the inclusion of Austrian usage, the Duden fits into older traditions of exonormative codification in Austria. It is based on the perceived superiority of northern German usage, a notion that has been present since the eighteenth century (Stevenson et al. 2018: 62).

Drawing on previous efforts (Ammon 1995: 121), the idea of endonormative codification was revitalised by the Austrian all-party government of Social Democrats (SPÖ), the conservative People's Party (ÖVP) and, until 1947, the communist party (KPÖ) after World War II: An official Austrian Dictionary, the ÖWB, was commissioned by the Minister of Education, Felix Hurdes (ÖVP), in 1947. At the instigation of the Ministry, questionnaires regarding language use in Austria were sent to language authorities such as teachers, publishers, editors, typographers, literary writers, journalists and the association of Viennese professors of German (*Wiener Sprachverein*). Example pages were issued for peer-group reviewing in 1950 but found their way into public debate (Wiesinger 2014: 259). Following lively discussions, the ÖWB's publication was delayed until 1951. Since then, the ÖWB has been published by the *Österreichischer Bundesverlag* (öbv) 'Austrian federal publisher'. Contrary to what the name may suggest, the öbv is a private company. In 1967, the publishers established a *Wörterbuchstelle* 'lexicon office' (Sutter 2017: 167). The undisputed position of the ÖWB as the country's codex was controlled through various means until the 1970s, e.g. through provision in schools free of charge. The ÖWB editions of the 1970s seem to have drawn on the Duden volume *Wie sagt man in Österreich?* 'How do they say in Austria?' (Ebner 1969), a mixture of an endonormative codifier and exonormative publisher. Until the ÖWB's thirty-fifth edition in 1979, there seem to have been only minor changes compared to the first edition (Wiesinger 1980: 368). For the thirty-fifth edition, the ÖWB was completely revised under the editorship of Erich Benedict, Maria Hornung and Ernst Pacolt. This included the reselection of lemmata, method of stylistic markings and the introduction of new spellings. However, subsequent editions abandoned many of these innovations. Furthermore, the ÖWB has followed the supra-regional spelling reforms of the 1990s (notwithstanding traditional Austrian spellings also exonormatively codified in the Duden). Since 2016, there is an online version. As of 2024, the ÖWB online service can be used without registration. In theory, the adherence to the ÖWB is mandatory in schools and government bodies (Wiesinger 2014: 259), although in practice speakers may rely on other sources such as the Duden codex (Schmidlin 2011: 120).

The ÖWB is subject to regular criticism which historically has centred on the following main issues:

(a) destandardisation:[3]

It is argued that the ÖWB endangers its prescriptive function and lowers stylistic standards via the codification of colloquial and dialect variants. This argument appears to have been particularly strong with the thirty-fifth edition (1979) as it left several non-standard variants unmarked (Huesmann 1997). From a critical sociolinguistic perspective, it will be important to investigate why non-standard features are featured in an otherwise standard codex and how public discourse has influenced the selection.

(b) dominance of codified East Austrian, particularly Viennese language use:

The ÖWB is edited in the national capital Vienna and – despite differing standards of usage – has not sufficiently relied on West Austrian sources in its history, thus introducing a Viennese monocentrism and national homogenisation 'by the back door' ('durch die Hintertür', Moser 1995: 176), creating an imbalance of power between West and East Austrian speakers as well as normative authorities within the country. Far from being unusual during standardisation processes, this imbalance is relevant to a critical sociolinguistic evaluation of democratic language policy as corpus planning in that vein often results in a perceived inauthenticity of standard (Costa, De Korne & Lane 2017: 2).

3 For a comprehensive discussion of a possible destandardisation of German cf. Auer & Spiekermann (2011).

(c) national variety purism (Ammon 1995: 184):

The ÖWB devalues 'German German' (GG) influence. It includes perceived GG variants, sometimes also to paraphrase AG lemmata. In cases where GG variants such as *Sahne* 'cream', *Kloß* 'dumpling' and *Tomate* 'tomato' receive separate entries, the ÖWB indicates via asterisk that they are considered 'non-domestic' or 'uncommon in Austria' in contrast to *Obers*, *Knödel* and *Paradeiser* (Ammon 1995: 181; Reiffenstein 1995: 163). Effectively, this is an ethnicist stigmatisation of Austrian speakers (and German migrants) who converge towards or regularly use these linguistic forms (cf. Glauninger 2015; Greth & Köllen 2016).

(d) lack of methodical consistency and transparency:

The ÖWB has relied much on the personal knowledge and preferences of the editors and not on an extensive corpus of Standard German in Austria (Moser 1995; Ammon 1995: 134, 189–90; Sutter 2017: 174–5; Langer 2021: 6), which time and again has resulted in inconsistent diatopic and stylistic markings as well as omissions (Wiesinger 1980: 370; Reiffenstein 1995: 162). Another frequent point of criticism concerns the non-transparent methods and planning process as well as the anonymity of either editors, consultants, or members of the lexicon office (Reiffenstein 1995: 158; Sutter 2017: 175). This entails top-down language planning which particularly fosters the imbalance described in (b).

The fact that the criticism has been voiced in public by linguists as well as non-linguists highlights the complex discursive interplay of norm authorities. An analysis should take this complexity into consideration as well as the objects of debate, i.e. the selection of linguistic signs which qualify for codification purposes: 'The variety which emerges in discourse is neither a coherent linguistic system nor the consistent depiction of an empirically tangible variety' (Huesmann 1997: 196, my translation). What, then, is AG in discourse?

3. AG as a socio-semiotic register (*enregisterment*)

Enregisterment describes 'processes and practices whereby performable signs become recognised (and regrouped) as belonging to distinct, differentially valorised semiotic registers by a population' (Agha, 2007: 81). The semiotic resources of registers in Agha's sense comprise linguistic elements like lexical, phonetic, or typographic variants and index 'stereotypic features such as interlocutors' roles, relationships and the type of social practice in which they are engaged' (Agha, 2007: 148). Instead of empirical evidence, linguistic ideologies lie at the heart of enregisterment, i.e. sets 'of beliefs about language articulated by users as a rationalisation or justification of perceived language structure and use' (Silverstein 1979: 193).

Agha (2007) uses the concept of enregisterment to explain the emergence of standardisation in English, a process in which dictionaries played a central role as they indicate the existence of a certain 'civilised' and 'cultivated' style of speech (Agha 2007: 207–9), as is the case with pronunciation, i.e. Received Pronunciation, RP (Agha 2007: 190–1).

Both RP and AG can be referred to as 'registers' in Agha's use of the term. A register is not so much a variety with a fully grown linguistic structure but a socially meaningful set of select linguistic features. AG is chiefly construed via linguistic 'double stereotyping' of Austrian and German speakers, often focused on pronunciation and lexis (Auer 2013: 41). Common examples include words like *Semmel* 'bread roll' and *Knödel* 'dumpling' which, against empirical evidence, are presented as purely 'Austrian' words in opposition to perceived GG *Brötchen* and *Kloß* (Auer 2013: 32–4). The register and its negative GG counterpart are meant to complement each other and regularly appear simultaneously in context. The ÖWB's explanations of AG lexemes via GG equivalents can be viewed as an example of such an enregisterment.

Registers do not primarily serve instrumental, communicative purposes but social differentiation and the construction of a national identity (Auer 2013: 41). Austrian speakers draw bi-national linguistic comparisons to stress their non-Germanness and to criticise Germany's factual as well as perceived dominance in population, economy and culture – in fact, the very principle of AG seems to be that it is *not* GG (Auer 2013: 20; 41).

Closely related to that is the issue of standardness: Although the register AG includes several standard features – not least administrative terms like *Matura* 'A-levels', or words for food –, it also encompasses supposedly un-German practices and features. These are perceived as typically Austrian not despite but because of their nonstandardness, e.g. the conscious shifting to dialect in public speeches (Glauninger 2015: 46–7) or phrasemes which index a down-to-earth, cosy lifestyle but also a certain bawdiness of AG (de Cillia et al. 2020: 102; Stevenson et al. 2018: 63). Thus, Austrian speakers can use *österreichisches Deutsch* 'AG', *Österreichisch* 'Austrian', or *österreichische Sprache* 'Austrian language' to refer to a rather diverse repertoire of standard and non-standard features (de Cillia et al. 2020: 100). Efforts to codify perceived non-standard features in an otherwise standard codex such as the ÖWB's approach reflect this: 'Austrian' as a particular social meaning is indicated here not only through standard variants but also through features that are perceived to be non-standard.

Compared to variation in standards of usage (as well as non-standard), the enregisterment of AG

> leads, on the one hand, to an *ideological 'blow up'*, when geographically restricted forms are upgraded to nationwide standard forms (as in the example of [the East Austrian variants *Fleischhauer/-hacker*] 'butcher' [...]), and on the other hand, to an *ideological 'shrinking'*, when wider geographical distributions are reduced to fit the state territory [e.g. in the case of *Semmel* 'bread roll' which is actually used in Austria as well as southern Germany, K.N.]. (Auer 2021: 44)

Both ideological strategies are relevant to national identity: Dressler & Wodak (1983: 252–4) speak of a separatist function towards GG and a solidarity function of AG within Austria. The result with Austrian speakers is purism towards national varieties, *national variety purism* (Ammon 1995: 184): GG should be clearly separable from AG. Purist speakers favour the idea of a homogenous language use on both sides of the national border. Again, the ÖWB provides examples of this ideology with the preference for East Austrian features (ideological blow-up), the inclusion of widely spread, mostly southern features (ideological shrinking) and the marking of GG origins of certain words (see Section 2).

Based on these premises, we will now turn to our diachronic analysis to see how the enregisterment of AG and the prevalent ideologies unfolded in historical discourse.

4. Historical CDA and data

Both enregisterment and CDA concentrate on metapragmatic discourse. The difference between enregisterment and CDA is that enregisterment so far has rather asked how language ideologies create socially meaningful discourse objects. In contrast, CDA in the tradition of Fairclough (1995) is more interested in a critique of 'institutionally reproduced power' (Blommaert 2005: 24), e.g. a critique of the authorities who influence the public perception of a discourse object (such as a language or a variety) by implementing and enforcing their language ideologies. However, to impose power onto an object, authorities need to enregister it. In that sense, codification processes *per se* as well as the affirming or intervening public discourses should be seen as parts of enregisterment.

Any discourse data should stem from authentic social practices, i.e. communication in context (Reisigl 2013: 75). This often means focusing the investigation on public written data such as opinion pieces, political speeches, etc., as opposed to elicited group discussions or interviews. Context is understood here as intertextual and interdiscursive. An historical CDA investigates the historical progression of intertextuality, e.g. the history of quotes or historical social meaning referenced or alluded to (Reisigl 2013: 79). This method ties in with enregisterment in which the process of 'enregistering' language and linguistic features is rarely ever a *tabula rasa* creation of meaning but in principle follows a quote-of-quote-structure to modify existing social indices (Hall-Lew, Moore & Podesva 2021: 12).

In addition, interdiscursivity considers the layering of macro-political discourses, e.g. left-wing vs. right-wing stances and the individual discursive topic. Evaluations of AG in general and endonormative codification in particular have been described as an index of political affiliation: left-wing 'patriotically Austrian' supporters and right-wing pan-Germanic opponents (e.g. Wodak 1994). Yet, this indexicalisation needs to be validified through further evidence.

Although the current approach differs from typical enregisterment studies which use qualitative interviews (e.g. Johnstone, Andrus & Danielson 2006) or a multimodal analysis (Agha 2007), a few studies combining enregisterment and CDA have been conducted before (e.g. Gal 2019). Oriented towards these, the following analysis will focus on public written data taken from national and regional newspapers and magazines, including all sorts of texts (mostly opinion pieces) such as squibs, letters to the editor and commentaries.

As explained in Section 1, the sources stem from two crucial periods of discourse on the codification via the Austrian Dictionary: (1) its origins, formation and establishment during 1946–1951; (2) the completely reworked edition and the ensuing debate in 1979/80. Data for the period of 1946–1951 (see appendix) stem from *Austrian Newspapers Online* (<https://anno.onb.ac.at>, accessed 23 March 2023) where I collected relevant texts via a simple text search for AG (*österreichisches Deutsch*), the ÖWB, or similar search items such as *österreichisches Schriftdeutsch* 'Austrian written German'.[4] Sufficient data for the period 1979/80 was hard to obtain. I mostly relied on hard copies of the newspaper and magazine articles mentioned in Möcker (1980), Clyne (1985) and Wiesinger (1980). Letters to the editor are rather short, i.e. of approx. 100–200 words. Squibs and reports typically comprise 300–700 words. Commentaries, especially from the period 1946–1951, are longer and reach approx. 1,000–2,000 words (for sources and labelling system see appendix). It is clear that further texts may exist, which is why this investigation should only be considered a first attempt of uncovering the historical codification discourse.

4 Doublets such as agency texts were collected but excluded from analysis.

What follows is a thematical analysis focusing on (a) the intertextuality of enregistering AG (e.g. register names, single linguistic features) and (b) the interdiscursivity of (linguistic) ideologies which concern AG. My focus will lie on sources with an explicit intertextuality, i.e. references to preceding texts. For the present study this means that sources with multiple layers of coding will be treated with analytic priority since they provide a more complex intertextuality and interdiscursivity. This leads to a slight bias towards longer pieces such as reports and commentaries during the presentation of examples, however the study tries to compensate for this as much as possible and considers shorter texts wherever space allows.

MAXQDA was used to code the data with manifest content tags such as *österreichisches Deutsch* ('AG'), *Unterrichtssprache* ('language of instruction', see 5.3) or other register names. Additionally, all variants which were subject to metalinguistic discourse were tagged individually (e.g. *Obers* 'cream', PRESSE_1979_report_editors). After that, another layer of codes was added for language policy themes such as 'federalism', 'linguistic justice', positionings for or against an endonormative codification ('pro' or 'contra'), or 'educational relevance of codices'. Finally, themes related to linguistic ideologies were coded based on typical words and phrases, with categories such as 'loan word purism' (*Reinerhaltung* [sic] 'keeping clean and preserving', PRESSE_1979_lett1_Wolfram), 'purism of national variety' (*unösterreichisch* 'un-Austrian', NZ_1950_report_editors) and 'linguistic conservatism' (e.g. *kultureller Verfall* 'cultural decay', DKV_1950_comm_ra).

The analysis means to provide evidence for the historical context from which AG emerged and offer insights on the power balance in respective discourses, including reflections on the contemporary scientific critique (Reisigl 2013: 75) in the case of 1979/80.[5] I will first present a chronological overview and then analyse the discourse on specific issues of language policy. Of course, the interdiscursivity of the data does not allow for a full separation of 'stand-alone' sub-discourses, so the analysis is organised around four specific ideological aspects present in both periods of

5 To my knowledge no equivalent scientific debate in the period 1946–1951 existed.

time: anti-German purism, Austrian federalism, warnings of destandardisation and economic concerns. These will be analysed based on Reisigl's 'five simple questions' on discursive strategies (Reisigl 2017: 52):

1. How are persons, objects, phenomena, etc. named and referred to? (nomination)
2. What characteristics or qualities are attributed to social actors, objects, phenomena etc.? (attribution/predication)
3. What arguments are employed? (argumentation)
4. From what perspective are these nominations, attributions, arguments expressed? (perspectivisation)
5. Are the respective utterances articulated overtly, are they intensified or mitigated? (modification, or 'intensification and mitigation', Reisigl 2017: 52)

After that, I will turn to a more general layer of discourse and embed the results in a critical reflection of socio-cultural implications, in particular references and associations with either conservative or progressive politics. Note that this study should be regarded as a pilot study rather than a full-blown CDA research project and further investigations will be needed (see Section 6.).

5. Analysis

5.1 Discursive chronology

Early coverage in the perceived communist newspapers *Neues Österreich* (all-party medium but editor in chief was Ernst Fischer of the KPÖ)[6]

6 Information on individuals and media taken from Pfalzgraf (2019), ANNO, and Stadt Wien (online).

and *Die Neue Zeit* (newspaper of the KPÖ in Upper Austria) in 1948/49 focused on the presumed denazification function of the ÖWB and praised the concept of codified national standards of usage, called *Umgangssprache* 'colloquial language' or *echtes Schriftösterreichisch* 'real written Austrian' (e.g. NOS_1948_report_APA, NOS_1948_report2_editors, NOS_1949_report3_editors). Similarly, Edwin Rollet, chairman of the Austrian Writers' Association and anti-Nazi journalist with the republic-owned newspaper *Wiener Zeitung* (WZ_1950_comm_Roll), wrote an article in favour of the ÖWB and what he called *Österreichisches Deutsch* 'AG'. He was against the linguistic separatism of the political left but in support of a codification of national standards of usage.

The monthly culture magazine *Die Wiener Bühne* reacted to the reports of *Neues Österreich* in January 1950 and contributed to the discussion with complex reasoning pieces by editor-in-chief Franz Taucher and journalist Erik G. Wickenburg against simplifications and national stereotypes (BUEHNE_1950_comm_Egw, BUEHNE_1950_comm_Tau); although also a perceived leftist paper, the articles argued against the ÖWB and for an elitist preservation of norms. They also voiced economic concerns for Austrian companies depending on exports to Germany if a separate national orthography should be established.

Hans Weigel, Viennese author and journalist, apparently spoke in person to the ÖWB commission and authored a longer critical commentary in the Viennese national newspaper *Weltpresse* (moderate leftist) which dismissed codified national spelling alternatives and called for an open debate in mid-January 1950 (WP_1950_comm_Weigel1). Numerous letters to the editor expressed opposition to the creation of *Österreichisch* 'Austrian' or *österreichische Schriftsprache* 'Austrian (written) language' during the next weeks (WP_1950_lett1_Kunzl to WP_1950_lett13_Bachler). At the same time in Salzburg, the major regional newspaper *Salzburger Nachrichten* under assistant editor-in-chief Bruno Kornel Hillebrand-Skrehunetz, an ex-Nazi propagandist (Rathkolb 2015: 247), published several critical pieces (SN_1950_comm_Hille, SN_1950_comm_Hala, SN_1950_squib_Peris1–2) which ridiculed the planned ÖWB. Further critique and a call for a more democratic process was encouraged by an anti-communist piece in the *Oberösterreichische*

Nachrichten (OON_1950_Fl). In Tyrol, Eduard Widmoser criticised the dominance of the East Austrian dialect in the ÖWB and doubted the usefulness of a national codex (TBZ_1950_comm_Wid).

Meanwhile, in conservative media such as the magazine *Die Furche* and the newspaper *Das Kleine Volksblatt*, the debate focused on economic concerns: Hugo Zienert, secretary in the Ministry of Finance, voted against a national codification, partly met with approval by letters to the editor but countered by Alfred Missong, a founding father of the ÖVP, anti-Nazi Austrian nationalist and member of the commission of the ÖWB (FURCHE_1950_comm_Zie, FURCHE_1950_lett1_anonym, FURCHE_1950_lett2_anonym,_FURCHE_1950_comm_Miss). An anonymous opinion piece as well as an article by Fritz Bock, Viennese politician of the conservative ÖVP, denied the economic threat of linguistic separatism for Austrian publishers and focused on what he perceived as benefits of national codification such as the fostering of an un-Prussian and non-Nazi national identity (DKV_1950_ra/Bock/report).

In the wake of these protests and a formal complaint by the westernmost federal state Vorarlberg, the Minister of Education Felix Hurdes and his Secretary Stur denied the plans to nationalise spelling codification and to codify a significant number of dialect variants in late February and early March 1950 (WZ_1950_report_editors, DKV_1950_report_editors, NOS_1950_report4_editors, WP_1950_report_editors). Nonetheless, complaints about the Viennese dominance, the non-transparent editing and the economic plausibility of the ÖWB continued especially outside East Austria until the publication of the ÖWB in 1951 (OBC_1950_comm_Stein, VN_1950_comm_editors, VOW_1950_lett_editors, OON_1950_comm_editors, SN_1950_squib_Peris3).

Let us now turn to the progression of discourse in 1979/80. In contrast to the earlier discourse 1946–1951 – probably also due to the few sources available – the debate in 1979/80 runs towards almost unequivocal rejection of the reworked edition:

Although the dictionary came out earlier in 1979 (Weitschacher 1981: 41), the discussion first gained momentum in November/December 1979 when negative press coverage in the then-leading national conservative newspaper, *Die Presse*, set in. Karl Hirschbold, a Viennese journalist specialised in language cultivation, published a shorter negative comment on the uselessness of the ÖWB in his regular squib (PRESSE_1979_squib_Hirsch1), quickly followed by a longer, sarcastic piece in the newspapers' miscellaneous section (PRESSE_1979_report_editors) which utilised perceived non-standard additions in the ÖWB to ridicule both the critic Erwin Mehl, Viennese sports scientist and ethnonationalist, and the ÖWB's spokesperson Maria Hornung, a Viennese linguist. *Die Presse* went on to print letters to the editor which were critical of the ÖWB, including one by Fritz Wolfram, Viennese teacher and right-wing politician with the FPÖ (PRESSE_1979_lett1_Wolfram), both favouring a normatively stricter codification. A large opinion piece by the Viennese magistrate official Edwin Hartl in *Die Press*e in January 1950 (PRESSE_1980_comm_Hartl) reacted to the critique and is the first in our data to defend the new ÖWB and its promoted 'more realistic' codification of language use. The week after, a verbal exchange between Hirschbold and Hornung on specific variants developed (PRESSE_1980_squib_Hirsch2, PRESSE_1980_lett4_Hornung, PRESSE_1980_squib_Hirsch3), besides further support of Hirschbold's critique by language cultivator Wolfgang Hardt-Stremayr (PRESSE_1980_lett5_Stremayr).

Hans Weigel once again entered the scene in March 1980 and published a commentary in the progressive national news magazine *profil*. As before in 1950, he emphasised the need for an endonormative codification but dismissed the ÖWB's thirty-fifth edition (PROFIL_1980_comm_Weigel). This was met with consent in letters to the editor (e.g. PROFIL_1980_lett1_Wolf), however finally countered by the ÖWB editors in May 1980 (PROFIL_1980_comm_Hornung).

The discourse in 1979/80 can be considered strongly personalised in the sense that it involves a smaller network of participants based in Vienna. Some, like Hornung, Hardt-Stremayr and Wolfram, were members of the extreme right-wing, purist[7] *Verein Muttersprache* under Mehl's chairmanship, and Hirschbold and Weigel had been occasional speakers

7 Purism here is used in a broader sense, i.e. '"free from foreign words" but also as "correct" in the sense of "grammatical" according to a norm established by a standard variety' (Pfalzgraf 2021: 61).

at association meetings (Pfalzgraf 2019: 57, 242, 277, 509). At this point, it remains unclear if personal animosities between discourse participants were an underlying motivation (Pfalzgraf 2019: 57). Afterwards, we may assume that the public debate cooled down, but critical discussions within the scientific community had only just begun (Ammon 1995: 133 for an extensive bibliography).

5.2 National variety purism and stereotypes of 'the Germans'

As expected (see Section 2), one of the central ideological assumptions in discourse is the national variety purism. Ammon classifies it as a special case of variety purism: It resembles the idea to keep a language free from 'foreign' influence (loan words) but, in our case, applied to a national standard variety: AG should be free from 'foreign' GG (Ammon 1995: 184).

The origins of this purism involve a political agenda: The initial intent of endonormative codification in Austria after World War II was understood as an additional effort of renationalisation, which at the time meant: denazification in Austria, supported by a broad political consensus between conservatives, social democrats and communists (Rathkolb 2015: 186) but pursued more rigorously by the political left: Founding fathers of the ÖWB such as the editors Ernst Pacolt and Otto Langbein had been persecuted by or had fought against the Nazi regime, which arguably explains their motivation to work on a presumably 'denazified' codex (Ammon 1995: 128). The case of the conservative Alfred Missong (not an editor, but a member of the commission) proves similar to those of Pacolt and Langbein.

In these circles, Austrian colloquial language (*Umgangssprache*) was not simply enregistered by contrasting it with GG but by implying that it is a non-fascist use of German. The original idea of (socialist) school superintendent Albert Krassnigg – together with his (conservative) co-superintendent Anton Simonic among the leading editors – was to mark GG variants as *reichsdeutsch* (r.d.) 'Reich German' to emphasise the anti-Nazism of the project (Ammon 1995: 182–3). Press announcements of an *Amtliches Österreichisches Wörterbuch* 'Official Austrian Dictionary' spoke of a 'reissue of the orthographic and lexical register [!]' in which 'all Non-Austrian will be disposed of and the Austrian vocabulary and language use will once again [!] be authoritative' (SN_1948_report_APA, OBC_1950_advert_öbv, my translation).

Perceived GG variants such as the phraseme *zur Auszahlung bringen* 'to pay so., to pay wages' (vs. AG *auszahlen*) or the complex compound *Lichtbildausweis* 'picture I.D'. (vs. AG loan word *Legitimation*) were considered Nazi-officialese as opposed to how Austrians actually spoke before the so-called Anschluss of 1938.

> The downfall of the [so-called] Greater Germanic Reich has left us with an almost Babylonic linguistic confusion, also in Austria. It was not the written [standard] German that was dictated to us from the outside after the occupation, but we had to adapt to Reich German officialese with Reich German officials, to military slang with Reich German soldiers, to party speak with party bigwigs. (NOS_1949_report3_editors, my translation[8])

On the other hand, stereotypes of Austrians as civilised language users as opposed to uncivilised, rude Germans led to the stigmatisation of a colloquial style, e.g. *abhauen* vs. *fortgehen* ('to bolt' vs. 'to leave'), thought to have been 'imported' to Austria (NOS_1949_report3_editors).

Neues Österreich was particularly supportive of Franz Wollmann, director of the Viennese Teachers' Academy, former member of the extreme right-wing purist *Allgemeine Deutsche Sprachverein* and later honorary member of the succeeding Viennese *Verein Muttersprache*. The purism here extended into national variety purism against 'Reich German' and 'un-Austrian' variants as well as 'Prussian mistakes':

8 For reasons of space, I will refrain from quoting the original German alongside the translations here and in the following. All translations of source texts are my own unless otherwise noted.

> Court Counsellor Dr Wollmann on his purifying language book: after the Prussians, their 'Káffee' [compared to AG *Kaffée*] shall be expatriated as well.[9] The new Austrian Dictionary bans northern German idioms into footnotes. [...] The new dictionary aims at liberating our Austrian colloquial language from Prussian expressions which were imported during the Nazi period and unfortunately have become common to some extent. [...] One particular characteristic peculiarity of Prussian jargon which is vigorously fought against in the new dictionary, is the wrong use of the auxiliary verb 'haben' ['to have', instead of *sein* 'to be'] in the formation of analytic tenses. We all are still able to hear phrases such as 'Wo hat das Ding gelegen?' [lit. 'Where has this thing lain?'] (NOS_1948_report2_editors)

Note that the enregisterment of AG means an ideological blow-up for Germany and an ideological shrinking for Austria: Not all regions in Germany use auxiliary *haben* for the formation of the present perfect of *legen* 'to lie'; especially in southern Germany, standard usage is the same as in Austria (Elspaß & Kleiner 2019: 162–3).

GG was virtually always represented by northern German, by supporters as well as opponents of the ÖWB: *Neues Österreich* (NOS_1949_report3_editors) contrasts Austriacisms with 'Berlinisms', 'northern German', or 'Prussicisms' (SN_1950_comm_Hala, FURCHE_1950_comm_Zie, TBZ_1950_comm_Wid) although the simplified equation of 'German = Nazi = Prussian' was criticised by some (BUEHNE_1950_comm_Egw, WP_1950_comm_Weigel2). Additionally, the enregisterment of GG as an artificial and stilted style is used even without implications of Nazi administration, e.g. GG *rechtschreiblich* 'orthographic' and *abcelich* 'alphabetic' vs. AG *orthographisch* and *alphabetisch* (BIF_1950_comm_N).

The anti-fascist agenda was soon toned down, and respective register pairs of AG and GG were deregistered (lost their previous social meaning). During the discourse in 1979/80, both supporters and critics of the ÖWB relied on the established purist assumptions nonetheless, e.g. GG *Kartoffel* ('potato') is perceived as an 'intruder' while AG *Erdapfel* 'fights for dear life' (Möcker 1980: 421, my translation). Moreover,

9 So-called Reich Germans were deported in 1945/46 in an act of 'Austrification', often tearing bi-national families apart (Rathkolb 2015: 38).

'the Prussian' surfaces again as a stereotype albeit overall less prominently (PRESSE_1979_lett2_Wladar, PRESSE_1980_comm_Hartl). Sometimes the implication of uncivilised GG was reiterated with other colloquial variants (e.g. PROFIL_1980_comm_Weigel) but more importantly, the arrogant German who speaks artificially but 'more correctly' and belittles the Austrians for their 'dialect' was fostered and today has become a stable stereotype (de Cillia et al. 2020: 103–5). Such feelings of inferiority seem to follow a tradition that goes as far back as the eighteenth century (Stevenson et al. 2018: 62), although they were also criticised at times (BUEHNE_1950_comm_Egw, BUEHNE_1950_comm_Tau, PRESSE_1980_lett3_Schachner). What is left today of the original political implications is a certain anti-Germanism present across all political camps and the purism towards GG influence with many Austrian speakers (Greth & Köllen 2016).

5.3 Federalism and democratic language planning

Two further and closely connected topics in the discourse on AG codification are federalism and democratic language planning. Austria is a federation of states in principle. In practice, however, Vienna and East Austria constitute the nation's political, cultural and economic centre. Ammon (1995: 134, 211–3) states that it is typical during standardisation processes that such centres enjoy certain privileges. The *Geltungsregion* 'area of social validity' is enlarged by codification, i.e. it is larger than the actual *Verwendungsregion* 'area of use' (Ammon 1995: 212–3). It seems understandable in theory that variants used in the capital are well known in the whole country and preferred by the Viennese majority of the ÖWB editors, not least when there is a lack of large corpora. Still, a principal issue in language policy, linguistic equality, remained unresolved at least until the 1990s, i.e. the consideration of West Austrian language use in Vorarlberg and Tyrol (Moser 1995). This imbalance is demonstrated by the fact that East Austrian variants may be marked as such in the ÖWB but no West Austrian equivalent is given or does not exist, e.g. in the case of East Austrian *stangeln* 'to punt' (Ammon 1995: 212–3). In other cases, East Austrian variants are left unmarked although they are unknown or considered unusual in West Austria, e.g. East Austrian *Fleischhauer/-hacker* vs. West Austrian *Metzger* 'butcher' (Moser 1995: 170–2).

The criticism of Viennese dominance is noticeable throughout our data and across all regions. Viennese features were sarcastically inserted into commentaries to achieve an alienation effect. Note that not all commentaries were against non-standard codification *per se*, which points to the central role regional non-standard dialects play in the construction of Austrian national identity:

> We [Tyroleans] do not want to have a dialect that is not at all ours imposed on us in a centralist way. After all, there are other dialects besides 'Viennese' which are also spoken in Austria. (TBZ_1950_comm_Wid)
>
> But there it is just with double zeal to dig out the official vocabulary! Who is not a mental *Armitschkerl* ['poor thing'] (or a Tyrolean who has never understood so much Viennese), will learn it with time already – and why should one not do something for once and *abstrudln* ['work hard'] for the rescue of the national culture? (SN_1950_comm_Hala)
>
> Furthermore, it is an imposition on the Austrian federal states that expressions of Viennese jargon should have universal value for the whole of Austria. (PRESSE_1979_lett1_Wolfram)
>
> The new ÖWB is [...] Viennese-biased. It contains a plethora of regional words and sayings, and this is all very well. Yet, no adequate number of words and sayings from other federal states matches the surplus of Viennese. [...] we Viennese once again make ourselves unpopular with the other states. One must hope that the new ÖWB is rejected by the regional superintendence of education on justified grounds of federalist complaints. (PROFIL_1980_comm_Weigel)

In addition to a federalist agenda, a general criticism of an undemocratic language policy comes into play (Wiesinger 1980: 367). The political communication between the Ministry of Education and the public did not run smoothly between 1946 and 1951: it was noted that the Ministry did not communicate with critical media (WP_1950_comm_Weigel2, SN_1950_squib_Peris1). Although Secretary Stur spoke of more than 2,400 filled-out questionnaires which were reviewed and given to a dictionary commission of experts under his supervision, he conceded that a smaller separate commission, a committee of only six persons[10], was the actual authority (DKV_1950_report_editors, WP_1950_report_APA, ETB_1950_report_editors). It was unclear what the working mode in the commission looked like and how exactly its members had been recruited. Robert Stein, a Viennese publishing manager, lamented:

> When Dr Missong speaks of publishers who are members of this commission it has to be established that these men only spoke as private individuals within this commission. […]. Eventually, we would publicly like to urge the administration of the Ministry of Education to get in touch with the responsible professional associations in matters which concern the publishing industry and the book trade in the future. (OBC_1950_comm_Stein)

As a sort of concession to the protests, the attribute *amtlich* 'official' in the dictionary title was dropped (Ammon 1995: 131–2). The refusal of an official authorisation of the ÖWB may also have been motivated by negative experiences with the temporary official renaming of German as a school subject to *Unterrichtssprache* 'language of instruction' 1945–1951 (Rathkolb 2015: 37). Hans Weigel and the regional newspaper *Salzburger Nachrichten* regularly jabbed at this regulation and Minister Hurdes as its perceived personification, also using *Hurdestanisch* 'Hurdestanian' as a derogative allusion to Minister Hurdes and the ÖWB's register (e.g. WP_1949_comm_Weigel, SN_1951_comm_Hille, cf. Fig. 11.1):

The enregisterment of AG as a ridiculous artificial product was, however, short-lived. By 1979, the social meaning of AG appears to have been consolidated and, to a certain degree, based on positive linguistic stereotypes. The discourse participants did not openly doubt the existence of an Austrian standard variety and the usefulness of an endonormative codification anymore. At the same time, they denied the plausibility of an Austrian language separate from the German language, which conforms to Ammon's observations (Ammon 1995: 132–3). The critics of 1979/80 focused more on the presumed attempts of destandardisation by the ÖWB. A closer inspection shows that the corresponding issues were

10 Presumably Stur himself, Pacolt, Langbein, Krassnigg, Simonic, and Wollmann.

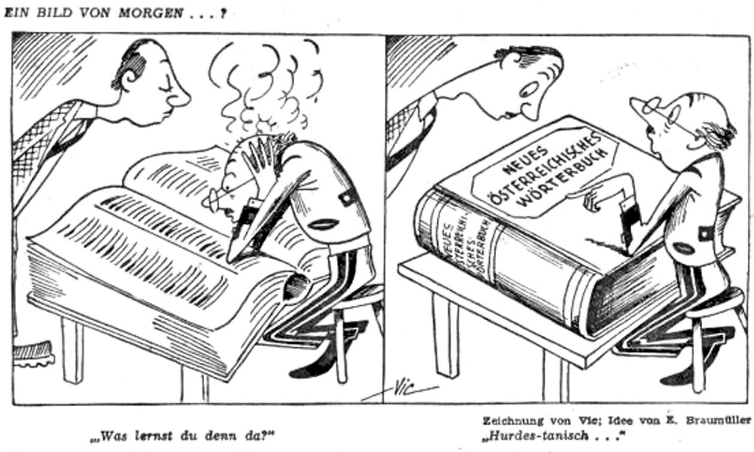


Figure 11.1 *A picture of tomorrow …*? '*Gee, what are you studying*?' (New Austrian Dictionary) – '*Hurdes-tanian …*' (SN_1950_squib_Peris2).

already at play 1946–1951. I will therefore turn to this continuity between 1979/80 and 1946–1951 in more detail.

5.4 Codification of dialect and 'colloquial' standards of usage: Destandardisation?

The broad-scale inclusion of features that were perceived as non-standard by writers can be termed an extreme form of anti-German separatism which was driven partly by anti-Nazism and partly by traditional Austrian nationalism (Ammon 1995: 52, 130–1, 186–8).

In the spirit of the former, the editors of the original ÖWB had initially intended to include many colloquial and non-standard variants to maximise the linguistic distance from Germany (Wiesinger 2014). One of the arguments put forth in favour of this practice was that the exonormative Duden, which was being used in Austria, codified non-standard northern German so that GG dominated language use on several varietal levels. For reasons of linguistic equality, opinions were expressed suggesting that Austria should codify national non-standard variants even if they belonged to a shared southern German usage and led to an ideological shrinking:

> […] for it was indeed the case that the Duden, this Bible of all correctors, very often closed its doors to the organic results of development and innovations coming from southern Germany, Switzerland and Austria, while the linguistic traits originating in Prussian colloquial language, which often did not grow organically at all and were often valid only locally, found welcome acceptance. This led to an official over-Prussification of the German language, to which, of course, the southern regions […] did not submit. […] for to take only one quite obvious example, for instance the diminutive syllable, the Swabian *-le* or *-li*, the colourless and general *-chen* or *-lein* and the southern German *-erl* will differ very considerably from one another and each of them will contain sentimental values which are lost when they are all normalised and equalised to the common, that is to say, the least expressive form. A *Maidli*, a *Mädele*, a *Mäderl* are something different from a *Mädchen* ['girl']. (WZ_1950_comm_Roll)
>
> How often do we find expressions in German dictionaries that originate from the North German dialect. These are to be eradicated in our parts and purely Austrian expressions are to be included. Surely that is right. (FURCHE_1950_lett2_anonym)

The idea of non-standard codification was revived in 1979/80, but this time the editors stressed the argument that it reflected colloquial standards common among Austrian speakers as, for example, the editor Maria Hornung argued (PRESSE_1980_lett4_Hornung).

In both phases, 1946–1951 and 1979/80, the idea received fundamental criticism. The aversion towards Viennese dominance and official language planning (see Section 5.3 for details) mixed with fears of a possible destandardisation and *Verhunzung* 'adulteration' of German:

> […] it is finally time to state quite decisively that it is a sin against the [regional non-standard] dialect if it is violated in its autonomy by official decisions. If dialect is elevated to a written language, even if it is only a few expressions, it is no longer a dialect. […] Only great poets like our Grillparzer are authorised to vernacularise certain dialect expressions in written German. (TBZ_1950_comm_Wid)
>
> The 'incorporation' of clearly dialectal expressions or idioms into an official dictionary, i.e. the mixing of written German and dialect in an official publication,

> would not only lead to an alarming lowering of the standard of Austria's written language; the result would be an adulteration, even a real adulterating of the nobler German written language and would be quite capable of exposing Austria to ridicule and endangering the reputation of its culture. (FURCHE_1950_comm_Zie)
>
> Nothing against dialect and colloquial language; but by carelessly blurring the boundaries to standard language, our 'ÖWB' could quickly become a 'SWB' (= *SCHLUDERISCHES WÖRTERBUCH*) [Sloppy Dictionary]. (PRESSE_1980_squib_Hirsch3)

Notably, the enregisterment of 'colloquial' AG features in both periods concerned not only lexical but also grammatical variants. While the media coverage in 1946–1951 focused on the umlaut plural (e.g. AG *Wägen* vs. GG *Wagen* 'cars') and diminutives (AG *-l* vs. GG *-chen*), in 1979/80 a more diverse range of variants were presented and rejected (Wiesinger 1980: 392–3), among them 'wrong' neuter with *das Monat* 'the month' (vs. 'correct' masculine *der Monat*), preposition *trotz* + dative (vs. genitive), or the use of *am* (vs. *auf dem* 'on the, at the') – all of which nowadays are considered standard variants in use (cf. Variantengrammatik 2018).

A few more cosmopolitan commentators in both periods also saw the danger that codifying 'colloquial' variants could corroborate negative stereotypes about Austrians and their supposedly inferior German. They complained that the dictionary made Austria the laughingstock among Europe's civilisations.

> The large number of dialect expressions, for example, is by no means an enrichment, but rather appears to be the product of an idiosyncrasy that has been clinging to the country since 1945 [...]; an open-minded, lively country fed by many cultural circles seems to be developing more and more into a nation of people in traditional costume who do not care about living to the rhythm of the world, which has always been a characteristic of genuine Austrians. (BUEHNE_1950_comm_Egw)
>
> The heirs of a great Empire with many nationalities are ill-suited to lapse into petty cultural and linguistic (see Austrian Dictionary) narrow-minded particularism instead of pan-European thinking. (PRESSE_1980_lett3_Schachner)

In sum, the incorporation of perceived non-standard features in a standard codex was successful in principle but not on a broader scale.

Constructions of national identity via nonstandardness were considered, but in both phases, 1946–1951 and 1979/80, the number of corresponding variants was eventually reduced according to strict prescriptivist and conservative views of standard language. It cannot be decided here if a minor inclusion of perceived non-standard variants speaks for destandardisation. At any rate, some grammatical variants have been standardised in use rather than by codification in the ÖWB, which gives an impression of the limits of prescriptive codices in practice.

5.5 Endonormative orthography and economic concerns

From 1946–1951, the negative reactions towards a liberal inclusion of perceived non-standard features intermingled with the emotional criticism of a proposed spelling 'reform', in fact mostly regarding additional phonetic spelling variants of French loan words, e.g. *Blamasch* 'disgrace', *Büro* 'office', *Frisör* 'barber', *Oranschen* 'oranges', *Scharm* 'charm'. Some publishers feared that an endonormative orthography could result in economic troubles, which the ÖWB commission dismissed as exaggerations:

> Are the editors of the dictionary aware that the entire graphic arts industry and the entire publishing industry in Austria would be confronted with insoluble problems if, in short, the Duden were to be officially abolished in Austria? Should books printed in Austria perhaps in future be printed in an edition with Austrian spelling and in an edition with that spelling which is considered correct not only in Germany but also in Switzerland, among all Germans abroad and everywhere else in the world where German is spoken? (FURCHE_1950_lett1_anonym)
>
> However, one will easily be able to dispel the concerns of this interest group by pointing out that the few specific Austriacisms in question here can be understood in the entire German-speaking area, especially since they are also included in the Duden with the note 'ö.R'. (Austrian rulebook). (FURCHE_1950_comm_Miss)

The discussion in 1979/80 also featured some criticism of orthographic codification (e.g. Möcker 1980), but in contrast to earlier did not link them to economic concerns anymore. Instead, grammatical and lexical variants were highlighted as unique selling propositions:

> Of course, German editors should be given such an 'Austrian Dictionary' as a guide, so that Austrian authors who are printed outside can still be recognised as such. (PRESSE_1980_comm_Hartl)
>
> We say that we 'are' stood and sat, not 'have' stood and sat. *Sessel* is something different here ['chair'] than outside Austria ['armchair'] [...]. Yes, we need such a book, partly because non-Austrians want to look it up when they read Nestroy or Doderer, for example, or because we want to look it up when we want to communicate with non-Austrians – especially in the culinary lexical field. (PROFIL_1980_comm_Weigel)
>
> In this context, the Austrian food industry must also be criticised for producing packages solely for export with domestic German product names. In view of the large quantities, no one can argue that the few letters for the additional Austrian name would be too expensive. If we were to find *Tomatenmark* = *Paradeismark* ['tomato paste'], *Kartoffelpüree* = *Erdäpfelpüree* ['mashed potatoes']; on compote tins, jam jars and juice bottles *Aprikosen* = *Marillen* ['apricots'], *Johannisbeeren* = *Ribisel* ['currant'], ... etc., ... , one would on the one hand contribute this to the stabilisation of the Austrian terms in Austria and on the other hand be able to use the incentive of 'exotic' product terms for export. (An analogous approach on the menus of our restaurants would be desirable). (Möcker 1980: 421)

Again, the enregisterment featured both the ideological blow-up of East Austrian variants such as *Sessel* 'chair' (vs. Western Austrian standard *Stuhl*) and *Paradeiser* 'tomato' (vs. West Austrian standard *Tomate*) and an ideological shrinking of the southern standard auxiliary verb *sein* for present perfect of *stehen* 'to stand' and *sitzen* 'to sit' (vs. northernly *haben*) to Austrian dimensions.

In the end, the respective ÖWB editions of 1951 and 1985 abandoned both the idea of liberal spellings and the idea of a broad incorporation of 'colloquial' features without further comment (Ammon 1995: 136).

5.6 Socio-cultural implications

Throughout the historical evidence, an underlying cultural conservatism of the ÖWB's opponents is noticeable.

For instance, a problem inherent to the Ministry's communication strategy 1946–1951 is *Umgangssprache* 'colloquial German', a notoriously vague term in German: It can refer to variants on the verge to standard, or *de facto* standards of usage. The newspaper *Neues Österreich*, in favour of the original progressive and anti-fascist concept of the ÖWB (and obviously well-informed by the original editors), used the term in this meaning (e.g. NOS_1949_report3_editors). This stage represents an early example of the solidarity function of AG and a puzzle piece in the so-called *Opfermythos* 'victim myth', i.e. the myth of Austria as the 'first victim' of Nazi Germany's occupation politics, a sovereign country which had unwillingly been occupied and had had no part in Nazi crimes. The myth was authorised as a doctrine by the Austrian government after the war to reconcile left-wing and right-wing parties with each other in an all-party government and integrate Austrian ex-Nazis into a democratic society, and the myth remained more or less unquestioned until the 1980s (Art 2005: 104).

However, when anti-communism superimposed anti-fascism (Rathkolb 2015: 33–5) after the communist KPÖ withdrew from the government in 1947, many responded negatively to a codification of *Umgangssprache* (e.g. letters to the editor in the *Weltpresse*, WP_1950_lett1_Kunzl to WP_1950_lett13_Bachler), referring to another common meaning in Austria, i.e. regional and local vernaculars. Conservative newspapers outside Vienna, like the *Salzburger Nachrichten* (SN_1950_comm_Hille) and the *Oberösterreichische Nachrichten* (OON_1950_comm_Fl), exploited this aversion to vernacular codification in order to defame Austrian cultural independence as a form of 'communist' separatism (in reference to the majority of ÖWB editors leaning towards the political left). *Umgangssprache* remained a disputable concept without consensual enregisterment.

Although the anti-communist/-socialist implications appear less relevant in later years, the discourse on the 1979 ÖWB and its approach of colloquial standards reveals similar tendencies when the teacher Hans-Jürgen Weitschacher (1981) cried out *Hilfe, die Roten kommen!* 'Help, here come the reds!'. Although he explained that the colour should indicate the new red book cover of the ÖWB rather than socialists or social democrats (Weitschacher 1981: 41), he must have been aware of his insinuation during times of the Cold War.

Likewise, the early discussion in 1979/80 was fixated on the unmarked codification of sexualised perceived non-standard words like *lepschi* ‘erotic adventure’ (PRESSE_1979_report_editors, PRESSE_1979_lett1_Wolfram) and vulgar terms like *scheißen* ‘to shit’ (Weitschacher 1983: 44) as part of an exaggerated progressive agenda and sexual liberation (Dressler & Wodak 1983: 255). The accusations went so far as to condemn the inclusion of terms for body parts in a school dictionary as a sign of moral decay (PRESSE_1979_report_editors, PRESSE_1980_comm_Hartl), but ceased when critics pointed out its inherent prude bigotry (PRESSE_1980_comm_Hartl).

Reiffenstein (1983: 19) assumed that the inclusion of considerably more and stylistically unmarked features which he perceived as non-standard was the consequence of a progressive socio-liberal policy towards lower classes (Ammon 1995: 134). This alludes to the absolute dominance of the Social Democratic Party of Austria (SPÖ) during the 1970s and early 1980s which brought about many changes in the educational system such as the democratisation of schools and universities (Vocelka 2019: 119). Interestingly, the criticism mirrored the elitist scepticism towards the political and socio-cultural situation in Germany at the same time: The inclusion of colloquial features in the Duden and a more usage-based approach in codification during a politically dominant social liberalism (coalition of social democrats and liberals in Germany) were met with a bias towards the language of ‘third-rate authors, the yellow press [*Bild-Zeitung* in Germany, *Kronen Zeitung* in Austria, K.N.] and certain weekly magazines [*Der Spiegel* in Germany, *profil* in Austria, both perceived progressive, K.N.]’ (Möcker 1980: 443, my translation). In Austria, overt elitist critics like Hirschbold stated that it should not only depend on ‘by how many so-called language users these forms are used’ but also ‘by whom’ (PRESSE_1980_squib_Hirsch3, my translation), and Weigel complained about ‘philologists stylising themselves’ as overly progressive and ‘close to the people’ (PROFIL_1980_comm_Weigel, my translation). Reiffenstein (1983: 19–20, 24) on the other hand agreed with a more liberal take on codification, but still thought tolerance towards diatopic standard variation and stylistic variation should not be conflated. Similarly, Dressler & Wodak claimed that the inclusion of diatopic and stylistic variation could lead to more tolerance and reflexivity; hence, the problem was the ÖWB's poor execution, not its liberal ideology (Dressler & Wodak 1983: 258).

6. Summary

AG emerged as a national register after World War II and, in the case of the ÖWB, was initially enregistered as a strongly anti-fascist language. In that sense, many features which indicated this social meaning appear irrelevant in 1979/80; national stereotypes of cuisine and corresponding lexemes, however, stabilised, as did the enregisterment of a few grammatical variants. Diachronically stable register features in our data are *Knödel* 'dumpling', *Karfiol* 'cauliflower', *Ribisel* 'currant', *Paradeiser* 'tomato', *Obers* 'cream', *Erdapfel* 'potato', *Sessel* 'chair', the *-el-/-l-*diminutive, the auxiliary *sein* for part. perf. of *stehen* 'to stand' and some cases of gender variation. The lexical variants are exactly the 'national' features often mentioned in ethnographic and interactional studies (de Cillia et al. 2020: 101–3), they have become national linguistic stereotypes (Stevenson et al. 2018: 61–3) and enregistered AG features.

While the register AG had been firmly established by 1979/80, the same ideological issues of language planning as in earlier times (excluding denazification) lingered on in public discussions. National stereotypes of AG and GG were spread and contributed to the codification as well as contemporary social meanings of the register AG. The purism against GG influence in 1946–1951 was mostly motivated by anti-Nazism, and in 1979/80 it was legitimised as means of linguistic justice among German-speaking nations. Further research is needed to determine to what extent the codification (or the discourse on it) has led to a nationalisation of standards of usage in Austria (Ammon 1995: 189). In any case, the endonormative codification had transformed into an important symbol of national identity by 1979/80, at least among elites.

The reactions towards the specific ÖWB edition depended on various aspects and not only political affiliation. The evidence neither implies a cohesive political left, ranging from communist propagandists to social democratic magistrate officials, nor a unified pan-Germanic opposition on the other side of the political spectrum. Rather, the criticism of the ÖWB was partly motivated by the exclusion of individuals from the Academy's Commission and the ÖWB's Editorial Board, and partly by socio-cultural criteria such as economic concerns, European cosmopolitanism, or cultural elitism. Inner-Austrian equality also appears as a motivation for criticism in both periods, but the political issue of federalism was even more pressing during the rebuilding of a nation-state in 1946–1951. Support for the ÖWB and endonormative codification must be differentiated likewise: a welcoming stance towards a more 'usage-based' codification could still reject the codification of perceived non-standard or 'colloquial' features – elitist purism is a strong reoccurring theme in our data.

7. Outlook

This analysis could only highlight selected aspects of the historical social relevance of AG and can only be a first step towards a broad historical CDA with detailed case studies and the formulation of a comprehensive critique (Reisigl 2017: 54–5). Strikingly, the position of AG as a central part of national identity seems to have grown even stronger after Austria's accession to the EU in 1995, and non-standard seems to have become even more central to the register (de Cillia et al. 2020: 112), rendering the issue of inconsistent codification of perceived non-standard features even more relevant: where do sociolinguistic stereotypes counter-productive to linguistic justice end and where does democratically justified, non-elitist standardisation begin, especially without retreating to a bias towards East Austrian usage?

Future analyses could thus turn towards this more recent period during the 1990s and examine if and how a change in the political landscape altered the social meaning of AG: did the rise of the extreme-right FPÖ and its 'move away from overt neo-fascist discourse' towards Austrian nationalism (Wodak 2015: 25) allow for a reindexicalisation of AG? How is the increasing number of German migrants in Austria dealt with in metalinguistic discourse and, more specifically, in codification?

In addition, concepts such as dominance and asymmetry (Clyne 1995: 21–2) do not suffice to capture the historically prevalent linguistic ideologies in Austria and need refinement via discursive evidence. Traditional feelings of linguistic inferiority towards GG can have a strong influence on codification and usage in Austria, but they alone do not explain satisfactorily how the complex discourse on AG ultimately created the current double standards.

Bibliography

Agha, Asif. 2007. *Language and Social Relations*. Cambridge: Cambridge University Press.

Auer, Peter. 2013. Enregistering pluricentric German. In Augusto Soares da Silva (ed.). *Pluricentricity. Language Variation and Sociocognitive Dimensions*. Berlin, Boston: de Gruyter. 19–48.

Auer, Peter. 2021. Reflections on linguistic pluricentricity. *Sociolinguistica* 35(1), 29–47.

Auer, Peter & Helmut Spiekermann. 2011. Demotisation of the standard variety or destandardisation? The changing status of German in late modernity (with special reference to southwestern Germany). In Tore Kristiansen & Nikolas Coupland (eds). *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus Press. 161–76.

Ammon, Ulrich. 1995. *Die deutsche Sprache in Deutschland, Österreich und der Schweiz. Das Problem der nationalen Varietäten*. Berlin, New York: de Gruyter.

Art, David. 2005. *The Politics of the Nazi Past in Germany and Austria*. Cambridge: Cambridge University Press.

Ayres-Bennett, Wendy. 2021. Modelling language standardization. In Wendy Ayres-Bennett & John Bellamy (eds). *The Cambridge Handbook of Language Standardization*. Cambridge: Cambridge University Press. 27–64.

Cillia, Rudolf de, Ruth Wodak, Markus Rheindorf & Sabine Lehner. 2020. *Österreichische Identitäten im Wandel*. Wiesbaden: Springer VS.

Clyne, Michael. 1985. Reactions to the 1979 Austrian Dictionary – Conservatism or Cultural Cringe? In John E. Clark (ed.). *The Cultivated Australian. Festschrift in Honour of Arthur Delbridge*. Hamburg: Buske. 263–71.

Clyne. Michael. 1995. *The German Language in a Changing Europe*. Cambridge: Cambridge University Press.

Costa, James, Haley Jean De Korne & Pia Lane. 2017. Standardising minority languages: Reinventing peripheral languages in the 21st century. In Pia Lane, James Costa & Haley De Korne (eds). *Standardizing Minority Languages: Competing Ideologies of Authority and Authenticity in the Global Periphery*. London: Routledge. 1–23.

Dressler, Wolfgang U. & Ruth Wodak. 1983. Soziolinguistische Überlegungen zum 'Österreichischen Wörterbuch'. In Maurizio Dardano, Wolfgang Dressler & Gerhard Held (eds). *Parallela. Akten des 2. österreichisch-italienischen Linguistentreffens*. Tübingen: Narr. 247–60.

Ebner, Jakob. 1969. *Wie sagt man in Österreich? Wörterbuch der österreichischen Besonderheiten*. Mannheim, Wien, Zürich: Dudenverlag.

Elspaß, Stephan & Stefan Kleiner. 2019. Forschungsergebnisse zur arealen Variation im Standarddeutschen. In Joachim Herrgen & Jürgen Erich Schmidt (eds). *Language and Space. An International Handbook of Linguistic Variation*. Vol. 4: *German*. Berlin, Boston: de Gruyter. 159–84.

Fairclough, Norman. 1995. *Critical Discourse Analysis*. London: Longman.

Gal, Susan. 2019. Making registers in politics: Circulation and ideologies of linguistic authority. *Journal of Sociolinguistics* 23(5), 450–66.

Glauninger, Manfred. 2015. (Standard-)Deutsch in Österreich im Kontext des gesamtdeutschen Sprachraums. Perspektiven einer funktional dimensionierten Sprachvariationstheorie. In Alexandra N. Lenz & Manfred M. Glauninger (eds). *Standarddeutsch im 21. Jahrhundert – Theoretische und empirische Ansätze mit einem Fokus auf Österreich*. Göttingen: V&R Vienna University Press. 11–57.

Greth, Julia & Thomas Köllen. 2016. Perceived Anti-Germanism in Austria. *Studies in Ethnicity and Nationalism* 16(1), 40–62.

Hall-Lew, Lauren, Emma Moore & Robert J. Podesva. 2021. Social meaning and linguistic variation: Theoretical foundations. In Lauren Hall-Lew, Emma Moore & Robert J. Podesva (eds). *Social Meaning and Linguistic Variation. Theorizing the Third Wave*. Cambridge: Cambridge University Press. 1–23.

Huesmann, Anette. 1997. Standardisierung und Destandardisierung am Beispiel des österreichischen Wörterbuchs. In Klaus J. Mattheier & Edgar Radtke (eds). *Standardisierung und Destandardisierung europäischer Nationalsprachen*. Frankfurt am Main: Lang. 193–9.

Johnstone, Barbara, Jennifer Andrus & Andrew E. Danielson. 2006. Mobility, indexicality, and the enregisterment of 'Pittsburghese'. *Journal of English Linguistics* 34(2), 77–104.

Langer, Nils. 2021. Review of: Stefan Dollinger. 2019. The pluricentricity debate. On Austrian German and other Germanic standard varieties (Routledge Focus). Abingdon: Routledge. *Zeitschrift für Rezensionen zur germanistischen Sprachwissenschaft* 13(1–2), 1–8.

Kroskrity, Paul V. 2001. Identity. In Alessandro Duranti (eds). *Key Terms in Language and Culture*. Malden: Blackwell. 106–9.

Möcker, Hermann. 1980. Fahren sie schon Rad oder fahren Sie noch rad? Grammatische und orthographische Beobachtungen am neuen Österreichischen Wörterbuch (35. Aufl.). *Österreich in Geschichte und Literatur* 24, 416–45.

Moore, Emma & Podesva, Robert P. 2009. Style, indexicality, and the social meaning of tag questions. *Language in Society* 38(4), 447–85.

Moser, Hans. 1995. Westösterreich und die Kodifizierung des 'österreichischen Deutsch'. In Rudolf Muhr, Richard Schrodt & Peter Wiesinger (eds). *Österreichisches Deutsch. Linguistische, sozialpsychologische und sprachpolitische Aspekte einer nationalen Variante des Deutschen*. Vienna: Hölder-Pichler-Tempsky. 166–77.

Oakes, Leigh. 2021. Pluricentric linguistic justice: A new ethics-based approach to pluricentricity in French and other languages. *Sociolinguistica* 35(1), 49–71.

ÖWB 1951 = *Österreichisches Wörterbuch*. 1951. Vienna: Österreichischer Bundesverlag.

ÖWB 1979 = *Österreichisches Wörterbuch*. 1979. 35th edn. Vienna: Österreichischer Bundesverlag.

ÖWB 1985 = *Österreichisches Wörterbuch*. 1985. 36th edn. Vienna: Österreichischer Bundesverlag.

Pfalzgraf, Falco. 2019. *Der Verein ‚Muttersprache' Wien unter Vorsitz von Karl Tekusch und Erwin Mehl (*1949–1984). Heidelberg: Winter.

Pfalzgraf, Falco. 2021. Karl Tekusch (1890–1977), his concept of Sprachechtheit, and the purism movement in Austria. *Historiographia Linguistica* 48(1), 60–82.

Rathkolb, Oliver. 2015. *Die paradoxe Republik. Österreich 1945 bis 2015*. Vienna: Zsolnay.

Reiffenstein, Ingo. 1983. Deutsch in Österreich. In Ingo Reiffenstein, Heinz Rupp, Peter von Polenz & Gustav Korlén (eds). *Tendenzen, Formen und Strukturen der deutschen Standardsprache nach 1945*. Marburg: Elwert. 15–27.

Reiffenstein, Ingo. 1995. Das Österreichische Wörterbuch. Zielsetzungen und Funktionen. In Rudolf Muhr, Richard Schrodt & Peter Wiesinger (eds).

Österreichisches Deutsch. Linguistische, sozialpsychologische und sprachpolitische Aspekte einer nationalen Variante des Deutschen. Vienna: Hölder-Pichler-Tempsky. 158–65.

Reisigl, Martin. 2013. Critical discourse Analysis. In Robert Bayley, Richard Cameron, and Ceil Lucas (eds). *The Oxford Handbook of Sociolinguistics*. Oxford: Oxford University Press. 67–90.

Reisigl, Martin & Wodak, Ruth. 2009. The discourse-historical approach. In Ruth Wodak & Michael Meyer (eds). *Methods of Critical Discourse Analysis*. 2nd edn. London: Sage. 87–121.

Schmidlin, Regula. 2011. *Die Vielfalt des Deutschen. Gebrauch, Einschätzung und Kodifizierung einer plurizentrischen Sprache*. Berlin, New York: de Gruyter.

Silverstein, Michael. 1979. Language structure and language ideology. In Paul R. Clyne, William F. Hanks & Carol L. Hofbauer (eds). *The Elements: A Parasession on Linguistic Units and Levels*. Chicago: Chicago Linguistic Society. 193–247.

Stadt Wien (online) = <https://www.geschichtewiki.wien.gv.at/Wien_Geschichte_Wiki> (accessed 23 March 2023).

Stevenson, Patrick, Kristine Horner, Nils Langer & Getrud Reershemius. 2018. *The German-speaking World. A Practical Introduction to Sociolinguistic Issues*. 2nd edn. Abingdon: Routledge.

Sutter, Patrizia. 2017. *Diatopische Variation im Wörterbuch. Theorie und Praxis*. Berlin, Boston: de Gruyter.

Variantengrammatik 2018 = *Variantengrammatik des Standarddeutschen. 2018. Ein Online-Nachschlagewerk*. Verfasst von einem Autorenteam unter der Leitung von Christa Dürscheid, Stephan Elspaß und Arne Ziegler. Open-access publication. <http://mediawiki.ids-mannheim.de/VarGra/index.php/Start> (accessed 1 July 2024).

Vocelka, Karl. 2019. *Österreichische Geschichte*. 5th edn. Munich: C. H. Beck.

Wiesinger, Peter. 1980. Zum Wortschatz im ‚Österreichischen Wörterbuch'. *Österreich in Geschichte und Literatur* 24, 367–96.

Wiesinger, Peter. 2014. *Das Österreichische Deutsch in Geschichte und Gegenwart, 3rd revised and extended edn*. Vienna: Lit-Verlag.

Weitschacher, Hans-Jürgen. 1981. Hilfe, die Roten kommen! Keineswegs vollständige Überlegungen zur Neuauflage des Österreichischen Wörterbuches. In Werner Kristoferitsch (ed.). *Festschrift 75 Jahre Bundesgymnasium Graz Pestalozzi*. Graz: Bundesgymnasium Graz Pestalozzi. 41–5

Wodak, Ruth. 1994. Wir sind nicht Duden-Land. *Wiener Journal* (June 1994), 26–7.

Wodak, Ruth. 2015. Anything goes! The Haiderization of Euope. In Ruth Wodak, Majid KhosraviNik & Brigitte Mral (eds). *Right-wing Populism in Europe: Politics and Discourse*. London: Bloomsbury. 23–38.

Appendix

Sources in alphabetical order; abbreviations follow the pattern: short title_year_short title of text type (short name of) author; in case an author is featured with several contributions, this is indicated with numbers.

abbreviation	source	text type	author	date
BIF_1950_ comm_N	Berichte und Informationen	commentary	Dr N.	03.05.1950
BUEHNE_1950_ comm_Egw	Wiener Bühne	commentary	Erik G. Wickenburg	January 1950
BUEHNE_1950_ comm_Tau	Wiener Bühne	commentary	Franz Taucher	January 1950
DKV_1950_ comm_ra	Das Kleine Volksblatt	commentary	r.a.	01.02.1950
DKV_1950_ comm_Bock	Das Kleine Volksblatt	commentary	Dr Fritz Bock	14.02.1950
DKV_1950_ report_editors	Das Kleine Volksblatt	report	editorial staff	24.02.1950
ETB_1950_ report_editors	Erlaftal-Bote	report	editorial staff	07.01.1950
FURCHE_1946_ comm_Zel	Die Furche	commentary	Hans Zelkinger (Hans Brečka)	02.11.1946
FURCHE_1947_ comm_Zel	Die Furche	commentary	Hans Zelkinger (Hans Brečka)	27.09.1947
FURCHE_1949_ comm_Frei	Die Furche	commentary	Dr Franz Freitag	01.01.1949
FURCHE_1950_ comm_Zie	Die Furche	commentary	Dr Hugo Zienert	28.01.1950

(*Continued*)

abbreviation	source	text type	author	date
FURCHE_1950_lett1_anonym	Die Furche	letter to the editor	anonymous publisher	04.02.1950
FURCHE_1950_lett2_anonym	Die Furche	letter to the editor	anonymous teacher	04.02.1950
FURCHE_1950_comm_Miss	Die Furche	commentary	Dr Alfred Missong	18.02.1950
NOS_1948_report_APA	Neues Österreich	report	APA correspondence	02.03.1948
NOS_1948_report1_editors	Neues Österreich	report	editorial staff	02.03.1948
NOS_1948_report2_editors	Neues Österreich	report	editorial staff	05.03.1948
NOS_1949_report3_editors	Neues Österreich	report	editorial staff	25.12.1949
NOS_1950_report4_editors	Neues Österreich	report	editorial staff	03.03.1950
NOS_1950_report5_editors	Neues Österreich	report	editorial staff	23.06.1950
NZ_1950_report_editors	Neue Zeit	report	editorial staff	05.01.1950
OBC_1950_comm_Stein	Anzeiger des österr. Buch=, Kunst= und Musikalienhandels	commentary	Dr Robert Stein	01.03.1950
OBC_1948_advert_öbv	Anzeiger des österr. Buch=, Kunst= und Musikalienhandels	advertisement	Österreichischer Bundesverlag Wien	15.03.1948
OON_1950_comm_Fl	Oberösterreichische Nachrichten	commentary	F. L.	06.02.1950
OON_1950_comm_editors	Oberösterreichische Nachrichten	commentary	editorial staff	22.07.1950
PRESSE_1979_squib_Hirsch1	Die Presse	squib	Karl Hirschbold	24./25.11.1979
PRESSE_1979_report_editors	Die Presse	report	editorial staff	01./02.12.1979

(*Continued*)

abbreviation	source	text type	author	date
PRESSE_1979_ lett1_Wolfram	Die Presse	letter to the editor	Prof Dr Fritz Wolfram	22./23.12.1979
PRESSE_1979_ lett2_Wladar	Die Presse	letter to the editor	Prof J. O. Wladar	22./23.12.1979
PRESSE_1980_ comm_Hartl	Die Presse	commentary	Edwin Hartl	05./06.01.1980
PRESSE_1980_ squib_Hirsch2	Die Presse	squib	Karl Hirschbold	12/13..01.1980
PRESSE_1980_ lett3_Schachner	Die Presse	letter to the editor	Otto Schachner	02./03.02.1980
PRESSE_1980_ lett4_Hornung	Die Presse	letter to the editor	Univ.-Prof Dr Maria Hornung	09./10.02.1980
PRESSE_1980_ lett5_Stremayr	Die Presse	letter to the editor	Wolfgang Hardt-Stremayr	09./10.02.1980
PRESSE_1980_ squib_Hirsch3	Die Presse	squib	Karl Hirschbold	08./09.03.1980
PROFIL_1980_ comm_Weigel	profil	commentary	Hans Weigel	10.03.1980
PROFIL_1980_ lett1_Friedrich	profil	letter to the editor	Dr Friedrich Friedrich	24.03.1980
PROFIL_1980_ lett2_Wolf	profil	letter to the editor	Dr Gunter Wolf	31.03.1980
PROFIL_1980_ comm_Hornung	profil	commentary	Prof Dr Maria Hornung, Prof Ernst Pacolt, Prof Erich Benedikt	19.05.1980
SN_1948_report_ APA	Salzburger Nachrichten	report	APA correspondence	03.03.1948
SN_1950_comm_ Hille	Salzburger Nachrichten	commentary	B. K. Hillebrand (-Skrehunetz)	13.01.1950
SN_1950_comm_ Hala	Salzburger Nachrichten	commentary	Halawachl (pseudonym)	27.01.1950

(*Continued*)

abbreviation	source	text type	author	date
SN_1950_squib_Peris1	Salzburger Nachrichten	squib	Periskop staff	04.02.1950
SN_1950_squib_Peris2	Salzburger Nachrichten	squib	Periskop staff	11.02.1950
SN_1950_squib_Peris3	Salzburger Nachrichten	squib	Periskop staff	04.11.1950
SN_1951_squib_Hille	Salzburger Nachrichten	squib	B. K. Hillebrand (-Skrehunetz)	19.03.1951
TBZ_1950_comm_Wid	Tiroler Bauernzeitung	commentary	Dr Eduard Widmoser	16.02.1950
VOW_1950_lett_editors	Volkswille	letter to the editor	anonymous teacher	19.04.1950
VN_1950_comm_editors	Vorarlberger Nachrichten	commentary	editorial staff	28.03.1950
WP_1949_squib_Weigel	Weltpresse	squib	Hans Weigel	31.08.1949
WP_1950_comm_Weigel1	Weltpresse	commentary	Hans Weigel	13.01.1950
WP_1950_lett1_Kunzl	Weltpresse	letter to the editor	J. F. Kunzl, Austrian Writers' Association	24.01.1950
WP_1950_lett2_Uttner	Weltpresse	letter to the editor	Dr Friedrich Uttner	24.01.1950
WP_1950_lett3_Edlitzberger	Weltpresse	letter to the editor	Erna Edlitzberger	24.01.1950
WP_1950_lett4_Renner	Weltpresse	letter to the editor	Klaus Remmer	24.01.1950
WP_1950_lett5_Faltus	Weltpresse	letter to the editor	Josef Faltus	24.01.1950
WP_1950_lett6_Steininger	Weltpresse	letter to the editor	Max Steininger	24.01.1950

(*Continued*)

abbreviation	source	text type	author	date
WP_1950_lett7_ Twaruzek	Weltpresse	letter to the editor	Hans Twaruzek	24.01.1950
WP_1950_lett8_ Kratochwill	Weltpresse	letter to the editor	Prof Dr R. Kratochwill	03.02.1950
WP_1950_lett9_ Orthofer	Weltpresse	letter to the editor	Karl Orthofer	03.02.1950
WP_1950_lett10_ Huber	Weltpresse	letter to the editor	Franz Huber, typesetter	03.02.1950
WP_1950_lett11_ Raudnitz	Weltpresse	letter to the editor	Oskar Raudnitz, language teacher	03.02.1950
WP_1950_lett12_ Berg	Weltpresse	letter to the editor	Erich Alban Berg	10.02.1950
WP_1950_lett13_ Bachler	Weltpresse	letter to the editor	Emma Bachler	10.02.1950
WP_1950_comm_ Weigel2	Weltpresse	commentary	Hans Weigel	10.02.1950
WP_1950_report_ editors	Weltpresse	report	editorial staff	02.03.1950
WP_1950_report_ APA	Weltpresse	report	APA correspondence	17.03.1950
WZ_1950_ comm_Roll	Wiener Zeitung	commentary	Edwin Rollet	06.01.1950
WZ_1950_ report_editors	Wiener Zeitung	report	editorial staff	24.02.1950

MONIKA WINGENDER

12 Types of Slavic standard languages – with a focus on the Russian standard language between mono- and pluricentricity

ABSTRACT
The collapse of the USSR and Yugoslavia resulted in dramatic changes in the respective language situations. Against the methodological background of comparative standardology and Wingender's (2013) tetrahedron model, this article outlines the main developments in current Slavic standard languages before focusing on Russian as a standard language type. In Russia, language culture and language planning traditionally aim at a monocentric type of the literary standard. The break-up of the Soviet Union and developments in its successor states, however, have been accompanied by an increasing regionalisation and variation of the Russian language. This article analyses this tension between monocentricity and pluricentricity, using the Russian standard language in Ukraine as a case study. The database for analysis comprises (socio)linguistic publications on standard languages as well as developments in language policy and culture. Furthermore, it also incorporates a qualitative media analysis which focusses on debates on Standard Russian in Ukraine.

1. Introduction[1]

With 258 million people speaking it as their mother tongue, Russian ranks eighth among the world's most spoken languages (Moldovan 2020: 183). It is widespread across the globe and, accordingly, has a broad range of functions: Russian is a mother tongue and state language in the multinational Russian Federation, a mother tongue and language of interethnic communication (to varying degrees and in different situations) in the other fourteen successor states of the USSR, a language in the diaspora and, in many countries, a foreign language.

Given this wide distribution of Russian in the world and its multiple functions, questions about the development of pluricentricity in the Russian standard language are obvious. It is these questions that this article will explore. Kamusella (2018: 153) writes that Russian constitutes an exception to other major languages, since languages such as English, French or Spanish 'are pluricentric in their character, meaning that official varieties of these languages are standardised differently in those states where the aforesaid languages are in official use'.

The question of monocentricity or pluricentricity is of particular importance in the case of Russian, since Russian language culture, with its codified norms, is traditionally oriented towards the concept of literary language, which is associated with the language of (classical) writers and, in general, with the model role of writers, and follows a monocentric standard type. Following the collapse of the Soviet Union, however,

1 The methodological basis for this article is the tetrahedron model developed by the author within the framework of a project, funded by the German Research Foundation, on 'Types of Slavic Standard Languages. Theoretical and Empirical Investigations of Current Feature Correlations in Croatian and Russian'. The results of this project were published in Müller & Wingender (2013). The content-related parts of this article are based on the plenary talk on 'Types of Slavic standard languages – with a focus on the Russian standard between mono- and pluricentricity', presented by the author at the conference 'Double Standards: Codified norms and norms of usage in European languages (1600–2020)' at the University of Salzburg in April 2022.

Russian in the successor states of the USSR found itself in a new ethnolinguistic situation. Thereafter, from the point of view of the Russian Federation, it has been referred to as 'near abroad', especially in political terminology. In this context, the ideology of the *Russkij Mir* ('Russian world')[2] has become increasingly important since the turn of the millennium. Originally an ideology promoting Russian language and culture, this ideology has recently developed into a geopolitical concept. In this ideology, the 'velikij i mogučij russkij jazyk' ('great and mighty Russian language')[3]plays a central role. This concept runs through Russian cultural and language history.

With regard to Standard Russian in the successor states of the USSR, the question of language contact and the development of national varieties in the post-Soviet states arises, as does that of its management through the language policy of these respective states. Are there efforts underway to develop Russian as a pluricentric language with different national standards? Among the successor states of the USSR, this article will focus on Ukraine in particular, since it is here that new developments in the language situation, in the use of languages and, especially, in attitudes towards languages are particularly evident. The Russian-Ukrainian conflict, particularly since 2014, demonstrates that the sociolinguistic framework of Russian in Ukraine has changed significantly in recent years. A real turning point in attitudes towards Russian was the launch of the full-scale invasion of Ukraine by Russia in February 2022. These latest developments also raise questions about the detachment from the normative centre and norm authority of the Russian Federation and, accordingly, the development of pluricentricity in the Russian standard language. Since the spread of *Russkij Mir* ideology and the intensification of the

2 With regard to Russia's external language policy and *Russkij Mir* cf. Brüggemann (2017) and the sections 'Exporting Russian: Language as "Soft Power"' and 'The Russian World Foundation' in Gorham (2014: 156–165).

3 Cf. the following quote from Turgenev's 'Poems in Prose': 'In days of doubt, in days of dreary musings on my country's fate, thou alone art my stay and support, mighty, true, free Russian speech!' (<https:&/en.wikisource.org/wiki/Page:The_Novels_of_Ivan_Turgenev_(volume_X).djvu/334>, accessed 05.12.2023)

Russia–Ukraine conflict, the topicality of the issue of the development of pluricentricity in Russian has clearly increased (Kamusella 2018: 155).

Against this background, this article aims to situate questions concerning the development of pluricentricity in the Russian standard language within the larger context of the types of Slavic standard languages. Methodologically speaking, this will be done against the background of comparative standardology and, in particular, Wingender's (2013) tetrahedron model, which is based upon it. The article is predominantly literature-based. To examine the discussion of Russian in Ukraine, we also draw on media analysis. Overall, this article is committed to qualitative methods.

The time frame of the study focusses on the period following the collapse of the Soviet Union and examines in particular the development of pluricentricity in the Russian standard language in the successor states of the USSR. Individual retrospective analyses of the earlier history of standardisation of the Russian language help to clarify the background to discussions on Standard Russian. Since we are focusing on Russian in Ukraine as a case study, a brief examination of the most recent developments following Russia's full-scale invasion of Ukraine in February 2022 will also be included at the end.

Accordingly, the article takes as its starting point (2.) a characterisation of the types of Slavic standard languages according to Wingender's (2013) tetrahedron model, then (3.) moves on to outline traditional views on Russian as a standard language type before (4.) turning to the academic debates on pluricentricity with regard to the Russian standard language. Finally (5.), it will analyse the case study of Standard Russian in Ukraine. In the end, all the results of the investigation will be classified in the tetrahedron model.

2. Types of Slavic standard languages within the framework of comparative standardology

Following the collapse of the USSR and Yugoslavia as well as the so-called Eastern Bloc, numerous factors have shaped the development of Slavic standard languages. The central tendencies can be outlined as follows:

(a) Effects of socio-political developments since the upheavals of 1989:
  - promotion of the status of an existing standard language to a national or state language: functional expansion of previously disadvantaged languages in terms of language policy (e.g. Ukrainian, Belarusian, Slovak);
  - restriction of the scope of a previously widespread (or dominant) standard language (Russian, Czech);
  - division of a pluricentric standard language and standardisation of independent standard languages (Bosnian, Croatian, Serbian, Montenegrin).

(b) Various sociolinguistic developments (general developments that are not limited to the period from 1989 onwards):
  - written language projects: development of standard/literary micro-languages (Kashubian in Poland, Burgenland Croatian in Austria, Rusyn in Vojvodina, Transcarpathia, Slovakia and Poland);
  - changes in the language corpus: accentuation of certain aspects of antinomies such as internationalisation (anglicisation) vs. nationalisation (purism) and democratisation (allowing non-standard language elements) vs. intellectualisation (efforts to differentiate from the vernacular; terminology according to the Prague School) (all Slavic languages);
  - loss of function of standard languages, including (standard) language death (Lower Sorbian).[4]

This outline shows that since 1989, developments in Slavic language situations have been very dynamic and that, accordingly, the question of how many Slavic standard languages there are is not so easy to answer. Hentschel takes this as the occasion for his overview of the Slavic languages (Hentschel 2003). One of the factors in determining the number of standard languages is the question of the development of pluricentricity, as in the case of Serbo-Croatian, as also demonstrated in the overview above.

4 This overview is a translation from Wingender (2010: 193). For a further typological overview of developments in the Slavic standard languages, see Marti (2014).

If we want to further link these general developmental tendencies of Slavic standard languages with their historical developments and thus identify types of Slavic standard languages, then we require a comparative standardology. The task of comparative standardology is to define an inventory of standard language features in relation to which different standard languages can be compared with each other in order to arrive at statements about their particular type. Comparative standardology has a long tradition in Slavic studies (cf. Wingender, Barkijević & Müller 2010: 126–7). The Prague linguistic circle and, above all, the course of the intensive discussion between D. Brozović and N. I. Tolstoj produced milestones in the research on types of Slavic standard languages (exemplary: Brozović 1967; Tolstoj 1985). These milestones were incorporated into the model of types of Slavic standard languages (Wingender 2013), which is designed in the form of a tetrahedron so as to be able to depict the non-hierarchical structure and the correlations between the various features.

The multi-factorial tetrahedron model is based on four components: linguistic, functional, social and situational (see the description of the model in Wingender 2013: 26–31, which includes a figure of the model). To explain it briefly: a language that has an insufficiently developed linguistic component cannot fulfil the functions assigned to it. If the functional component is not developed, the language will neither be supported by language policy nor will it be used. If the social component is not developed, the language is not supported by the language community, so that acceptance problems lead to language abandonment. If the situational component is not developed, the language is threatened by a lack of autonomy and sociolinguistic embedding. If a language only has a linguistic basis, it is neither supported by the creators nor the users and is not used, i.e. it is a mere artefact. These four components each comprise different features. The linguistic component includes features such as heterogeneity of the language base, influences of other languages and influences of other varieties, as well as norms and codification. The functional component includes the features functional spheres, vitality and official attitudes. The social component includes features such as user attitudes, symbolic value and tradition/historicity. The situational component includes the features autonomy and sociolinguistic embedding.

The tetrahedron model is conceptualised as a dynamic model and functions as a force field depicting the types of standard language as a whole.[5] A further fundamental basis for the conceptualisation of the features of the model is the inclusion of both the creators (features including norms, codification, official attitudes) and the users (features like vitality, user attitudes, symbolic value) of standard languages. The model therefore also covers questions of double standards in the sense of norm and usage.

The relationship between codified norms and norms of usage plays an important role in this, not least in the case of the development of pluricentricity. Here we will briefly outline the role this plays in the tetrahedron model. In Section 4 we will then take up the discussion again in order to further pursue the concept of pluricentricity and then, finally, to apply it to the question of pluricentricity in the Russian standard language. How, therefore, can the tetrahedron model capture the dynamic of monocentricity versus pluricentricity? We will illustrate this with two sample fields.

First, in the project that led to the development of the tetrahedron model (see acknowledgements in Fn. 1), we examined two prototypes of Slavic standard languages: Serbo-Croatian as a pluricentric standard language, with all its chequered history, and Russian as a traditionally monocentric one in terms of language culture. While the case of Serbo-Croatian as a pluricentric standard language type is obvious, the question of the development of the Russian standard language with regard to mono- and pluricentricity remains very much under discussion. We address this question in the following subsections, focusing on the current developments of the Russian language in Ukraine.

Secondly, we will adopt the basic definitional criteria of pluricentricity, taking here as an example the widely used conception of Clyne (2004).[6] As a result, the tetrahedron model can reveal key insights into

5 Rabus (2015) applied this model to the characterisation of Slavic standard microlanguages.

6 We return to Clyne in Section 4 and place his concept within the context of further definitions of pluricentricity.

pluricentricity in all of its four components. Thus, the linguistic component addresses pluricentricity in relation to the features of language contact (with regard to the feature of the influence of other languages), norms and codification. The functional component covers pluricentricity with reference to the feature of official attitudes, which is concerned on the one hand with the status of the standard language and on the other with 'the relation of language to power' and language policy. In the case of the social component, the feature of the symbolic value in particular plays a key role, since this concerns 'the relation of language to national identity' and also the role of languages as 'unifiers and dividers of people'. Finally, in the situational component, the sociolinguistic embedding is decisive, insofar as this relates to the interplay between state, nation(s) and language(s) and to the dynamics of language contacts on a social level (all direct quotations are from Clyne 2004: 296).

Before moving on to investigate the question of monocentricity vs. pluricentricity in the Russian standard language, the following section will explore the question of what type of standard language Russian is traditionally considered to be and how this is connected to the question of monocentricity versus pluricentricity.

3. Traditional views on Russian as a standard language type

To be able to characterise Russian as a standard language type, it is important first to examine the underlying concepts and terminology.[7] While, in international linguistics, the codified variety is usually referred to as a standard language, the term 'standartnyj jazyk' is little used in Russian, with that of 'literaturnyj jazyk' ('literary language')

7 Sufficient attention has already been paid to the differences in designation and terminology in standard language research. As an example, we refer to Baum (1987: 35–55).

being much more widespread.[8] Germanova (2011: 7) justifies this through differing conceptions regarding the standardised and codified form of existence of language. Thus, she distinguishes between a 'культурно-ценностная' ('cultural and value-based') and an 'инструментально-пргагматическая' ('instrumental-pragmatic') conception, whereby, according to Germanova, the first is characteristic of Russian linguistics, the second of English sociolinguistics.[9] Accordingly, it is less of a question of terminological differences between standard and literary language than it is one of conceptual differences, and these are culturally conditioned and shaped, primarily, by the respective academic traditions.

The issue of literary language vs. standard language therefore impacts upon the conceptualisation of Russian as a type of standard language. This is because the designation of Russian as a 'literary language' emphasises the 'Topos der Literarizität des Russischen' ('topos of the literariness of Russian') (Kuße 2018: 472). This refers, above all, to nineteenth-century literature as the source of literary language and as a normative model, and it is important to underline this.

Traditional views of Russian as a literary language are reflected in normative dictionaries and grammars, which take the Russian language from Puškin onwards as their basis. This is also reflected in the Academy Grammar books.[10] Here we cite the Academy Grammar (1980) as an example. It aims to provide a scientific description of the current state of the grammatical construction of the Russian literary language ('описание современного состояния грамматического русского литературного языка'). This grammar is based on a rich material base, including

8 An early exception is Brozović (1967).

9 In Western linguistics, variational linguistics is widespread, and with it the view of standard language as a variety in the overall system of language. Although Russian linguistics also traditionally speak of various 'forms of existence' ('formy suščestvovanija jazyka'), including the literary language, the view of the literary language as a cultural (key) concept shapes attitudes to the codified form of language.

10 We will return to the fact that this normative model is currently changing in the further course of this Section.

literature, both current and classical ('художественная литература – современная и классическая'), mentioned above all.[11]

When it comes to Academy dictionaries, subsequent or revised editions published after the collapse of the Soviet Union also built upon this tradition. The four-volume so-called Small Academy Dictionary of 1999 draws upon materials taken from literature, ranging from Puškin to the present day ('из произведений художественной литературы от Пушкина до наших дней', *Slovar' russkogo jazyka 1999*: § 1). The so-called Great Academy Dictionary of 2004 covers the lexis of the Russian language in its classical and new periods ('его классического (XX в.) и нового (XX–XXI) периодов', *Bol'šoj akademičeskij slovar'* 2004: 3) and, like the aforementioned Small Academy Dictionary of 1999, also covers the lexis in general use in the Russian language from Puškin onwards.

This conception of the Russian standard language is thus perpetuated by the model role of writers in Academy publications. Accordingly, discussing the role of writers and literature in the negotiation of linguistic norms, Lunde (2018) states:

> Both issues have a prehistory in Russian culture, since social commitment, as well as moral and philosophical authority, traditionally accompany the task of being 'a great Russian writer', on the one hand, while literary texts have long played a norm-maintaining role in education, on the other. (Lunde 2018: 1)

Mustajoki (2023) describes this model, too:

> Early texts written in a language are often fictional and, in the Russian tradition, texts of renowned writers are regarded as a manifestation and model for 'literary language'. (Mustajoki 2023: 368)

Furthermore, Moldovan's (2020) characterisation of Russian also highlights such aspects when he writes of Russian language and literature:

> The Russian language is one of the most perfect languages in the world, a language with rich writing and great literature, which has made it one of the most

11 The quotations are from the preface to the Academy Grammar (1980: 3).

> authoritative and studied languages in the world. It was Russian classical literature that associated ideas about outstanding achievements of humankind in the search for spiritual values with the Russian language. When speaking about Russian identity and about the Russian language as a unifying element of this identity, we primarily mean the powerful moral and humanistic potential imprinted in Russian that is accumulated in Russian literature. (Moldovan 2020: 183–4)

Here he identifies the authority of the Academy of Sciences as the primary norm authority regarding the literary language:

> We are speaking primarily about the Academy's expert assessment of speech innovations and the codification of the norms of the literary language in normative dictionaries, grammars, and reference books on the culture of speech. (Moldovan 2020: 184)

Although the research repeatedly emphasises that there is a broad variety of norm authorities which currently impact upon Russian, this does not mean that literature and writers no longer play any role in this respect. As Lunde (2018) points out in her monograph, this importance has changed today. Through the example of the characterisation of the language of post-Soviet Russian authors, the author identifies and outlines multiple forms of relationship between standard, norm deviation and variation. In this, Lunde combines linguistics with literary studies in her approach, and analyses the sociolinguistic change and the role of writers in these processes so as to demonstrate the changing role of literature in norm and usage of Standard Russian.

In order to highlight how the literature-oriented normative model has changed in the publications of the Academy, Mustajoki (2023) takes up the example of the Academy Grammar:

> Today we are facing a revolution, or at least a new era, in Russian grammatology. In the Academy of Science Institute of Russian language, a new type of academic grammar is under construction. [...] Instead of using classical literature as a source of examples, it relies on large contemporary corpora. Moreover, the grammar is digitised and regularly updated. (Mustajoki 2023: 372)

What do such traditional views of Russian as a 'literaturnyj jazyk' and the orientation towards the written texts of model writers reveal? The focus on this ideal, supported by corresponding Russian language culture and language policy, is associated with a monocentric type of language – traditionally with the 'literaturnyj jazyk', whereby the norm authority lies in the Russian Federation. However, the situation of the Russian language has been transformed following the collapse of the USSR. Developments in the Russian language outside of Russia, in the fourteen successor states of the USSR, raise the question of the development of pluricentricity with regard to norm and usage. The following section will therefore address the tension between mono- and pluricentricity.

Concluding this section, it should be pointed out that the conceptual meaning of the Russian standard language has undergone numerous socially relevant expansions, particularly after the turn of the millennium. For example, in its status as a state language, the Russian standard language is conceptualised within the context of national identity and, increasingly in recent years, in terms of national security ('lingvističeskaia bezopasnost', cf. Wingender 2018). These framings are accompanied by a whole series of language-related laws and programmes, for example the 'Law on Russian as a State Language of 2005' and the now five 'Federal Target Programmes *Russkij jazyk* ['Russian language']'. In general, these tend to hinder the development of variation and pluricentricity.

4. Russian standard language – academic debates on pluricentricity

In this section, against the background of fundamental conceptions of pluricentricity, we analyse how the Russian standard language has usually been classified in research literature. The focus here is on the Russian standard language outside Russia, and in the successor states of the USSR in particular, rather than on the situation of Russian in the Russian

Federation itself.[12] This is because this article examines developments following the disintegration of the USSR and the relationship between language, nation and identity with regard to pluricentricity. Section 5 then focusses on a case study, Russian in Ukraine.

Since its introduction by William Stewart in the 1960s and Heinz Kloss in 1978 (see Norrby, Lindström, Nilsson & Wide 2020: 202; Clyne 2004: 296), the concept of pluricentricity and its application to different language situations has been intensively discussed in research. The literature on this is now correspondingly rich, as evidenced e.g. by Muhr's summary article on fourty years of 'Misconceptions about pluricentric languages and pluricentric theory' (Muhr 2018: 17). So as not to go beyond the scope of this article, we focus in the following section on a few selected conceptions (for a detailed overview of the state of research see Muhr 2016 and Norrby, Lindström, Nilsson & Wide 2020) and investigate definitional criteria for pluricentricity in particular. In his definition, Clyne (2004) lists various criteria:

> The term pluricentric was employed, fairly incidentally, by Kloss (1978, 66–7) to denote languages with several interacting centres, each providing a national variety with at least some of its own (codified) norms. Such national varieties take an intermediate position between national languages and regional varieties (cf. art. 19, 30). In pluricentric languages we see both the relation of language to national identity and the relation of language to power. They are both unifiers and dividers of people in that they enable different nations to communicate in the same language but express their distinctiveness within that language. Any national variety of pluricentric language is potentially a separate language. As language status is tantamount to a claim to nationhood, many national varieties of pluricentric languages have been declared to be languages in their own right according to the ausbau principle (cf. art 26) (Hindi/Urdu, Malay/Indonesian, Croatian/Serbian/Bosnian) (Clyne 2004: 296).

Clyne's definition thus encompasses criteria for pluricentricity that are repeatedly taken up in the literature: several centres, national varieties, own norms, national identity, language and power, status and

12 Issues of variation in Russian in Russia are analysed, for example, in Krause (2013); for a further example, see Andrews (1995).

nationhood. A major difference in various definitions of pluricentricity is the question of norm and codification, i.e. whether the national varieties each have their own *codified* norms. In his definition above, Clyne includes non-codified norms, which he marks by putting the brackets ('own (codified) norms'). Wikipedia, in contrast, begins its definition with: 'A *pluricentric language* or *polycentric language* is a language with several codified standard forms, often corresponding to different countries'.[13] Muhr, on the other hand, emphasises the process character of the formation of pluricentricity and does not see codification as a condition. Accordingly, he takes up Clyne's definition and strengthens it further:

> It is no coincidence that he put 'codified' into brackets, as the existence of codified norms is not at all a prerequisite for a language being pluricentric. This is particularly not the case in the early stages when a language community/territory splits off or a new settlement is founded. (Muhr 2018: 33)

Muhr therefore arrives at a distinction of three stages of the development of pluricentricity, dependent upon which of the six criteria he mentions (occurrence, status, linguistic distance, relevance for identity, acceptance of pluricentricity, codification of national norms) are fulfilled: minimal, basic or full pluricentricity (Muhr 2018: 35).

Such a gradual conception of pluricentricity is fundamentally suited for the description of language situations in the Slavic-speaking world to help us comprehend the sociolinguistic developments following the disintegration of Yugoslavia and the Soviet Union in their processual nature and dynamics. According to this conception, Muhr describes Russian as a new pluricentric language:

> Russian is a 'new' PCL that came into existence after the dissolution of the Soviet Union in 1991. This resulted in a number of countries outside the Russian Federation where Russian has an official or quasi-official function. (Muhr 2020: 39)

13 <https://en.wikipedia.org/wiki/Pluricentric_language> (last accessed 30.12.2024).

However, since the focus of this paper is explicitly on *standard language* types, we will concentrate here on pluricentric standard languages (rather than pluricentric languages in general) and focus upon the feature 'codified norms' in our further considerations.[14] This will be discussed both in the further course of this section and in the case study on Ukraine in Section 5.

While the definitions of pluricentricity listed at the beginning of this section were developed against the background of German and English in particular, we will now turn to more recent research that focusses on Russian outside of Russia.[15] Since developments in the Russian language outside Russia are anything but homogeneous, there is a wide spectrum of differing language situations (cf. the country-specific contributions in Mustajoki, Protassova & Yelenevskaya 2020; Norman & Kusse 2020; Ryazanova-Clarke 2014). This applies to the so-called near abroad, the successor states of the USSR, and to countries with a large number of Russian migrants such as Germany, Israel or the USA. One reason for the varying developments in Russian outside of Russia are issues of status which, according to Muhr's aforementioned conception, form one of the criteria of pluricentricity. The status of Russian in the successor states of the USSR ranges from state language in Belarus and official language in Kazakhstan and Kyrgyzstan, through language of interethnic communication, e.g. in Uzbekistan, to national minority language, e.g. in Estonia, and foreign language in Latvia. It is clear that, in terms of vitality, attitudes, identity and use, these different state and social functions are accompanied by differing sociolinguistic developments.

14 Our approach to include codified norms is thus conditioned by the object of study of this article (standard language types). In other studies, the inclusion of non-codified norms may be useful when dealing generally with national varieties of pluricentric languages and their stages of development (according to Muhr).

15 With regard to the study of pluricentricity in Russian, there are sometimes certain terminological differences or other developments (such as 'Georusistika') in Russian Studies compared to other research contexts. These are discussed in detail in Del Gaudio (2013: 344–8), also with regard to contrary positions. In addition, Del Gaudio (2013: 349) contrasts the partially divergent criteria of pluricentricity in Krjučkova (2011) and Muhr.

The question of monocentricity vs. pluricentricity in the Russian language is controversial in Russian society, not least because of the conception of literary language discussed in Section 3. Thus, with reference to the 'great and mighty Russian' discussed here, Mustajoki, Protassova & Yelenevskaya (2020) write:

> The dilemma of the same and different, the norm and a variant is a trigger for never-ending discussions in the field of Russian Studies [...]. When it comes to identity, and the so-called cultural code, when debates touch on the present and future of the 'great and mighty' Russian language, all layers of Russian society get involved, defending their privilege and exclusive right to say 'our Russian'. (Mustajoki, Protassova & Yelenevskaya 2020: 8)

Especially in relation to pluricentricity and the norm authorities, discussions about pluricentricity in Russian repeatedly raise the question of ownership, i.e. who (which institution, group, etc.) has the right to set the norm (Mustajoki 2020: 22; Del Gaudio 2013: 348).[16] Language policy in the successor states of the USSR tends to focus on planning the respective titular and state language, rather than Russian. Thus, comparing Russian in Belarus, Kazakhstan and Ukraine, Del Gaudio (2013) concludes the following:

> At the present stage there remain a few open issues. The first has to do with the variety which should theoretically be chosen as a standard for the national variety. Secondly, as is also the case of Kazakh Russian, there is so far no institutional recognition of these varieties. Thirdly, there is no widespread language consciousness among the average speakers of using a specific national variety. Therefore, at the present state, the question of Russian national varieties remains a scholarly research interest that could, in the future, lead to their codification. (Del Gaudio 2013: 360)

16 Kamusella identifies an intensification of the discussion on ownership in recent years: 'This lukewarm approach to the question whether Russian is a monocentric or pluricentric language, whether only Russia has the right to "own" and control it, or perhaps this privilege should be shared at least with all the interested post-Soviet states and Israel, changed decisively in the mid-2010s'. (Kamusella 2018: 177)

We will return to the question of norm authorities and norm setting in the case study on Ukrainian in Section 5.

A further criterion of pluricentricity relates to linguistic differences. How does Russian in the successor states of the USSR differ from Russian in Russia? Since language contact between Russian and the languages in the successor states of the USSR presents a very heterogeneous picture (cf. the comparative contact linguistic study by Verschik 2010), it is only possible to provide a general characterisation: Differences exist in the lexical sphere, which is obvious, since administrative, social, cultural and other peculiarities in Kazakhstan, Belarus or Latvia, among other states, require their own designations. Furthermore, to varying degrees, there are also phonetic, morphological, syntactic and pragmatic differences to Russian in Russia.[17]Examining the situation in the successor states as a whole, Mustajoki (2020) distinguishes three groups. He writes:

> The varieties of Russian used in Belarus and Kazakhstan and to a lesser extent in some other former Soviet republics, where Russian appears in official documents, differ from Russian Russian, especially in lexical terms. In most cases, no systematic normalisation measures are taken, but we see an inherent, 'natural' norm creation. (Mustajoki 2020: 29)

Mustajoki thus refers once again to the aforementioned question of norm-setting centres. He goes on to say that the second group includes countries with a large number of Russian migrants, and the third covers the '"Island Russian" varieties' (Mustajoki 2020: 29; e.g. Alaska Russian).

With regard to the criteria of pluricentricity, it should also be added that language questions in Eastern Europe have often interacted, both

17 Linguistic differences between Russian in Russia and Russian outside Russia are described on a country-specific basis, since, as mentioned above, the language contact situations in the successor states of the USSR are very different. It should be noted that the volumes by Mustajoki, Protassova & Yelenevskaya (2020), Ryazanova-Clarke (2014) and Norman & Kusse (2020) include country-specific studies, cf. as an example the study on Latvian Russian by Berdicevskis (2014), who also explicitly includes the question of pluricentricity, and on Belarusian Russian and Ukrainian Russian by Zeller & Sitchinava (2020).

historically and in the present day, with political aspects. Kamusella (2018) outlines what this means for discussions of pluricentricity in Russian. For example, he highlights the political standardisation of monocentricity by Russia, cf., in his article, the sections 'Linguistic Monocentrism in the Service of Russian Neo-Imperialism?' (Kamusella 2018: 162–8) and 'Peace and Pluricentric Russian' (2018: 182–4), in which he also devotes much space to the analysis of the *Russkij Mir* ideology.[18]

Before moving on to the case study of Russian in Ukraine, it can be noted as an interim conclusion that, in order to understand pluricentricity, a set of different criteria is required. Evaluating the preceding discussions, and with a view to the tetrahedron model mentioned at the beginning, we can, for the time being, outline the following criteria: Language contact, language variation, national varieties, own norms, codification and language policy including corpus, status, prestige, learning/teaching planning. In addition, in Section 5 below we will explore the criteria of symbolic value, identity, attitudes and acceptance. At the end of Section 5, using the tetrahedron model, we will then discuss these criteria in the context of all the results of this article.

5. Debates on pluricentric developments in the Russian standard language in the successor states of the USSR: The case of Ukraine

Muhr's conception of pluricentricity (2018: 35) includes both (socio)linguistic criteria (occurrence, status, linguistic distance, codification of national norms) and criteria that deal with symbolic values and attitudes in a broader sense (Muhr: relevance for identity, acceptance of pluricentricity). In Wingender's (2013) tetrahedron model, these criteria cover the social component, as well as the linguistic and functional components. In this section, without losing sight of the linguistic and functional

18 We will return to his remarks on Russian in Ukraine in Section 5 of this chapter.

components, we will use this as a basis for examining the social component in the case of Russian in Ukraine. In addition to literature-based studies on Russian in Ukraine, we will also conduct a media analysis, which will focus in particular on attitudes to pluricentricity in Ukraine.

We will briefly return to the country-specific research on pluricentricity in the successor states of the USSR mentioned in Section 4. The recent edited volumes by Mustajoki, Protassova & Yelenevskaya (2020), Norman & Kusse (2020), Ryazanova-Clarke (2014), as well as the literature review in Kamusella (2018: 176), show that research thus far has tended to be on Belarus, Estonia, Latvia, Lithuania, Kazakhstan, Moldova and Ukraine, with the other successor states less investigated with regard to pluricentricity. This means that for Ukraine, which will form the basis of our case study, there is a solid literature base, particularly since some of these studies are also comparative, e.g. Del Gaudio (2013), comparing Ukraine, Belarus and Kazakhstan, and Zeller & Sitchinava (2020) on Belarus and Ukraine.

In examining current research, the following picture emerges: although various pluricentric developments in Russian in Ukraine have been identified, the assessment of the extent to which Russian in Ukraine can be classified as a national variety of a pluricentric Russian or is on the way to becoming one varies. First it is important to outline some of the linguistic differences between Ukrainian Russian and Russian Russian, which are often mentioned:

- Lexical differences: '*zajava* instead of *zajavlenie* "application"', '*deržava* instead of *gosudarstvo* "state"' (Zeller & Sitchinava 2020: 117);
- Phonetic-phonological differences: 'fricative realisation for R-Russian /g/, Ukrainian Okanje (distinction of unstressed /o/ and /a/) vs. R-Russian Akanje' (Zeller & Sitchinava 2020: 116);
- Morphosyntactic variation: *dumat' za kogo-to* (думать за кого-то) + accusative case vs. *dumat' o kom-to* (думать о ком-то) + prepositional 'to think of/about' (example from the overview table in Del Gaudio & Ivanova 2015: 180).

Furthermore: how are these linguistic differences evaluated with regard to pluricentric developments in the Russian language?[19] Del Gaudio & Ivanova (2015) assume that there are linguistic features on all levels and see Russian in Ukraine as a 'NDV [non-dominant variety] in formation' (Del Gaudio & Ivanova 2015: 169). The authors conclude that

> The case study based on some recurrent grammatical patterns derived from the Russian press of Ukraine, and particularly evident in prepositions governing, supports the thesis that the formation of a potential Ukrainian variety of Russian could be underway (Del Gaudio & Ivanova 2015: 190).[20]

Combining language and politics, Kamusella (2018) analyses the politicised discourse surrounding Ukrainian Russian. He proposes the development of a Ukrainian Russian in terms of pluricentricity in order 'to defuse the conflictual nature of the current official Russian thinking on the Russian language' (Kamusella 2018: 182). The Russian language should be de-ethnicised, so that Russia cannot continue to associate it with political and territorial aims: 'The Russian language would stop functioning as "fate" or "destiny", defined through the prism of ethnolinguistic nationalism in the service of the neo-imperial ideology of Russkii

19 One problem in the discussions surrounding the evaluation of linguistic differences is the lack of broad variation-linguistic studies, cf. Zeller & Sitchinava (2020: 112): 'One has to emphasise, however, that at present, comprehensive variational-linguistic investigations, which on the one hand deal with the usage frequency and the degree of stabilisation of the B-Russian and U-Russian features, and on the other hand investigate language change, social differentiation, style variation, and social assessments of the variation in B-Russian and U-Russian, do not exist. There is a high degree of uncertainty regarding the distribution, frequency, and stability of most of the features discussed below'.

20 The opposite tendency, from pluricentricity to monocentricity, is identified by Del Gaudio & Dorofeev (2016) in Crimea. Due to the annexation of Crimea by Russia and the accompanying change in language policy, they state: 'One can witness an evident shift between a former pluricentricity of Crimean Russian (C-Russian), which was in many respects not really differentiated from its Ukrainian-Russian (U-Russian) counterpart, towards a monocentric model steered by the centre'. (Del Gaudio & Dorofeev 2016: 449)

Mir' (Kamusella 2018: 184). His starting point is the comparison of the situation of Russian with that of other major languages, such as English, French or Spanish, which are pluricentric. At the end of his study, he emphasises the political aspect of pluricentricity:

> A more stable and peaceful world is possible. A de-ethnicised pluricentric Russian language – thus overhauled into a colourful multiethnic and multicultural plenitude of world Russians – could be a versatile means to this end. The globe's pluricentric Frenches or Englishes are a clear case in point. (Kamusella 2018: 184)

Moser, on the other hand, does not see Russian in Ukraine as being on the way to becoming a national variety of Russian. He concludes on the basis of a broader standardological approach:

> Keines der Merkmale begründet jedoch zwingend die Existenz einer einheitlichen ukrainischen Variante der russischen Sprache, umso weniger die Notwendigkeit ihrer Standardisierung [...]
>
> 'None of the features, however, necessarily establishes the existence of a uniform Ukrainian variant of Russian, all the less the need for its standardisation [...]' (Moser 2022: 418).

Pluricentricity, furthermore, is primarily a result of language planning (Moser 2020: 187, 2022: 417). As a further argument, Moser also cites the features of prestige and attitudes, since local features are associated with low language prestige in the standard, and a broad consensus does not exist. The traditional Russian standard continues to be prestigious. What should also not be underestimated is that the standardisation of a national variety of Russian in Ukraine would 'auf eine Anhebung des Status des Russischen in der Ukraine hinauslaufen, welcher derzeit von einer deutlichen Mehrheit der Bevölkerung gar nicht gewünscht wird' ('amount to an elevation of the status of Russian in Ukraine, which is currently not at all desired by a clear majority of the population') (Moser 2022: 419).[21]

21 See also the detailed analysis of the discussion on a national variety of Russian in Ukraine in Moser (2020).

To offer an interim conclusion, it can be stated that the question of pluricentricity should not only be seen through the prism of variation (due to language contact in the successor states of the USSR) and political factors. As is generally the case with standard languages, including pluricentric standard languages, a whole set of criteria of different components must be taken into account. Accordingly, on the basis of the tetrahedron model (Wingender 2013), we come to a similar assessment to that of Moser (2020, 2022) with regard to the question of a national variety of Russian in Ukraine. Before we finally clarify this on the basis of the tetrahedron model, we now turn to the question of how a Ukrainian Russian is discussed in Ukrainian society.

As has been mentioned above, linguistic differences grounded in the Russian-Ukrainian contact situation are not sufficient to allow us to speak of a national variety of a pluricentric Russian standard language. What proposals have been made with regard to the standardisation of Russian in Ukraine, and what have the reactions to these been within the speech community? There have been several attempts, for example, to develop or implement a Ukrainian Russian. Discussion on the issue within the mass media has intensified, particularly in the past decade, which can be explained by the Russo-Ukrainian conflict and the role of the *Russkij Mir* ideology.[22] In the online media in Ukraine, as well as in contributions from Ukrainian news and information agencies, it is possible to identify certain blocks of events.[23] The media analysis that follows does not claim to be comprehensive: its function, rather, is to systematise certain blocks of topics and chronological sequences with regard to societal debates on a Ukrainian Russian. The following quotes are in Ukrainian and are translated into English.

22 On this, see also Kamusella (2018).

23 See also Moser's review of media texts (Moser 2020: 185–6, fn. 1).

I. Ukrainian standard of the Russian language

In 2014, Serhii Koshman (coordinator of the movement 'We are Europeans', Kharkiv) proposed the codification of the Ukrainian version of the Russian language: 'Не віддамо нашу російську мову росіянам!' ('We will not give our Russian language to the Russians!').[24]

This thematic strand also includes the 'Electronic petition on the website of the President of Ukraine to introduce the Ukrainian standard of the Russian language' (2015), which, however, received only nineteen votes ('Електронна петиція на сайті Президента України про запровадження україського стандарту російської мови').[25]

Here, one can discern a political (rather than standardological) argumentation, with a clear connection between war and language. One example of this is the recognition of a specific status for Ukrainian Russian as a counteroffensive tool in the hybrid war: 'Російська мова. Прибрати монополію Москви?' ('The Russian language. Remove Moscow's monopoly?', 2016).[26]

A further example of a political argumentation is the contribution 'Андрій Курков: Навіть за Союзу литовці змогли сховатися від радянської реальності' ('Andriy Kurkov: Even during the Soviet Union, Lithuanians were able to hide from the Soviet reality').[27] In his interview, Kurkov also proposes the establishment of an Institute of the Russian Language in Ukraine. This brings us on to our next thematic strand.

24 <https://www.istpravda.com.ua/columns/2014/04/14/142381/> (last accessed 05.12.2023).

25 <https://petition.president.gov.ua/petition/15853> (last accessed 05.12.2023). See also Moser (2020: fn. 26).

26 <https://ua.krymr.com/a/28084658.html> (last accessed 05.12.2023).

27 <https://rozmova.wordpress.com/2017/01/01/andrij-kurkov-4/> (last accessed 05.12.2023).

II. Institute of the Russian language in Ukraine

In 2019, the historian Timothy Snyder recommended the standardisation of Ukrainian Russian and the establishment of a corresponding institute. In an interview from 7 July 2019, Snyder spoke from his own experience with American English, thus bringing in the pluricentricity debate. His interview demonstrates the political argumentation already mentioned on several occasions above: 'If you officially had your own Ukrainian version of the Russian language that would be a very powerful argument against the Russian propaganda'. ('Historian Snyder on Language, Zelenskyy and Vakarchuk').[28]

Snyder reiterated the idea of establishing a State Institute of Russian Language and Culture in another interview (10 July 2019): 'Тімоті Снайдер: створіть в Україні інститут російської мови – і в Путіна будуть проблеми' ('Timothy Snyder: Create a Russian Language Institute in Ukraine – and Putin will have problems').[29]

Snyder's proposal for the creation of a state institute for the Russian language in Ukraine is reported in the following article from 10 July 2019: 'Інститут російської мови в Україні. Чи потрібен?' ('Institute of the Russian Language in Ukraine. Is it necessary?').[30]

Linguists and writers commented on Snyder's statement. A selection of comments can be found in the article from 20 July 2019: 'Чи потрібна українцям "власна" російська мова?' ('Do Ukrainians need their "own" Russian language?').[31] For example, linguists argue that the differences are not sufficient to consider Russian a variant of the literary

28 <https://hromadske.ua/en/posts/ukrainians-own-russian-but-dont-admit-it-historian-snyder> (last accessed 05.12.2023).

29 <https://www.bbc.com/ukrainian/features-48924775> (last accessed 05.12.2023).

30 <https://ua.korrespondent.net/ukraine/4116790-instytut-rosiiskoi-movy-v-ukraini-chy-potriben?fbclid=IwAR1JHtli7dad_Vu5h0KjByt03b_nZ7F0pgK0BnF9eM7M7qNXOYek6ymRsOY> (last accessed 05.12.2023).

31 <https://www.bbc.com/ukrainian/features-49022213> (last accessed 05.12.2023).

norm in Ukraine. They also argue that such a codification would be undertaken primarily for political motives.

III. A Ukrainian variety of the Russian language – part of the cultural diversity of Ukraine

In spring 2021, Arsen Avakov, Minister of Internal Affairs (2014–2021), and Yulia Mendel, Volodymyr Zelenskyi's press secretary (2019–2021), proclaimed that the Ukrainian variety of the Russian language is a part of the cultural diversity of Ukraine and that the Russian language belongs to Ukraine.[32]

In 2021, this provoked a wider debate in the media; for reasons of space, we can only mention a few examples below: 'Балачки про "українську російську" мову тягнуть державу в колоніальне минуле – Капранов' ('Kapranov: Talking about a "Ukrainian–Russian" language drags the state into the colonial past').[33] This includes statements from writers, linguists and politicians highlighting the political instrumentalisation of Russian.

The linguist Bohdan Ažnjuk argues that there is no serious scientific basis for the notion of a Ukrainian variant of the Russian language, since the linguistic differences are not significant enough: 'Термін "українська російська" мова не має наукового підґрунтя' ('The term "Ukrainian Russian" has no scientific basis').[34]

32 <https://www.youtube.com/watch?app=desktop&v=leM40qTf82M>; <https://www.youtube.com/watch?v=JMQ_7d2olfg> (last accessed 05.12.2023).

33 <https://www.radiosvoboda.org/amp/balachky-pro-ukrainsku-rosiysku-movu-tyahnut-derzhavu-v-mynule-kapranov/31291962.html> <https://www.radiosvoboda.org/amp/balachky-pro-ukrainsku-rosiysku-movu-tyahnut-derzhavu-v-mynule-kapranov/31291962.html> (last accessed 05.12.2023).

34 <https://www.ukrinform.ua/rubric-society/3234108-termin-ukrainska-rosijska-mova-ne-mae-naukovogo-pidgrunta-movoznavec.html?fbclid=IwAR1dgsonH1eJrOhT_H8uy1c1BaZMrkTqgg-3VzJ1ADC84Ir15th9HUqO-nOk (last accessed 05.12.2023).

The argument about the lack of a scientific basis is repeated and defended with reference to the Ukrainian–Russian hybrid variety Suržyk: 'Мовознавець Авраменко про "український варіант" російської мови: Це має свою назву – суржик' ('Linguist Avramenko on the "Ukrainian version" of the Russian language: It has a name – Surzhyk').[35]

Taras Kremin, the State Language Protection Commissioner, also issued a critical response: '"Українська російська мова" від Мендель. Як прессекретарка Зеленського спровокувала новий скандал' ('"Ukrainian–Russian Language" by Mendel. How Zelensky's press secretary provoked a new scandal'),[36] in which he refers to the Constitution of Ukraine.

The head of the Ukrainian Institute of National Memory commented on the debate in 'Що треба розуміти про російську мову в Україні' ('What to understand about the Russian language in Ukraine').[37] His argument is based on the concern that support for the Russian language would not lead to a situation like that in Switzerland, but rather one more akin to that in Belarus.

Evaluating the presented facts and arguments in our article, we conclude with a classification of Russian in Ukraine from a standardological perspective. Thus, using the tetrahedron model (Wingender 2013), the following picture emerges:

35 <https://detector.media/infospace/article/187000/2021-04-14-movoznavets-avramenko-pro-ukrainskyy-variant-rosiyskoi-movy-tse-maie-svoyu-nazvu-surzhyk/> (last accessed 05.12.2023). <https://detector.media/infospace/article/187000/2021-04-14-movoznavets-avramenko-pro-ukrainskyy-variant-rosiyskoi-movy-tse-maie-svoyu-nazvu-surzhyk/> (last accessed 05.12.2023).

36 <https://nv.ua/ukr/ukraine/politics/movniy-skandal-z-mendel-chi-ye-ukrajinska-rosiyska-mova-novini-ukrajini-50152215.html> (last accessed 05.12.2023).

37 <https://wz.lviv.ua/blogs/454946-shcho-treba-rozumity-pro-rosiisku-movu-v-ukraini> (last accessed 05.12.2023).

(1) Linguistic component
On the relationship between norm and usage: due to the Ukrainian–Russian language contact in Ukraine, there are differences in the usage of Russian in Ukraine, and these are evaluated differently in the research literature, but there has been no codification in dictionaries or grammars.

(2) Functional component
In language policy, a distinction is usually made between status, corpus, prestige, language acquisition, usage and discourse planning (after Marten 2016). Here it can be stated that no systematic strategies of language policy developing Russian in Ukraine into a national variety of a pluricentric standard language can thus far be discerned.

(3) Social component
With regard to Russian, the traditional Russian standard remains prestigious. Since the role of Russian in the successor states of the USSR is quite varied, and often also controversial, no institutions have yet emerged that could develop Russian as a pluricentric standard language. In Ukraine, the standardisation of Russian as a national variety is viewed critically. There is a tension when it comes to arguments surrounding the attitudes of users: the standardisation of Ukrainian Russian as a tool against Russia's propaganda on the one hand versus the danger it poses to the Ukrainian state language on the other.

This was the language situation prior to February 2022: Current developments in Ukraine's language situation following Russia's full-scale invasion have cast the question of the development of pluricentricity of Russian in a new light – Russian is increasingly stigmatised as the language of the enemy and aggressor. A debate on the standardisation of a national variety of Russian in Ukraine is currently inconceivable. The symbolic value of Ukrainian as the language of resistance is increasing. Clyne writes (2004: 296):

> In pluricentric languages we see both the relation of language to national identity and the relation of language to power. They are both unifiers and dividers of people in that they enable different nations to communicate in the same language but express their distinctiveness within that language.

Russia's war of aggression against Ukraine strengthens the link between the Ukrainian language and national identity in Ukraine and the distinctiveness from Russia and Russian as the language of the aggressor. This can be seen, not least, in the toppling of monuments to Russian writers such as Puškin and their banishment from the school curriculum in Ukraine. This underlines once again that a debate on the standardisation of a national variety of Russian in Ukraine is currently inconceivable.

(4) Situational component
The disintegration of the USSR and the establishment of independent successor states could have formed the basis for the standardisation of national varieties of Russian in the sense of a pluricentric standard language. Alongside Ukraine, further candidates for the emergence of such a variant in the research literature are particularly those successor states in which Russian has an official function (such as Belarus and Kazakhstan). In Ukraine, due to Russia's full-scale invasion and the stigmatisation of Russian as the language of the enemy, such developments are currently inconceivable. Kulyk (2022: 237) writes: 'Russisch gilt als Sprache des Feindes. Was die Ukraine in dreißig Jahren mit ihrer Sprachpolitik nicht erreicht hat, hat Russland mit dem Angriff auf das Nachbarland provoziert. Der große Krieg befördert den Sprachwechsel' ('Russian is considered the language of the enemy. With its attack on its neighbour, Russia has provoked that which Ukraine could not achieve in thirty years through its language policy. The great war promotes the shift of language'.)

The disintegration of the Soviet Union sheds light, once again, on the various factors that are at work in the standardisation of languages, and which, accordingly, can also constitute turning points in processes of standardisation.

Finally, I would like to comment on the 'pluricentric standard language' type. In the discussion of Russian in the successor states of the USSR, reference is often made to pluricentric languages such as English or, in the Slavic-speaking world, Serbo-Croatian. Kamusella (2018: 183) writes:

'Then, a new category of post-Russian languages could emerge, not dissimilar to that of the post-Serbo-Croatian languages of Bosnian, Croatian, Montenegrin and Serbian'. However, Serbo-Croatian constitutes a different type of standard language to Russian. The chequered history of the standardisation of Serbo-Croatian fluctuates between the unifying language type (joint standard language e.g. in the first phase of the first Yugoslavia and Tito's Yugoslavia) and the diverging language type (e.g. in the Independent State of Croatia (NDH), from the 1960/70s, and following the dissolution of Yugoslavia). By contrast, and as explained in Section 3, the Russian standard language is conceived as forming a continuity from Puškin to the present day, with corresponding model writers and further model norm institutions. Although the disintegration of the Soviet Union has created opportunities for the development of a pluricentric standard language (situational component in the tetrahedron model), the functional component in the form of the creators' attitudes (official attitudes in the tetrahedron model) is scarcely strategically active.

6. Conclusions

To conclude, I would like to summarise the aspects discussed in this paper on the pluricentricity of Russian against the background of types of Slavic standard languages: With regard to the question of a national variety of Russian, the development of the Russian language in Ukraine has thus far been very controversial in research. The basis for the conflicting assessments is the differing evaluation of language variation, as well as the different weighting given to linguistic, political, and standardological criteria. Standardological approaches to the question of pluricentricity have thus far tended to see no development towards a national variety of Russian in Ukraine. Since February 2022, furthermore, the discussion surrounding the development of a national variety of Russian in Ukraine has come to a halt since, with the growing stigmatisation of Russian as the language of the aggressor and the enemy, the role of Russian in Ukraine is becoming increasingly limited. With regard to types of Slavic standard languages, this very clearly underlines the role played by social and situational factors in correlation with the linguistic and functional component in the tetrahedron model (cf. Wingender 2013).

Bibliography

Academy Grammar. 1980. *Russkaja Grammatika* 1980. Tom 1. Moskva: Nauka.

Baum, Richard. 1987. *Hochsprache, Literatursprache, Schriftsprache. Materialien zur Charakterisierung von Kultursprachen*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Berdicevskis, Aleksandrs. 2014. Predictors of pluricentricity: Lexical divergences between Latvian Russian and Russian Russian. In Larissa Ryazanova-Clarke. *The Russian Language Outside the Nation*. Edinburgh: Edinburgh University Press. 225–45.

Bol'šoj akademičeskij slovar' russkogo jazyka 2004. Pod. Red. K. S. Gorbačeviča. Sankt Peterburg: Nauka.

Brozović, Dalibor 1967. Slavjanskie standartnye jazyki i sravnitel'nyj metod. *Voprosy Jazykoznanija* 1, 3–33.

Brüggemann, Mark 2017. Sootečestvenniki, diaspora und Russkij mir: Die externe Sprachpolitik Russlands nach 1991. In Kai Witzlack-Makarevich & Nadja Wulff (eds). *Handbuch des Russischen in Deutschland: Migration – Mehrsprachigkeit – Spracherwerb*. Berlin: Frank & Timme. 199–219.

Clyne, Michael G. 2004. Pluricentric language/Plurizentrische Sprache. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds). *Sociolinguistics. An International Handbook of the Science of Language and Society.* Vol. 1, 2nd. ed. Berlin, New York: de Gruyter. 296–300.

Del Gaudio, Salvatore 2013. Russian as a non-dominant variety in post-Soviet states: a comparison. In Rudolf Muhr, Carla A. Negre, Carmen F. Juncal, Klaus Zimmermann, Emilio Prieto & Natividad Hernández (eds). *Exploring Linguistic Standards in Non-Dominant Varieties of Pluricentric Languages/ Explorando estándares lingüísticos en variedades no dominantes de lenguas pluricéntricas*. Wien et al.: Lang. 343–62.

Del Gaudio, Salvatore & Jurij Dorofeev. 2016. The Russian language in Crimea: From pluricentricity to monocentricity. In Rudolf Muhr (ed.). *Pluricentric Languages and Non-Dominant Varieties Worldwide. Part I: Pluricentric Languages across Continents. Features and Usage*. Wien et al.: Lang. 449–68.

Del Gaudio, Salvatore & Olga Ivanova. 2015. A variety in formation? Morphosyntactic variation in Ukraine-Russian speech and press. In Rudolf Muhr & Dawn Marley, in collaboration with Anu Bissoonauth and Heinz L. Kretzenbacher (eds). *Pluricentric Languages. New Perspectives in Theory and Description*. Wien et al.: Lang. 169–92.

Germanova, Nataliia N. 2011. *Teorija i istorija literaturnogo jazyka v otečestvennom i anglojazyčnom jazykoznanii*. Moskva: Librokom.

Gorham, Michael S. 2014. *After Newspeak: Language Culture and Politics in Russia from Gorbachev to Putin*, Ithaca NY: Cornell University Press.

Hentschel, Gerd. 2003. Die slavische Sprachenlandschaft am Anfang des 21. Jahrhunderts. In Eva Gugenberger & Mechthild Blumberg (eds). *Vielsprachiges Europa: zur Situation der regionalen Sprachen von der Iberischen Halbinsel bis zum Kaukasus*. Frankfurt a. M.: Lang. 157–75.

Kamusella, Tomasz. 2018. Russian: A monocentric or pluricentric language? *Colloquia Humanistica* 7, 153–96.

Krause, Marion. 2013. Das Image regionaler Varietäten als Indikator soziolinguistischer Kompetenz und metalinguistischer Bewusstheit: HerkunftssprecherInnen und monolinguale MuttersprachlerInnen im Vergleich. In Sebastian Kempgen, Monika Wingender, Norbert Franz & Miranda Jakiša (eds). *Deutsche Beiträge zum 15. Internationalen Slavistenkongress Minsk 2013*. München, Berlin: Otto Sagner. 175–86.

Krjučkova, Tat'jana B. 2011. Problema vozniknovenija i funkcionirovanija nacional'nych variantov jazyka. In Rudjakov Aleksandr N. (ed.) *Georusistika*. Vyzovy XXI Veka. Simferopol': Antikva, 114–25.

Kulyk, Volodymyr. 2022. Die Sprache des Widerstands. Der Krieg und der Aufschwung des Ukrainischen. *Osteuropa* 72(6–8), 237–48.

Kuße, Holger. 2018. Review: Language on Display: Writers, Fiction and Linguistic Culture in Post-Soviet Russia by Ingunn Lunde. *Zeitschrift für Slavische Philologie* 74(2), 472–6.

Lunde, Ingunn. 2018. *Language on Display: Writers, Fiction and Linguistic Culture in Post-Soviet Russia*. Edinburgh: Edinburgh University Press.

Marten, Heiko F. 2016. *Sprachenpolitik. Eine Einführung*. Tübingen: Narr.

Marti, Roland. 2014. Rechtlicher und faktischer Status slavischer Standardsprachen und Sprachenkonflikte. In Sebastian Kempgen, Peter Kosta, Tilman Berger & Karl Gutschmidt (eds). *Die slavischen Sprachen. Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung*. Berlin: de Gruyter Mouton. 1972–85.

Moldovan, Aleksandr M. 2020. The Russian language: Current status and academic studies. *Herald of the Russian Academy of Sciences* 90(2), 183–8.

Moser, Michael. 2022. Geschichte und Gegenwart des Russischen in der Ukraine. Ein Überblick. *Die Welt der Slaven* 67(2), 393–423.

Moser, Michael A. 2020. Is a Ukrainian standard of the Russian language on the agenda? *Kyiv-Mohyla Humanities Journal* 7, 185–96.

Muhr, Rudolf. 2016. The state of the art of research on pluricentric languages: Where we were and where we are now. In Rudolf Muhr (ed.). *Pluricentric Languages and Non-Dominant Varieties Worldwide*. Frankfurt a. M.: Lang. 13–37.

Muhr, Rudolf. 2018. Misconceptions about pluricentric languages and pluricentric theory – an overview of 40 years. In Rudolf Muhr & Benjamin Meisnitzer (eds). *Pluricentric Languages and Non-Dominant Varieties Worldwide: New Pluricentric Languages – Old Problems*. Wien et al.: Lang. 17–56.

Muhr, Rudolf. 2020. European pluricentric languages in contact and conflict – an overview. In Rudolf Muhr, Castells Mas, Angel Josep & Jack Rueter (eds). *European Pluricentric Languages in Contact and Conflict*. Frankfurt a. M.: Lang. 11–63.

Müller, Daniel & Monika Wingender (eds). 2013. *Typen slavischer Standardsprachen. Theoretische, methodische und empirische Zugänge*. Wiesbaden: Harrassowitz.

Mustajoki, Arto. 2020. The democratization of Russian. In Arto Mustajoki, Ekaterina Protassova & Maria Yelenevskaya (eds). *The Soft power of the Russian Language: Pluricentricity, Politics and Policies*. London, New York: Routledge. 21–34.

Mustajoki, Arto. 2023. Literary norms in Russia: Past and present. In Joan C. Beal, Morana Lukač & Robin Straaijer (eds). *The Routledge Handbook of Linguistic Prescriptivism*. Abingdon: Routledge. 368–85.

Mustajoki, Arto, Ekaterina Protassova & Maria Yelenevskaya. 2020. *The Soft Power of the Russian Language: Pluricentricity, Politics and Policies*. London, New York: Routledge.

Norman, Boris & Kusse, Holger. 2020. *Russkij jazyk za predelami Rossii. Kollektivnaja monografija*. Ekaterinburg, Moskva.

Norrby, Catrin, Jan Lindström, Jenny Nilsson & Camilla Wide. 2020. Pluricentric languages. In Jan-Ola Östman & Jef Verschueren (eds). *Handbook of Pragmatics*. Amsterdam, Philadelphia: Benjamins. 201–20.

Rabus, Achim. 2015. Slavische Minderheitensprachen – Mikrostandardsprachen? *Zeitschrift für Slavische Philologie* 71(1), 147–74.

Ryazanova-Clarke, Lara. 2014. *The Russian Language Outside the Nation*. Edinburgh: Edinburgh University Press.

Slovar' russkogo jazyka 1999. V 4-ch t. Pod red. A. P. Evgen'evoj. Moskva.

Tolstoj, Nikita I. 1985. Slavjanskie literaturnye jazyka i ich otnošenie k drugim jazykovym idiomam (stratam) (opyt sravnitel'nogo rassmotrenija). In Mirra

M. Guchman (ed.). *Funkcional'naja stratifikacija jazyka*. Moskva: Nauka. 9–24.

Verschik, Anna. 2010. Contacts of Russian in the post-Soviet space. *Applied Linguistic Review* 1, 85–128.

Wingender, Monika. 2010. Die slavischen Sprachen in Europa. In Uwe Hinrichs (ed.). *Handbuch der Eurolinguistik*. Wiesbaden: Harrassowitz. 189–208.

Wingender, Monika. 2013. Modell zur Beschreibung von Standardsprachentypen. In Daniel Müller & Monika Wingender (eds). *Typen slavischer Standardsprachen. Theoretische, methodische und empirische Zugänge.* Wiesbaden: Harrassowitz. 19–37.

Wingender, Monika. 2018. Diskursive Konstruktion von Sprachenvielfalt und Sprachenkonflikten in Russland – Sprachenpolitik im Kontext von nationaler Sicherheit. In Sebastian Kempgen, Monika Wingender & Ludger Udolph (eds). *Deutsche Beiträge zum 16. Internationalen Slavistenkongress Belgrad 2018*. Wiesbaden: Harrassowitz. 353–64.

Wingender, Monika, Ivana Barkijević & Daniel Müller. 2010. Korpuslinguistische Untersuchungen von Standardsprachenmerkmalen. Ein Beitrag zur vergleichenden Standardologie. *Zeitschrift für Slavische Philologie* 67(1), 125–61.

Zeller, Jan Patrick & Dmitri Sitchinava. 2020. The Russian language in Belarus and Ukraine. In Arto Mustajoki, Ekaterina Protassova & Maria Yelenevskaya (eds). *The Soft Power of the Russian Language: Pluricentricity, Politics and Policies*. London, New York: Routledge. 108–22.

PART III

(Competing) Standards in formal education

SILVIA BALLARÈ

13 Competing standards in university students' formal production: A corpus-based study

ABSTRACT
In this paper, I aim to analyse formal texts written by Italian university students from a sociolinguistic perspective. More specifically, the linguistic nature of these texts with respect to the two different standards coexisting in the Italian scenario will be discussed; one standard is the old prescriptive and literary one, while the other is the so-called neo-standard which is the product of a process of downward convergence. Three linguistic features will be considered and their behaviour in three different corpora will be compared. More specifically, data will be extracted from a newly built resource consisting of university students' formal productions and from two sections of the CORIS corpus (i.e. academic prose and journalistic prose). Data show that university students' texts instantiate a very heterogeneous variety, which is distant from both 'standard' poles: while it is very prone to include some deviant linguistic features, it is reluctant to accept others.

1. Introduction[1]

Young speakers/writers are often considered to act as innovative agents in that they are prone to accepting and spread linguistic change. Sometimes, they are blamed for corrupting the prestigious 'good language' with their

1 This contribution is the result of the research carried out within the PRIN 2017 UniverS-Ita project Written Italian of university students: sociolinguistic framework, typological trends, didactic implications (ERC SH4, Prot. 2017 LAP429).

innovations. This is true also in the Italian scenario, where scholars started to address the issue some decades ago, focusing in particular on university students' formal productions (Lavinio & Sobrero 1991), where several substandard linguistic features were detected.

What complicates the picture is that at least from the 1980s onwards, two standards[2] have coexisted (see Cerruti, Crocco & Marzo 2017a for an overview) and thus, also nowadays, the presence of two models may lead to confusion, especially among young writers/speakers.

In this paper, university students' formal productions will be discussed from a sociolinguistic perspective, giving particular importance to the Italian *Architektur der Sprache* (Berruto 2012 [1987]: 23–30). Attention will be given to three linguistic features whose behaviour will be compared in university students' texts and in two corpora that exemplify the two competing standards.

After giving an account of the Italian sociolinguistic situation (Section 2), the data and the methods adopted for the analysis will be presented (Section 3). Section 4 is devoted to the data analysis, while Section 5 consists of some conclusive remarks.

2. Competing standards in contemporary Italian

In the Italian sociolinguistic scenario, two standards can be observed. The first one is the literary old standard (*italiano standard letterario*, in Berruto 2012 [1987]: 26), it is referred to as 'amended Florentine' (lit. *fiorentino emendato* in Galli de' Paratesi 1984: 10) and is the product of a standardisation process that started in the sixteenth century. It displays a number of literary features and is the variety on which the *language codices* are based: it is described by grammars and, at least ideally, taught in

2 In this paper, I will use the term *standard* not to designate an entire language but to refer to a standard variety (of a certain language), see Ammon (2004) for a thorough terminological discussion.

schools; for some authors, old-Standard Italian is 'actually crystalised in scholastic Italian at this point' (Antonelli 2011: 52, my translation). Thus, it is not to be considered as a Coserian norm (Coseriu 1971, inter alia), but rather as a prescriptive one. Nevertheless, it is important to stress that in contemporary Italian, this variety cannot be reduced to a mere abstract model, since it has been empirically proven that a cluster of its linguistic features occur coherently in the productions of educated speakers in highly formal communicative situations (Cerruti & Vietti 2022).

As mentioned before, alongside this old standard, a new one can be observed. Scholars started to notice this new variety in the 1980s (see Mioni 1983; Sabatini 1985; Berruto 2012 [1987]), and it can be considered as one of the outcomes of a major change that took place in the Italian sociolinguistic setting which started after World War II. As a matter of fact, it was only from the 1950s that (standard) Italian spread among a vast portion of the population that, until then, had spoken almost exclusively Italo-Romance dialects (De Mauro 1963, 2014). Italo-Romance dialects are not to be considered geographic varieties of Italian but rather, following Coseriu's (1980) taxonomy, 'primary dialects' and thus sister languages of Florentine on which the standard is based. That is to say that they, as Florentine, derive from Latin and 'evolved in parallel with Italian, and may display a noticeable degree of structural distance from it' (Cerruti, Crocco & Marzo 2017a: 5; see also Ledgeway & Maiden 2014; Maiden & Parry 2014).

The rapid spread of the standard language across contexts in which previously only Italo-Romance dialects were spoken and the increase in its use as a principal communication code led to the emergence of new sociolinguistic varieties, including the new standard. Over decades, some structural features originally marginalised to spoken informal productions began to be used also in formal exchanges and among educated speakers and therefore changed their sociolinguistic status. In order to give an account of this 'downward' convergence process ('restandardisation', *ristandardizzazione* in Berruto 2012 [1987]: 67; see also Berruto 2017), the term *neo-standard Italian* (Berruto [1987] 2012, see also Cerruti, Crocco & Marzo 2017b, Ballarè 2020) was coined. *Neo-standard Italian* is a 'new social norm, representing the shape of the Italian language as it is normally spoken and written by fluent, educated speakers today' (Berruto 2017: 32). The process that led to the emergence of this new norm can be explained by referring to the notion of *demotisation* (Coupland & Kristiansen 2011; Auer & Spiekermann 2011; see also *Demotisierung* in Mattheier 1997): 'the "standard ideology" as such stays intact while the valorisation of ways of speaking changes' (Coupland & Kristiansen 2011: 28). Neo-standard Italian mainly consists of a set of linguistic features that are to be considered as 'standard by usage' (Ammon 2003: 17): it is modelled on texts produced by speakers and writers considered to be prestigious by the members of the sociolinguistic community but is not codified. For this reason, some scholars have approximated neo-standard Italian to *journalistic Italian* (Antonelli 2011: 52), as opposed to the old scholastic one. It is important to emphasise that these productions display a set of linguistic features that are 'non-standard' in all respects and, thus, are perceived as 'errors', and stigmatised by some language experts and codifiers. In the Italian *Architektur der Sprache* this new standard did not replace the old literary one but rather coexists with it (Antonelli 2011: 51; see also Berruto 2012: 24).

2.1. University students' formal productions: between two standards

The above-described scenario results in educated speakers having two competing standards at their disposal as a model for their (highly and moderately) formal productions. This potential conflict might be more evident for younger speakers/writers who have been more exposed to the input of the new standard and might be less aware of the sociolinguistic characterisation of its linguistic features. University students have often been considered to be among those responsible for the alleged corruption of the 'good Italian language'. For example, the public debate on this issue was particularly lively in 2017, when a group of six hundred scholars and intellectuals co-signed an open letter (the so-called *Lettera dei Seicento*, lit. 'Letter of the six hundred') denouncing the lack of linguistic competence of university students and stating that three-quarters of undergraduates are to be considered 'semi-illiterate' (see De Santis & Fiorentino 2018 for a discussion).

However, these aspects have rarely been addressed from a sociolinguistic perspective, despite young educated speakers having a significant impact in defining which variety is to be considered prestigious and standard, both in the present and future. Linguists have often given an account of the use of non-standard linguistic features in university students' formal productions (see Lavinio & Sobrero 1991, and, more recently, Restivo 2022 for an overview), highlighting that 'the linguistic competence of university students is a serious problem, present – in varying degrees – in all universities and faculties'[3] (Sobrero 1991: 2, my translation). Nevertheless, these productions have not (or hardly ever) been discussed from a sociolinguistic perspective, also taking into account the characteristics of the sociolinguistic scenario and the coexistence of two standards. The issue of the consequences of the presence of these two competing standards may have for young educated speakers has never been addressed organically.

3. Data and methods

In the remainder of this paper, the sociolinguistic characterisation of university students' written and formal productions will be discussed, with respect to the old and the new standard; more specifically, the distribution of a set of three morphological and morphosyntactic features will be considered (see Ballarè 2020 for an overview). Data extracted from three corpora, which can be considered good examples of the three scrutinised varieties, will be analysed. The first corpus is a recently built

3 '[la] competenza linguistica degli studenti universitari è un grave problema, presente – in vario grado – in tutte le sedi universitarie e in tutte le facoltà'.

resource (Grandi et al. 2023), consisting of over 2,000 texts written by a representative sample (in terms of geographic location of the university and disciplinary area of the degree course) of students enrolled in the second year of bachelor's degree programmes. Data was collected during the academic year 2020–2021, and selected students were asked to write a short formal text (maximum 500 words), expressing their opinions about distance learning. The corpus has a rather small dimension (895,332 tokens), but, thanks to the adopted methodology (in terms of sampling and uniformity of the assigned task), it is highly representative. The second and the third corpora are two monitor sections (2017–2020) of the CORIS corpus (<https://corpora.ficlit.unibo.it/coris_ita.html>, last accessed 8 June 2024). One consists of journalistic texts (5,468,847 tokens) that, as mentioned before, are considered a prototypical example of neo-standard Italian in the literature (see Antonelli 2011, inter alia); the latter is made of papers and books written by scholars from various academic fields (1,707,703 tokens), and it can be considered a good example of the old standard because formal productions of highly educated speakers are the ones in which the linguistic features of the old-standard Italian have been attested in a cohesive and coherent way (Cerruti & Vietti 2022).

The aim of this study is to compare these three varieties in order to understand the nature of the university students' formal productions with respect to the two standards.

In particular, the analysis focuses on three linguistic features that can occur in a standard and (at least) a non-standard way, considering the latter typical of neo-standard texts.

Two features involve the selection of third-person singular pronouns (subject and oblique[4]), which is a point of great instability in the Italian linguistic system: as a matter of fact, they both display a complex paradigm that is going through a process of simplification in different varieties of Italian. More specifically, as far as subject pronouns are concerned, in Standard Italian there is a set of four forms that must be selected

4 In this paper, *oblique* is intended as opposed to *subject* and thus to include all the other syntactic functions (such as *direct* and *indirect objects* and *genitive*).

according to the gender and the semantic trait [+/- human] of the antecedent. If the antecedent is non-human, then the same pronouns can be employed also to express an oblique function; if it is human, the selection involves another set of pronouns differentiated for gender. Furthermore, there are some syntactic contexts (i.e. post-verbal position, coordinated clauses and focalisation) in which, even if the pronoun covers the syntactic function of the subject, the oblique form is required (see Cardinaletti 2021 inter al.). The whole picture is summarised in Table 13.1.

From a sociolinguistic perspective, *egli* and *ella* are perceived to be literary and obsolete by speakers and are almost absent in contemporary spoken productions (see Ballarè 2022). At the same time, *lui* and *lei* expanded their domains and began to be used in contexts formerly belonging to *egli* and *ella*. This overextension is far from being recent and is rather considered to be one of the key topics in the history of the Italian language: suffice to say that the issue was already addressed in the *Prose della volgar lingua* ('writings on the vulgar tongue') written by Pietro Bembo in 1525, which is considered to be the very first step of the standardisation process of Italian.

In the contemporary scenario, the employment of *lui*/*lei* as subject pronouns has been systematically counted among the key features of neo-standard Italian (Sabatini 1985: 159, 1990: 96; Berruto 2012 [1987]: 83–4). In these texts, *lui*/*lei* are occasionally used also to refer to a non-human antecedent, as has been noted. At the same time, *esso* and *essa* are sporadically employed when referring to a human antecedent. The whole paradigm seems to be involved in a variation process: case distinction between subject and oblique pronouns is nearly lost, and the semantic distinction between human and non-human antecedent is weakening.

Table 13.1 Third-person singular pronouns in Standard Italian.

	subject		obliques	
	+ hum	- hum	+ hum	- hum
M	*egli*	*esso*	*lui*	*esso*
F	*ella*	*essa*	*lei*	*essa*

The second feature linked with third-person singular pronouns involves singular dative clitics. Prescriptive grammars state that there is a form (*gli*) that must be used for male antecedents and another one (*le*) that is to be selected for female antecedents. However, it has been noted that in several varieties of Italian, *gli* is often used also with feminine antecedents. Furthermore, the masculine singular *gli* is often employed instead of the plural form *loro*, which is not marked by gender (Berretta 1985; Sabatini 1985: 158, 1990: 91–4, 97; Berruto 2012 [1987]: 84). That is to say that, at least in some varieties of Italian, the dative clitic *gli* can be used as a sort of 'general dative', with singular (feminine and masculine) and plural antecedents.

The last feature is the alternation between the indicative and the subjunctive mood in subordinate clauses introduced by a factive governor, namely *il fatto che* ('the fact that'). Even though grammars are not always clear about the correct mood choice for this context, subjunctive is systematically and explicitly described as conveying a meaning of uncertainty; as a consequence, factive contexts should arguably select the indicative.

Nevertheless, scholars have recently noticed an overextension of the subjunctive mood in subordinate clauses introduced by factive governors in non-standard varieties (Renzi 2012, 2019; see also Cerruti & Ballarè 2023). This expansion appears to be in line with the grammaticalisation process involving the subjunctive, which is leading to 'lexical routinization [...] and structural conventionalization' (Poplack et al. 2018: 217; see also Digesto 2019, 2021). It should also be mentioned that, in general, Italian speakers are well aware of the fact that the subjunctive mood should be preferred over indicative in formal contexts (see Digesto 2021): this sociolinguistic characterisation could help us understand why this pattern of variation is attested in moderately formal productions, whereas it is very rare in lower varieties.

These three linguistic features have been selected because they show substantial differences in their sociolinguistic distribution and characterisation, even though all three involve neo-standard Italian.

As mentioned before, the first two can be considered outcomes of a simplification process, despite displaying differences. The first feature concerns literary forms (i.e. *egli* and *ella*) that are perceived as such by the speakers and have thus been almost completely replaced by *lui/lei* in neo-standard Italian. The second is very frequently attested in lower varieties and, only sporadically, in the higher ones; however, it involves a process that takes place below the level of speaker's awareness. Finally, the over-extension of the subjunctive mood may have to do with its grammaticalisation process but also with being generally considered to be more appropriate in formal contexts. In other words, following the Labovian taxonomy (Labov 1994), we could say that while the first two linguistic features can be considered 'changes from below', because they are attested (also and originally) in lower varieties, the latter can be treated as a 'change from above', in that it can be observed only in (medium) high varieties. The whole picture must be kept in mind in the analysis phase.

In order to analyse the distribution and the behaviour of these three linguistic features in the three corpora, data were extracted as summarised in Table 13.2:

Table 13.2 Extracted data.

	extracted forms	university students' texts	CORIS corpus (journalistic and academic prose)
third-person subject pronouns	*egli*, *ella*, *esso*, *essa*, *lui*, *lei*	all occurrences	maximum 100 (randomised) occurrences per form
dative clitics	*gli*		
il fatto che + subj./ind.	*il fatto che*		

Finally, data were manually cleaned, and every occurrence was annotated according to its function and its being *standard* or *non-standard*. The only options that are accepted by prescriptive grammars were labelled as *standard*, while all the others, regardless of their frequency, were classified as *non-standard*.

4. Analysis

This section is devoted to data analysis. A short paragraph is dedicated to each linguistic feature, while an overall look at the university students' formal productions is provided in the last one.

4.1 Third-person subject pronouns

The first linguistic feature that is taken into consideration is third-person singular subject pronouns. Before moving on to the analysis of the distribution of their functions, some light will be shed on their sociolinguistic characterisation by discussing their (normalised) frequencies per million tokens in the three corpora. Data are shown in Figure 13.1.

First of all, not surprisingly, *egli* and *ella* are the less frequent pronouns. *Egli* is attested almost exclusively in the old standard, while it is quite rare both in neo-standard and in university students' texts. *Ella*

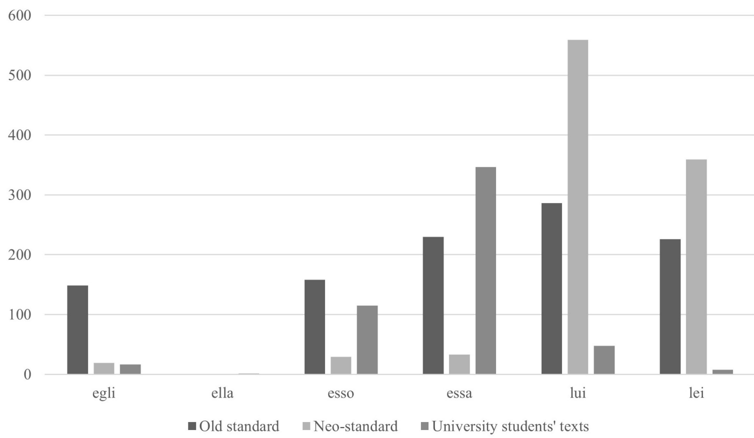


Figure 13.1 Third-person subject pronouns – frequencies (normalised per million tokens).

is nearly absent in all three corpora. *Esso* and *essa* are more frequent in the old standard than in the new one. As far as the masculine form is concerned, university students show a value that is placed almost in the middle between the other two; the high frequency of *essa* may be due to students very often referring to a feminine noun that was included in the assignment and is the key-topic of every text, i.e. *didattica a distanza* ('distance learning'). As expected, *lui* and *lei* are by far the most frequent forms in neo-standard texts, while their values are lower in the old one and, especially, in university students' texts.

Let us now consider the behaviour of the scrutinised forms in the three corpora. Data are shown in Table 13.3. As mentioned before, all the occurrences were manually annotated and, thus, we excluded cases in which the form did not behave as a third-person singular pronoun; for example, there were five cases in which *ella* was employed as a proper noun (as in *Ella* Fitzgerald) and two cases in which *lei* was selected as a courtesy form.

Table 13.3 Third-person subject pronouns – distribution.

	university students		neo-standard		old standard	
	standard	non-standard	standard	non-standard	standard	non-standard
egli	15 (93.75%)	1 (6.25%)	97 (97%)	3 (3%)	100 (100%)	0 (0%)
ella	1 (100%)	0 (0%)	3 (100%)	0 (0%)	10 (100%)	0 (0%)
esso	94 (91.26%)	9 (8.74%)	99 (99%)	1 (1%)	100 (100%)	0 (0%)
essa	309 (99.68%)	1 (0.32%)	98 (98%)	2 (2%)	100 (100%)	0 (0%)
lui	32 (74.42%)	11 (25.58%)	52 (52%)	48 (48%)	53 (53.54%)	46 (46.46%)
lei	4 (50%)	4 (50%)	42 (53.85%)	36 (46.15%)	54 (55.10%)	44 (44.90%)

Globally, the distribution indicates that, as expected, *non-standard* uses are more frequent when *lui* and *lei* are involved, in that they are often employed as subject pronouns. However, as mentioned before, even though this function is not prescribed by *ancien régime* grammars, it is widely accepted even in highly formal texts that should be an example of the old standard. If we take into account *egli/ella* and *esso/essa*, we can note that, at least in neo-standard and in university students' texts, there are some deviant uses, especially in the latter.

In order to provide a deeper understanding of the situation, there are some examples worth discussing. In neo-standard texts, *egli* is used two times in a post-verbal position and once with reference to a non-human antecedent, as shown in example (1). By contrast, in the non-standard occurrence of *egli* in university students' texts, the pronoun behaves as an oblique, as shown in example (2).

(1) C'è di peggio e di più, perché un programma nucleare iraniano fuori controllo va a incidere sulla sicurezza di Israele, che con un solo colpo atomico nella parte geograficamente più stretta del Paese potrebbe essere messo fuori causa pur disponendo ***egli stesso*** (mai negato, mai ammesso) di armamenti nucleari.

'There is more and worse, because an out-of-control Iranian nuclear programme affects the security of Israel, which with a single atomic strike in the geographically narrowest part of the country could be put out of action even though ***it itself*** has (never denied, never admitted) nuclear weapons'.

(2) A distanza vengono meno le interazioni con i professori, in quanto non si instaura più un colloquio diretto ***con egli*** ma sempre tramite uno schermo.

'Remotely, interactions with professors cease, because direct conversations are no longer established ***with him*** but always through a screen'.

Esso and *essa* are sporadically used with a human antecedent, both in the neo-standard example (3) and in university students' texts, see example (4). In order to explain these uses, one could say that this is a way of avoiding the choice between *egli/ella*, which are perceived to be highly formal, and *lui/lei* that, conversely, can be considered too informal. To support this hypothesis, it is worth mentioning that, in contemporary Italian and especially in formal production, examples of other strategies employed to avoid this selection, such as demonstrative pronouns (especially *questo/questa* 'this') used as personal pronouns are not rare. Interestingly, in one case, *esso* is used in the place of the demonstrative pronoun *questo* ('this') as an anaphoric encapsulator, see example (5).

(3) Si era sposato da poco con *Clara Wieck*, figlia del suo maestro di pianoforte e grandissima pianista ***essa*** stessa.

'He had recently married *Clara Wieck*, daughter of his piano teacher and a great pianist ***herself***.

(4) Anche per quanto riguarda il rapporto con *il professore*, per quanto ***esso*** sia bravo, non è come averlo fisicamente in aula.

'Even with regard to the relationship with *the professor*, no matter how good ***he*** is, it's not like having him physically in the classroom'.

(5) Tra i benefici della didattica online possiamo includere i seguenti: dormire un'ora in più al mattino, seguire l'intera lezione in pigiama, *bere una tazza di tè o mangiare mentre il professore spiega* (***esso*** è inammissibile in presenza: non solo è vietato mangiare all' interno dell'aula, ma risulterebbe maleducato da parte di chiunque mangiare durante la spiegazione!).

'Among the benefits of online teaching, we can include the following: sleeping an extra hour in the morning, following the entire lesson in pyjamas, *drinking a cup of tea or eating while the professor explains* (***this*** is unacceptable in presence: not only is eating inside the classroom forbidden, but eating during the explanation would be rude for anyone!)'.

4.2 Dative clitics

In this second section, we focus on the behaviour of the dative clitic *gli*. The distribution of its values is shown in Table 13.4.

In this case, as is shown, there is a relevant difference between the values attested in the old and neo-standard texts and the one detected in university students' productions, in that in the latter, the non-standard uses are by far more frequent. Furthermore, considering the functions covered by *gli* in the three corpora is relevant for the analysis. The whole picture is portrayed in Table 13.5.

Table 13.4 Dative clitics – distribution.

	university students		neo-standard		old standard	
	standard	non-standard	standard	non-standard	standard	non-standard
gli	50 (52.08%)	46 (48.92%)	98 (98%)	2 (2%)	97 (97%)	3 (3%)

Table 13.5 Functions of *gli*.

syntactic function	dative		direct object
gender/number of the antecedent	F.SG	M/F.PL	M.PL
university students	2	38	6
neo-standard	0	2	0
old standard	0	3	0

As shown, in the five non-standard occurrences attested in journalistic and academic prose, *gli* behaves like a plural dative and thus replaces *loro*. What is striking is that in university students' productions, *gli* covers this function in almost 40% of the cases. An example is provided in example (6), where the author correctly selects *loro* the first time and then, after a few words, opts for *gli*. Furthermore, it is only in this corpus that *gli* sporadically replaces *le* (dative feminine clitic), as shown in example (7), and, more surprisingly, *li* (plural dative object clitic), as in example (8). This last case, which is easier to observe in written than in oral production due to the phonetic similarity of the two elements (i.e. [ʎi] and [li]), has traditionally been considered to have a low sociolinguistic characterisation.

(6) *Molti studenti* grazie alla DAD sono riusciti a laurearsi nei tempi previsti, per chi si trova fuorisede sicuramente questa tipologia di didattica è stata molto utile, evitando ***loro*** spostamenti anche di lunga durata dalle loro abitazioni, ***gli*** ha permesso di seguire i corsi e procedere al sostenimento degli esami.

'Thanks to the distance learning, *many students* have managed to graduate on time, for those who are away from home this type of teaching has certainly been very useful, avoiding ***them*** even long-term travel from their homes, it has allowed ***them*** to follow the courses and proceed to take their exams'.

(7) Quest' anno, durante il periodo di pandemia ho assistito alle lezioni tramite *didattica a distanza* e sostenuto esami. Un vantaggio che ***gli*** si può attribuire è sicuramente la possibilità di poter rivedere le lezioni tramite la loro registrazione.

'This year, during the pandemic period, I attended lessons via *distance learning* and took exams. An advantage that can be attributed ***to it*** is certainly the possibility of being able to review the lessons from recordings'.

(8) *Gli esami* orali a distanza ***gli*** trovo meno difficili, questo non riguarda il carico didattico, ma è soltanto il fatto che mi vergogno meno.

'I find remote oral exams less difficult (lit. *Remote oral exams* I find ***them*** ...), this doesn't concern the study load, but it's just the fact that I'm less embarrassed'.

4.3 Il fatto che + *subjunctive/indicative*

This section is dedicated to the alternation between the subjunctive and the indicative mood in subordinate clauses introduced by the factive governor *il fatto che* ('the fact that'). The behaviour of the investigated feature is summarised in Table 13.6.

In this case, the non-standard uses (i.e. the use of the subjunctive mood) are consistently attested in the three corpora. However, it is possible to note that, while in university students' productions and in old-standard texts the non-standard occurrences are less than 30% of the total, in neo-standard texts this value exceeds 50%. An example of these non-standard uses is provided in (9).

Table 13.6 *Il fatto che* – distribution.

	university students		neo-standard		old standard	
	standard	non-standard	standard	non-standard	standard	non-standard
Il fatto che + subj/ind	79 (78.22)	22 (21.78)	46 (46%)	54 (54%)	71 (71%)	29 (29%)

(9) *Il fatto che* la pandemia ***sia avvenuta*** nel 2020, o comunque in un momento storico di forte sviluppo tecnologico, ha facilitato lo svolgimento di alcune attività che sarebbero state impossibili da svolgere in altre epoche, come appunto andare a scuola o all' università.

'*The fact that* the pandemic ***took place*** in 2020, or in any case in a historical moment of strong technological development, has facilitated carrying out some activities that would have been impossible to carry out in other eras, such as going to school or university'.

It is worth noting that, in addition to the sociolinguistic characterisation of the subjunctive, considered to be more suitable in formal contexts, the information structure of the sentence has been argued to be one of the reasons behind the overextension of the subjunctive in these cases (see Cerruti & Ballarè 2023 for a discussion). In fact, the subjunctive mood is more frequent in clauses whose content is thematic and given (as in the provided example): the subjunctive ensures that the complement clause is put in the background, while the content of the matrix clause – i.e. the speaker's stance – is foregrounded.

4.4 *An overall look*

In Table 13.7, the percentages of non-standard uses of some of the analysed linguistic features in the three corpora are reported. Some of them (i.e. *ella* and *lui/lei*) were excluded because they are not relevant to highlight the differences of the three varieties being analysed.

The first thing that can be noticed is that university students seem to be prone to accepting linguistic features that are considered 'changes

Table 13.7 Percentages of non-standard features.

	egli	*esso*	*essa*	*gli*	*il fatto che*
university students	6.25%	8.74%	0.32%	46%	22%
neo-standard	3.00%	1.00%	2.00%	2%	54%
old standard	0	0	0	3%	29%

from below'. The percentages of non-standard uses of third-person subject pronouns are almost always higher in these productions than in the other corpora. However, significant differences must be highlighted. In fact, frequencies are quite low with respect to non-standard uses of *egli*, *esso/essa* and *gli*. In this case, deviant uses are sporadic and students' behaviour is still quite conservative. The use of these pronouns can be explained taking into account the already discussed sociolinguistic markedness of the forms. *Egli*, as mentioned before, is perceived to be a literary form, while *esso/essa* are considered to be more suitable than *lui/lei* in formal contexts. As a result, they are sporadically overextended to functions that they are not intended to cover. If we now move to *gli*, it is strikingly not only the university students' value that is very high (46%) with respect to the other two (2% and 3%), but the range of functions that are covered by this form in these texts is widening. Students are less aware of the sociolinguistic characterisation of these features and the simplification of the paradigm here seems to be at a more advanced stage.

Finally, if we consider the only 'change from above', students surprisingly behave in a conservative way, in that their productions show percentages of 'non-standardness' lower even than the old-standard ones. In this case, the neo-standard texts display non-standard features more frequently.

Globally, the behaviour of these three features conveys an image of high heterogeneity: university students' productions are both conservative (see *il fatto che*), but also prone to accepting deviant features (*pronouns*).

5. Concluding remarks

The focus of this paper has been the comparison between university students' formal productions and the two competing standards that have been detected in the Italian sociolinguistic scenario. The analysis of the behaviour and the distribution of the three scrutinised linguistic features has revealed a rather complex picture. However, it must be borne in mind that, since the discussion is based on a rather small set of linguistic features, further studies may be required in order to provide a deeper understanding of the entire situation.

If we consider 'changes from below', university students' texts display more deviant values. In one case, it can be explained as an attempt to raise the formality of the productions: in order to avoid the selection of *lui/lei,* students prefer forms that are perceived to be more suited to the context, but they use them in a deviant way (i.e. *egli* as an oblique and *esso/essa* with human antecedents). In the other case, the dative clitic *gli* expands massively its functional domain as it is used with feminine and plural antecedents, and even with another syntactic function (i.e. direct object), as observed in considerably lower varieties.

Alongside this pattern, which places university students' productions below not only the old but also the new standard, there are cases in which these texts are more conservative; surprisingly, that is the case with the so-called 'changes from above' (i.e. *il fatto che* + subjunctive).

Overall, the scrutinised texts are hard to profile from a sociolinguistic perspective owing to the different characterisation of the detected linguistic features. This variety can be considered a sort of 'third way': students do not adapt to either of the two linguistic models. Over time, it will be possible to establish whether the current situation, which sees two competing standards coexisting, will result in a new and more stable scenario with a new linguistic model that differs from both of the old ones.

Bibliography

Ammon, Ulrich. 2003. On the social forces that determine what is standard in a language and on conditions of successful implementation. *Sociolinguistica* 17, 1–10.

Ammon, Ulrich. 2004. Standard variety. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds). *Sociolinguistics. An International Handbook of the Science of Language and Society*. Vol. 1. 2nd edn. Berlin, New York: Mouton de Gruyter. 273–83.

Antonelli, Giuseppe. 2011. Lingua. In Andrea Afribo & Emanuele Zinato (eds). *Modernità italiana. Cultura, lingua e letteratura dagli anni Settanta a oggi.* Roma: Carocci. 15–52.

Auer, Peter & Helmut Spiekermann. 2011. Demotisation of the standard variety or de-standardisation? The changing status of German in late modernity (with special reference to southwestern Germany). In Tore Kristiansen & Nikolas Coupland (eds). *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus. 161–76.

Ballarè, Silvia. 2020. L'italiano neo-standard oggi: stato dell'arte. *Italiano LinguaDue* 12(2), 469–92.

Ballarè, Silvia. 2022. Lo scritto formale di studenti universitari tra vecchio e nuovo standard. Il caso dei pronomi di terza persona singolare. *Studi Italiani di Linguistica Teorica e Applicata (SILTA), anno LI* (3), 362–84.

Berretta, Monica. 1985. I pronomi clitici nell'italiano parlato. In Günter Holtus & Edgar Radtke (eds). *Gesprochenes Italienisch in Geschichte und Gegenwart.* Tübingen: Narr. 117–33.

Berruto, Gaetano. 2012 [1987]. *Sociolinguistica dell'italiano contemporaneo*. Roma: Carocci (1987: Roma: La Nuova Italia Scientifica).

Berruto, Gaetano. 2017. What is changing in Italian today? Phenomena of restandardization in syntax and morphology: an overview. In Massimo Cerruti, Claudia Crocco & Stefania Marzo (eds). *Towards a New Standard. Theoretical and Empirical Studies on the Restandardization of Italian*. Berlin, Boston: de Gruyter. 31–60.

Cardinaletti, Anna. 2021. I pronomi deboli nel sistema pronominale dell'italiano. *Storie e linguaggi* 7, 2.

Cerruti, Massimo & Alessandro Vietti. 2022. Identifying language varieties: coexisting standards in spoken Italian. In Karen Beaman & Gregory R. Guy (eds). *The coherence of linguistic communities: Orderly heterogeneity and social meaning*. London: Routledge. 261–80.

Cerruti, Massimo, Claudia Crocco & Stefania Marzo. 2017a. On the development of a new standard norm in Italian. In Massimo Cerruti, Claudia Crocco & Stefania Marzo (eds). *Towards a New Standard. Theoretical and Empirical Studies on the Restandardization of Italian*. Berlin, Boston: de Gruyter. 3–28.

Cerruti, Massimo, Claudia Crocco & Stefania Marzo (eds). 2017b. *Towards a New Standard. Theoretical and Empirical Studies on the Restandardization of Italian*. Berlin, Boston: de Gruyter.

Cerruti, Massimo & Silvia Ballarè. 2023. Sociolinguistic variation, or lack thereof, in the use of the Italian subjunctive: mood selection under factive and semi-factive governors. *Sociolinguistica* 37(1), 75–93. [Special issue *Sociolinguistic*

Variation in Contemporary Spoken Italian, ed. by Silvia Ballarè & Massimo Cerruti.]

Coseriu, Eugenio. 1971. Sistema, norma e 'parole'. In Eugenio Coseriu. *Teoria del linguaggio e linguistica generale: sette studi*. Bari: Laterza. 19–103.

Coseriu Eugenio. 1980. Historische Sprache und Dialekt. In Joachim Göschel, Pavle Ivić & Kurt Kehr (eds). *Dialekt und Dialektologie*. Wiesbaden: Steiner. 106–22.

Coupland, Nikolas & Tore Kristiansen. 2011. SLICE: Critical perspectives on laguage (de)standardisation. In Tore Kristiansen & Nikolas Coupland (eds). *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus. 11–35.

De Mauro, Tullio. 1963. *Storia linguistica dell'Italia unita*. Roma, Bari: Laterza.

De Mauro, Tullio. 2014. *Storia linguistica dell'Italia repubblicana dal 1946 ai giorni nostri*. Roma, Bari: Laterza.

De Santis, Cristiana & Giuliana Fiorentino. 2018. La carica dei 600: la campagna mediatica sul declino della lingua italiana. *CIRCULA* 7, 1–28.

Digesto, Salvio. 2019. *A Variationist Analysis of Subjunctive Variability across Space and Time: from Contemporary Italian Back To Latin*. Ottawa: University of Ottawa PhD dissertation.

Digesto, Salvio. 2021. Lexicalization and social meaning of the Italian subjunctive. *Cadernos del Linguistíca* 2(3). E609. <https://cadernos.abralin.org/index.php/cadernos/article/view/609> (accessed 30 November 2024).

Galli de' Paratesi, Nora. 1984. *Lingua toscana in bocca ambrosiana*. Bologna: Il Mulino.

Grandi, Nicola, Silvia Ballarè, Francesca Chiusaroli, Francesca Gallina & Matteo Pascoli. *Corpus Univers-ITA* <https://corpora.ficlit.unibo.it/CUSP> (accessed 30 November 2024).

Labov, William. 1994. The study of change in progress: Observations in real time. In William Labov. *Principles of Linguistic Change*. Vol. 1: *Internal factors*. Oxford: Blackwell. 73–112.

Lavinio, Cristina & Alberto A. Sobrero. 1991. *La lingua degli studenti universitari*. Firenze: La Nuova Italia.

Ledgeway, Adam & Martin Maiden (eds). 2014. *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press.

Maiden, Martin & Mair Parry. 2014. *The Dialects of Italy*. London, New York: Routledge.

Mattheier, Klaus J. 1997. Über Destandardisierung, Umstandardisierung und Standardisierung in modernen europäischen Standardsprachen. In Klaus J. Mattheier & Edgar Radtke (eds). *Standardisierung und Destandardisierung europäischer Nationalsprachen*. Frankfurt am Main, Berlin: Lang. 1–9.

Mioni, Alberto A. 1983. Italiano tendenziale: osservazioni su alcuni aspetti della standardizzazione. In A.A.V.V. (eds). *Scritti linguistici in onore di Giovan Battista Pellegrini*. Pisa: Pacini. 495–517.

Poplack, Shana, Rena Torres Cacoullos, Nathalie Dion, Rosane de Andrade Berlinck, Salvio Digesto, Dora Lacasse & Jonathan Steuck. 2018. Variation and grammaticalization in Romance: A cross-linguistic study of the subjunctive. In Wendy Ayres-Bennett & Janice Carruthers (eds). *Manual of Romance Sociolinguistics*. Berlin, New York: Mouton de Gruyter. 217–52.

Renzi, Lorenzo. 2012. *Come cambia la lingua. L'italiano in movimento*. Bologna, Il Mulino.

Renzi, Lorenzo. 2019. Ancora su come cambia la lingua. Qualche nuova indicazione. In Bruno Moretti, Alin Kunz, Silvia Natale & Etna Krakenberger (eds). *Le tendenze dell'italiano contemporaneo rivisitate. Atti del LII Congresso Internazionale della Società di Linguistica Italiana (Berna, 6–8 settembre 2018).* Milano: Officinaventuno. 13–33.

Restivo, Maria Laura. 2022. L'italiano scritto degli studenti universitari: prime osservazioni sul corpus Univers-ITA. *Italiano LinguaDue* (1), 797–818.

Sabatini, Francesco. 1985. L''italiano dell'uso medio': una realtà tra le varietà linguistiche italiane. In Günter Holtus & Edgar Radtke (eds). *Gesprochenes Italienisch in Geschichte und Gegenwart*. Tübingen: Narr. 154–84.

Sabatini, Francesco. 1990. Una lingua ritrovata: l'italiano parlato. *Studi latini e italiani* 4, 215–34.

Sobrero, Alberto A. 1991, Prefazione. In Cristina Lavinio & Alberto A. Sobrero. *La lingua degli studenti universitari*. Firenze: La Nuova Italia.

HEIMIR F. VIÐARSSON

14 Doubling down on the scope and shape of standard norms: The role of compulsory education (1907–) in changing the tone of Icelandic language policy

ABSTRACT
Icelandic standardisation has its main roots in the nineteenth century but took on a radically different guise in the twentieth century, mainly through changes in the educational system. As Ottósson (1990) has hypothesised, the Education Act of 1907, which introduced compulsory schooling for children aged ten to fourteen, is thought to have brought about important changes: a harsher and simplified tone in discussion about standard norms, as well as in the broadening of the scope and goals of standardisation towards the spoken modality. These changes then fed into Icelandic language policy more generally. Based on a critical discussion of Icelandic 'purism' and standardisation as well as a study of a dozen grammar books from the period 1878–1944, Ottósson's hypothesis finds at least partial support in the available data.

1. Introduction[1]

The relationship between language standardisation and the public education system is a very important one as much previous scholarship

1 Parts of this paper were presented at *Menntakvika* 2022, an annual conference on educational sciences at the University of Iceland – School of Education. I would like to thank the organisers, audiences, and anonymous reviewers for valuable questions and comments. The usual disclaimers apply.

has pointed out (e.g. Haugen 1987; Deumert & Vandenbussche 2003; Kristmannsson 2004; Vandenbussche 2007; Wright 2016; Möller 2017; Ingvadóttir 2021). Given this significance, Deumert & Vandenbussche (2003: 458) highlight that research in this domain often 'concentrates on the "long nineteenth century"' (c. 1789 to c. 1914) when mass alphabetisation and a general education system contributed not only to a significant increase in literacy levels but also facilitated access and exposure to the norms of the standard language. The present paper views standardisation primarily through an educational lens, at the point in time when the implementation of standard norms had become an orchestrated, nationwide effort and was, therefore, aimed at the whole population.

Icelandic scholarship has emphasised the limited role and extent of Icelandic grammar teaching until the latter half of the nineteenth century and the beginning of the twentieth century (e.g. Viðarsson 2019). Prior to this period, elementary-level teaching for children, other than the privileged few, had only been offered at a small-scale, both intermittently and in a small number of areas, as schooling was not made compulsory for all children until the twentieth century (see Guttormsson 1993, Rastrick 2003, Þorsteinsson 2012). It is this specific change that will be our main focus throughout the paper, in particular a hypothesis put forth by Ottósson (1990) regarding the interaction between the public educational system and language policy.

The introduction of compulsory schooling in Iceland by legislation in 1907 has been argued not only to have resulted in the implementation of standard norms at a much wider national level than had previously been the case. It has been suggested that these legislative changes likely also had a visible and lasting impact on the tone and shape of Icelandic language policy as a direct result of now being geared towards younger children.

> Gerbreyting varð síðan í fræðslumálum þegar lögfest var skólaskylda 10–14 ára barna 1907. Þar með náði hin opinbera stefna í málfarsefnum, um rétt mál og rangt, til allra barna á mótunarskeiði. Í barnaskólum þeim, sem áður störfuðu, var reyndar mjög lítil kennsla í íslenskri málfræði. Jafnframt virðist svo sem málstefnan hafi tekið nokkrum blæbrigðabreytingum um leið og hún beindist að öðrum aldurshópi en áður. Stefnan virðist hafa verið einfölduð og löguð að

> þroska nemenda með því að talað var meira og afdráttarlausara um rétt mál og rangt. Meira hefur verið farið að reyna að breyta talmáli nemenda, en áður var öll megináherslan á ritmálinu. Þetta forvitnilega mál er hins vegar lítt kannað.
>
> 'When compulsory schooling for ten to fourteen-year-old children was legislated in 1907, a radical change took place in education matters. As a result [of the compulsory schooling legislation], the official policy in matters of right and wrong language use reached all children during their formative period. In the children's schools that had operated previously, there had actually been very little instruction in Icelandic grammar. At the same time, it seems that language-planning policy underwent certain changes in nuance as soon as it was aimed at a different age group than before. The policy seems to have been simplified and adapted to the maturity of the students in that there was more talk of right and wrong [language] and in a more direct manner. Greater emphasis was also placed on changing the spoken language of the students, but all the main focus had previously been on the written language. This intriguing issue, however, is largely unexplored'. (Ottósson 1990: 104–5, my translation)

The aim of this paper is to explore this aspect of Icelandic language history in more detail. The hypothesis will be evaluated mainly through a closer look at the primary school grammars available from the late nineteenth century up until the middle of the twentieth century. However, specific language-planning efforts within the primary school system will also be highlighted against these perceived shifts in tone and shape of the language policy. Previous studies have primarily focused on the emergence of the national standard language during the nineteenth century, in particular the role of the Reykjavík Grammar School, belonging to the secondary school level, which – during most of the nineteenth century – had been the only school of instruction on linguistics and the grammar of Icelandic (Möller 2017; Viðarsson 2019). The present study will serve to fill this gap to some degree.

The research questions that will be addressed are the following:

- To what extent can a shift in tone be discerned in Icelandic grammar books before and after the legislation of compulsory schooling in 1907?
- How has Icelandic language policy and language planning in a historical perspective been affected by changes imposed through the education system?

The structure of the paper is as follows: Section 2 offers a general background to the standardisation of Icelandic. Section 3 homes in on a particular aspect of the standardisation process, the widening in focus towards the spoken language. Section 4 presents the results of a study of prescriptive remarks in elementary grammars from the late nineteenth until the mid-twentieth century, suggestive of a gradual increase in explicit negative remarks in partial support of Ottósson's hypothesis. Section 5 offers a brief summary of the main points, concluding the paper.

2. Standardisation

The standardisation of Icelandic, like that of many other languages, tends to revolve around influential figures, typically named entities that were vocal about the present and future form of the language (Ottósson 1990; Árnason 2003a, 2003b; Kristmannsson 2004; Viðarsson 2019; Ingvadóttir 2021). Much of the previous scholarship frames this discourse not under the heading of 'standardisation' but rather 'language purification' ('málhreinsun') and while there is a relative lack of any actual research demonstrating its effects (beyond strictly edited material), we are typically led to believe that standardisation was, in the grand scheme of things, a very successful endeavour. Before going any further into detail, let us first have a bird's eye view of the historical background of these efforts.

2.1 The early modern period and the prelude to standardisation

Ideas or ambitions towards Icelandic language purification are widely considered to go back at least to Arngrímur Jónsson's *Crymogæa* (1609) (see e.g. Benediktsson 1985: 32–3; Ottósson 1990: 20–1; Pálsson 1995: 127; Jensson 2008). As Jensson (2008: 4) observes: 'In Iceland around 1600, the printing of books in the vernacular seems to have required not only a degree of standardisation of usage and orthography, but also, more importantly, a valorisation of the language itself'. A vague notion of purism and/or standardisation continued to play a role in scholarly discourse over the next couple of centuries.

For instance, the language-critical writing of the poet, writer and naturalist Eggert Ólafsson (1726–1768) can probably be said to have accumulated in the Society of the Learned Arts (*Hið íslenska lærdómslistafélag*) in 1779, that would later merge with the Icelandic Literary Society (*Hið íslenska bókmenntafélag*), as well as in the establishment of the Icelandic Society for the Education of the Nation (*Hið íslenska landsuppfræðingarfélag*) in 1794 (see further Ottósson 1990; Ingvadóttir 2021). The purpose of the Society of the Learned Arts was to publish practical and scientific works in Icelandic 'and attempt to cleanse [Icelandic] of foreign words and locutions, that are now starting to spoil it', as stated in the original societal charter published in 1780 (Ottósson 1990: 42, my transl.). While the literary societies aimed towards language purism, the Society of the Learned Arts went further in the direction of reviving archaic traits, as had also been the style of Eggert Ólafsson (Ottósson 1990: 35–8; Ingvadóttir 2021: 72–4).

None of these efforts, however, were particularly orchestrated nor would it be appropriate to regard them as being anywhere on par with the standardisation of the nineteenth century onwards, when these efforts became more centralised. In his overview of the discourse, Kusters (2003: 184) reports that standardisation effectively led to 'halting and reversing the changes in Icelandic', i.e. changes that had occurred during the early modern era. Thomason (2001: 9) also writes that 'the creators of Standard Icelandic deliberately archaised the language's structure, making it look older so as to bring it closer to the language of the Eddas', and Haugen (1987: 74), similarly, claims that Icelandic was 'recodified from the half-Danicised language of the 1584 Bible by reference to the classic models from Old Icelandic'.

The main nineteenth-century codifiers of Icelandic were associated with 'the Grammar School', located in the capital region – outside Reykjavík (in Bessastaðir, the current residency of the president) between

1805–1846, but in the city centre from 1846 onwards. A central figure in this enterprise early on was Sveinbjörn Egilsson (1791–1852), a teacher and later rector of the Grammar School, whose students would form the Copenhagen-based society and like-named periodical *Fjölnir* (1835–1847). The periodical set a particularly strict standard for authors and publishers alike in that the *Fjölnir* members, in particular the two linguists in the group, Konráð Gíslason (1808–1891) and Halldór Kr. Friðriksson (1819–1902), would write relentless book reviews in which they would very harshly criticise various aspects of grammar of the work in question, such as word order, morphology and lexis.

An example of their style is a review of a book called *Ljóðasmámunir, samt Emilíu Raunir* ('Trifles of Poetry and The Woe of Emilia', published 1839), in which Gíslason (1843: 62) writes:[2] 'This is much too short a title because the book should actually be called: An insignificant hotchpotch of barbarisms, solecisms, Danish spatters, paddings, clumsy expressions, improprieties and other such trifles' following which well over a hundred examples are provided, collected 'at random' in those parts of the book that were allegedly 'least bad' (Gíslason 1843: 64). The Fjölnir reviews were, nonetheless, extremely influential and would later even be used for reference as teaching material in the Grammar School, when Friðriksson took on the duties of an Icelandic teacher – a position he held between 1848 and 1895 (Viðarsson 2019: 40, 46–7). These reviews caused quite some stir, so much so, in fact, that their mentor-inspirator, Sveinbjörn Egilsson himself, was a bit taken aback, writing that it was difficult to keep up their standard (Ottósson 1990: 72).

Fjölnir's zero tolerance for Danish loans also carried over to grammatical features, embedded word order being a case in point. The 'Danish-like' ('dönskuleg') custom of placing the finite verb after sentence adverbs and the negation was rejected in favour of the traditional Icelandic word order: lit. 'all that which we *not had* expected' ('allt það, er vér *ekki áttum*

2 'Þetta nafn er mikjils til of stutt, því bókjin ætti reíndar að heíta: "Lítilfjörlegur samtíníngur af málleísum, bögumælum, dönskuslettum, hortittum, klaufalegum orðatiltækjum, smekkleísum og öðrum þess háttar smámunum"' (Gíslason 1843: 62).

von á') for the proper 'all that which we *had not* expected' ('allt það, er vér *áttum ekki* von á') (Gíslason 1844: 85). The now stigmatised word order had become frequent, in some texts even predominant, but quickly started to dwindle until becoming a rarely seen feature in (edited) Icelandic texts towards and beyond the turn of the nineteenth century (Viðarsson 2019).

Much of the focus of Icelandic codifiers revolved around the 'undoing' of historical change and 'revival' of extinct traits with reference to medieval forms (see Ottósson 1990, 2005), a choice which was based on the idea of a golden era of a more glorious past of independence. The early modern period from roughly the mid-sixteenth to eighteenth century had been one of increased contact with the European mainland through greater centralisation of the Danish monarchy and stronger hold of the Danish administration in Iceland after the Reformation. (It is probably no coincidence that the word order variation shown in the preceding paragraph emerged during this very period.) Both in socio-historical and in linguistic respects, the early modern period has often been described in terms of 'decline' or even 'humiliation', culminating in the Danish trade monopoly on the one hand, abolished (in part) in 1787 (see e.g. Gunnarsson 1983), and in a greater use of Danish in official spheres as well as the further 'Danicisation' of the language on the other, at least in the upper echelons (see Hauksdóttir 2016: 81–3; Ingvadóttir 2021: 46–59). However, present-day historians have increasingly criticised or rejected this supposed Danish-inflicted deterioration (Ólafsson 1999; Kristinsson 2018) and the same goes for the Icelandic language which the Danes, including the Danish authorities, actually helped to preserve, if anything (Hauksdóttir 2016; Ingvadóttir 2021: 91–2).

The traditional discourse is very visible in Jón Þorkelsson's (1870) outline of standardisation ('purification') efforts during the period between 1830 and 1870, which is of particular interest because codifiers continued to disagree on how far one should go towards re-establishing medieval linguistic forms and patterns. In Þorkelsson's (1870: 82) view, the Icelandic language had reached a state of 'great humiliation' ('mikla niðrlæging') in the middle of the eighteenth century resulting in attempts to 'restore' and 'reshape' it ('endrreisa og endrskapa'). Jón Þorkelsson (1822–1904) was certainly not an insignificant person when it came to this topic as he was a head teacher at the Reykjavík Grammar School, and later rector. He argues for the position that much more can be achieved in regard to restoring the supposed former glory of Icelandic – his paper, in fact, being an exploration of modern deviations from medieval Icelandic, discussed under the heading of 'wrong' word forms and word orders. Throughout the paper, his definition of correctness is the language of the Sagas of Icelanders and other medieval manuscripts, i.e. Old Norse. While standard norms continued to be based on (or at least motivated by) medieval Icelandic, 'purification' efforts tied to the formation of the standard language gradually eliminated earlier tendencies towards excessive archaism (Ottósson 1990: 75).

Surprisingly, perhaps, Þorkelsson's appeal for a more archaic standard language was actively counteracted from within his own workplace, the Reykjavík Grammar School, by a fellow teacher of Icelandic, Halldór Kr. Friðriksson. Friðriksson's (1871) response shows how the archaicising ideals continued to spark controversy and debate, presumably as scholars were actually discovering that Old Norse and contemporary Icelandic were, in fact, distinct grammatical systems – more so than people then (and even to this day) cared to acknowledge in public. Friðriksson considered Þorkelsson's stance an extreme and referred to it not only in terms such as dogmatism and miswander but even considered it childish (see Viðarsson 2019: 35–6).[3] It appears that rector Þorkelsson later abandoned his apparent equivocation of historical change and corruption, offering no such evaluation, for instance, in his wholly descriptive study of diachronic changes in the Icelandic subjunctive (Þorkelsson 1887).

3 Unsurprisingly, perhaps, the two were not on good terms. In his private letters to Jón Sigurðsson, politician and president of the Icelandic Literary Society, Friðriksson vowed to resign if Þorkelsson were to become rector of the Reykjavík Grammar School (see Viðarsson 2019: 36, and references therein). Unlucky for Friðriksson, Þorkelsson did become rector but Friðriksson remained, as head teacher from 1874, until the two retired in 1895. Later, in Friðriksson's obituary, Þorkelsson (1903: 23) would remark that he was 'a loyal enemy to his enemies, fierce and inflexible' but 'true, dependable and yielding to his friends'. While the obituary praises Friðriksson in most respects, one cannot help but sense some resentment here.

In Smári's (1920: 12–14) syntax of contemporary Icelandic, the early modern period is described as one of increasing foreign effects on the written language and he speaks of the 'gobbledegook' ('hrognamál') of the educated elite up until the 'restoration' ('endurreisn') of the language from the mid-eighteenth century onwards. Similarly, Jón Þorkelsson forni ('the Elder', 1859–1924), not to be confused with Jón Þorkelsson the rector, describes the language of an eighteenth-century narrative of a 1727 natural disaster as being 'very awkward' and adds that the author (a clergyman) 'attempted to be as archaic as possible, but the resulting style is a pesky doggerel ('leiðindahnoð eitt') amid the snot taste ('horbragð') of the language that people generally wrote during the eighteenth century' (Þorkelsson 1918: 54, my translation).

It is insightful to consider in this context the differences between Icelandic on the one hand, sometimes referred to as 'Insular' Scandinavian, and 'Mainland' Scandinavian languages on the other. The two language groups differ greatly in terms of visible foreign effects, especially in terms of lexis. As Teleman (2010) has observed for Swedish, a member of the latter group of Scandinavian languages, there was limited spread of Low German import words into Swedish historically if one takes the different position of the rural dialects into account. The propagation of Low German imports, he argues, came about primarily through the weakening of rural dialects in more recent times, which were less exposed to foreign loans, as well as the establishment of a national standard language during the nineteenth century. As the standard language prevailed to some extent over the rural dialects, so did the imported features from the standard that originally had less influence in these dialects.

Similarly, Pettersson (1988) describes different choices in the canonical embedded word order of Swedish (which holds for Mainland Scandinavian in general) vis-à-vis Icelandic with reference to the different ways in which these languages were standardised. The basis of the Icelandic norms was mostly rural and decidedly not the language of the cultural elite, but the opposite held in Sweden (Pettersson 1988). While Icelandic scholarship has also pointed out the rural and non-elitist or even anti-elitist origins of the Icelandic standard, due mostly to the Grammar School and its spawns (in particular the *Fjölnir* movement), this discussion tends to blend into more ideologically laden statements about the exceptional conservatism of Icelandic, which in my view is greatly overstated (Viðarsson 2019). The differences in selection and codification practices between these language communities, however, serve to explain, at least in part, the vastly different shape or source of the standard and much less visible effects of multilingualism in (especially late) nineteenth-century Icelandic onwards, as opposed to early modern Icelandic.

2.2 Beyond the elite: Standardisation in modern times

The twentieth century stands in clear contrast to previous periods. While elementary literacy skills (both reading and writing) were already mandatory by the late nineteenth century (see Viðarsson 2019: 51–2), 'Icelandic' first appeared as a separate subject at the elementary level in the diaries of the Reykjavík Elementary School (*Barnaskóli Reykjavíkur*) in 1899 (cf. Rastrick 2003: 98). Some grammar was already being taught in children's schools in the late nineteenth century (Guttormsson 1993: 10), but more systematic grammar or 'mother tongue teaching' ('móðurmálskennsla') was not introduced until the early twentieth century and became a compulsory subject in the Education Act of 1907.

The Education Act made elementary schooling compulsory for children aged ten to fourteen years, and this was also the first time that a special provision on the teaching of 'correct' language appeared for this age group. As Ottósson (1990: 105) observed, the twentieth century marked a shift in the scope of standard norms, both in terms of the modality (written and spoken language) and in that grammar books became more explicit and less nuanced to better fit the psychological development of children. Ottósson (1990) maintained that the standard norms now not only reached the whole population more easily during this formative period, but that the changes in 'tone' of grammatical prescriptions that this brought about also influenced Icelandic language policy in general.

The determination of the shape of the standard language, selection of norms and subsequent codification (Haugen 1987) obviously does not automatically lead to widespread effects across the language community at large. Well aware of this fact, of course, Haugen (1987: 61) points out that standard norms are 'relatively simple to implement' at a restricted scale when 'a small, elite group has a monopoly on education', but once we start to operate on larger scales, it becomes 'a major educational issue' (Kusters 2003: 184 makes a similar point). Research where common targets of standardisation have actually been explored empirically reveals that 'reality' tends to be quite different from the standard narrative.

This small detail often tends to be conveniently ignored, or, as stated more recently by Leonard & Árnason (2011: 94): 'When it comes to defining "non-standard usage", the myth has been that there is no such thing'. For exactly these reasons, the theoretical literature sometimes describes standardisation as 'hallucinations' (Elspaß 2014: 310, with references) or 'fiction': 'Standardisation is in part a fiction. We have imagined languages in the same way that we have imagined communities' (Wright 2016: 60). In the interest of space, only a few examples from the history of Icelandic standardisation and language-planning efforts can be presented here. Nonetheless, these selected cases should serve to illustrate the mechanisms and motivations underlying the standardisation process.

At least since Jónsson (1900), a change involving the spread of dative case to former accusative experiencers of verbs such as *langa* 'to want' and *vanta* 'to need, to lack' has been described as a 'language blemish' ('málllýti'). In the relevant literature this variation usually goes under the heading of 'Dative Sickness' or 'Dativitis' ('þágufallssýki', see Árnason 2003b: 201–3). According to various studies, despite having been actively counteracted in the schooling system for most of the twentieth and twenty-first centuries, making a frequent appearance on standardised grammar tests and vanishingly rare in edited material, its use only continues to increase (e.g. Svavarsdóttir et al. 1984; Jónsson & Eyþórsson 2003; Svavarsdóttir 2013).

While Icelandic standardisation in principle does not discriminate much between the various linguistic levels, most of it – especially in school grammars – targets the declination of irregular nouns. A study by Heimisdóttir (2008) re-evaluated a frequently made claim in the literature on the supposed revival of the medieval inflection of so-called *ija*-stems, i.e. words like *læknir* 'doctor', *kíkir* 'binoculars' and *hirðir* 'shepherd', also affecting many proper names like *Fjölnir*, *Reynir* and *Víðir*, where the nominative suffix *-r* was reanalysed as belonging to the stem (e.g. Benediktsson 1969). By studying nineteenth- and early twentieth-century private letters, in addition to a questionnaire administered to both elementary and high school students, Heimisdóttir (2008) found that the supposedly revived inflectional pattern was not only still at least partly attested before the effects of standardisation would have been felt, it was also not as successfully implemented as often claimed, based on the twenty-first-century questionnaire data.

If the results of Heimisdóttir (2008) are suggestive of the effects of standardisation at the community level, which is not unlikely, this is obviously a considerable blow to the standard narrative, especially when considering that the 'proper' inflection of *ija*-stems has been preached to students for over 150 years; Friðriksson's (1861: 20) Grammar School grammar only gives the old/standard paradigms, whereas the popular elementary grammar of Ásmundarson (1878: 14) explicitly states that most people do not know how to 'correctly inflect' ('beygja rjett') this class, adding that the alternative is 'wrong' ('rangt').

Árnadóttir & Einarsdóttir (2007) studied the declination of the irregular noun *hönd* 'hand', which, similar to the *ija*-stems above, has been a regular in grammar books since the nineteenth century. Examining 171 elementary school students in addition to a group of university students, they found no less than twenty-one different declension paradigms for this common everyday word. A striking result was that the standard pattern was used by only 8.8% of the elementary students. At a rate of 12% for the standard declination, however, the university students did outperform the children.

As much previous research suggests, the use or adoption of standard grammatical features clearly attests to variation at a social and socioeconomic level (cf. e.g. Viðarsson 2019). Further evidence of class-based effects in Icelandic purism – a double standard if you will – has also been observed in terms of lexis. As Pálsson (1995) maintains, Icelandic language policy appears to make a distinction based on class:

> Thus, the lexical paraphernalia of the University [of Iceland] and the policy makers much of which is of foreign origin, for instance words denoting academic rank,

> i.e. *lektor*, *dósent*, and *prófessor*, has largely been left intact while lower-class slang has been systematically attacked from above. Words like *troll* (from 'trawl') and *hol* or *hal* (from 'haul'), popular in the everyday language of fishermen, have generally been rejected by the policy makers who insist (sometimes unsuccessfully) upon more 'Icelandic' expressions. (Pálsson 1995: 135)

Scholars have also made the claim that there is a double standard to be found in the fact that lexical borrowings are generally not considered acceptable in writing or in the media, even if they are used frequently in speech:

> My experience at the NSR [National State Radio] proved a clear double standard at work in regards to language use, hosts and guests would simply change codes once on air, so as to live up to the request for language purity. This double standard is – not surprisingly – at work elsewhere in Icelandic society. (Thórarinsdóttir 1999: 296; see further Thórarinsdóttir 1999: 279–84 and Kristinsson 2008 on the role of state broadcasting; Kristmannsson 2004: 160–61 on a more general note).

It should be borne in mind that such double standards are, of course, nothing new. As I point out elsewhere, none other than the main Icelandic teacher at the Reykjavík Grammar School, Halldór Kr. Friðriksson, allowed himself, at least occasionally, to produce both lexical and grammatical features in writing in the private sphere that he meticulously corrected in student assignments for decades on end (see Viðarsson 2019: 36, 190–91).

Based on the discussion above, the question naturally arises as to why the success story of the standardisation narrative has been so prevalent in the literature. Besides comments such as those by Elspaß (2014) and others, such narratives being histories of prescription rather than practice, the answer also lies in perhaps the only linguistic variable in Icelandic where standardisation actually was particularly successful. What should be kept in mind, however, is that this was no normal undertaking.

3. Doubling down on prescription: The epidemisation of *flámæli* 'skewed speech'

As already mentioned, the twentieth century brought with it a widening in focus towards the spoken language. The best known and in the view of contemporary Icelandic teachers and scholars alike probably most urgent linguistic feature to become the subject of systematic suppression through language standardisation via the school system was the near-complete eradication of 'skewed speech' ('flámæli') or 'sound error' ('hljóðvilla'), a phonological merger of the vowels /ɪ, ɛ/ and /ʏ, œ/ (written <i, e> and <u, ö>), respectively (for overview and discussion, see e.g. Þráinsson & Árnason 1984; Hjálmarsdóttir 2021; Guðmundsdóttir 2022). Scholars have argued that the stigmatisation of 'skewed speech' was, at least to some extent, class-based (see e.g. Pálsson 1995; Thórarinsdóttir 1999; Sigurðsson 2006 for critical discussion along these lines). More recently, Hjálmarsdóttir (2021) even argues that the actions and attitudes towards 'skewed speech' may be characterised as epistemic injustice. While the merger resulted in some words becoming homophonous,[4] it was really no different from the many phonological mergers that have previously occurred in Icelandic over the centuries. What is different, however, is how the process unfolded.

The successes in halting/reversing the merger are rather remarkable and are taken by some to suggest that reversing language change is not only a common goal in language standardisation but a realistic one, even at such abstract levels, provided it is at the same time associated with a strong ideology (see Árnason 2003b on this point). Scholars agree that

4 One story that runs in the family is of a farmer who had attempted to order *blek* 'ink' at the district warehouse in Borgarnes (Western Iceland) during the early twentieth century, apparently pronouncing it *blik* 'bleach(ing)'. Naturally, the warehouse clerk associated this with a very different product they also had in stock. Unhappy with the goods presented to him, the farmer responded: 'No, *bliiiik*, you know, for writing!' – 'Oh, you mean *blek*!' the clerk then responded, to which the farmer replied: 'That is what I have been saying!' Similar anecdotes are mentioned in Hjálmarsdóttir (2021).

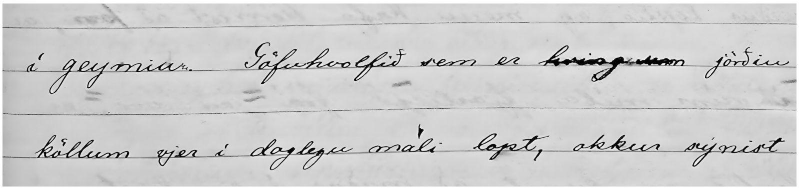
such was the case with 'skewed speech', which was referred to as if it were a disease or even an epidemic in public discourse (see e.g. Kristinsson 2008: 37; Hjálmarsdóttir 2021: 14, with references). As a first-hand testimony by a female speaker born in 1946 reveals, elementary school students who failed to distinguish between the two vowel pairs were forced to repeatedly reread texts out loud to the rest of the class until they managed to pronounce them properly (Hjálmarsdóttir 2021: 47).

The roots of 'skewed speech' can be traced at least to the nineteenth century (see e.g. Hjálmarsdóttir 2021: 12; Guðmundsdóttir 2022: 135). We know this because it was brought to North America by Icelandic immigrants and became a widespread feature of heritage Icelandic from the late nineteenth century onwards, relatively unaffected by standardisation and the Icelandic educational system (Arnbjörnsdóttir 2006). In terms of overt remarks, 'skewed speech' as a trait is mentioned in Þórarinsson (1888: 10), an article on reading and literacy training among children, where it is described as 'the mixing of the *e* and *i* sound and *ö* and *u* sound', said to be 'rather common among people in daily speech' (my translation; see also Þorsteinsson 2012: 39).

There are also attestations of 'skewed speech' (and teacher corrections thereof) in nineteenth-century student assignments from the Reykjavík Grammar School, preserved at the National Archives of Iceland (see Viðarsson 2018, 2019). An example is shown in Figures 14.1 and 14.2, presumably predating any systematic erasure of the feature within the school system, which was a twentieth-century phenomenon.

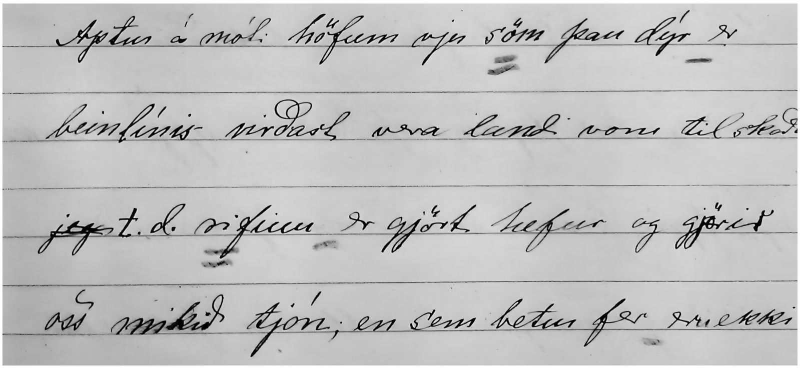
More explicit and harsh remarks on 'skewed speech' are found in the elementary grammar *Móðurmáls-bókin* (Ólafsson 1915: 19). The first edition of *Móðurmáls-bókin* (1911) only has a vague remark about 'wrong' pronunciation; 'skewed speech' being the actual phenomenon is only made explicit in the second edition (1915), which gives clear examples along with an extremely negative evaluation:[5]

5 Curiously, Ólafsson's (1911, 1915) grammar also passes judgement on variation in pronunciation more generally, where even the majority pronunciation spoken in and beyond the capital region, both westward and southward towards the north, is judged harshly. It appears that the pronunciation endorsed by Ólafsson is some



á geymiar. Göfuhvolfið sem er ~~kringum~~ jörðina
köllum vjer í daglegu máli lopt, okkur sýnist

Figure 14.1 Skewed speech in *göfuhvolfið* (= *gufuhvolfið*) 'the atmosphere', underlined by the teacher (Pálmi Pálsson) in a 2nd-grade essay (Axel Schierbeck, 1890).



Aptur á móti höfum vjer söm þau dýr er
beinlínis virðast vera landi voru til skaða
~~jeg~~ t.d. rifinn er gjört hefur og gjörir
oss mikið tjón; en sem betur fer eru ekki

Figure 14.2 Skewed speech in *söm þau dýr* (= *sum...*) 'some (of) those animals' and the noun phrase *rifinn* (= *refinn* 'the fox'), corrected by the teacher (Pálmi Pálsson) in a 2nd-grade essay (Helgi Jónsson, 1890).

> In some places in the country, the pronunciation is so corrupt that people do not distinguish between *e* and *i*, e.g. say *Jisús* for *Jesús*, *fenna* for *finna* ('find'), while others mix together *u* and *ö*: *gjura* for *gjöra* ('do'), *öpp* for *upp* ('up'). One often sees this wrong pronunciation in letters. – Although grown-up people in some places have this ruddy awful [Icel. 'endemis'] pronunciation, teenagers should not be allowed to make a habit of such indecency, and it is the self-evident duty of every teacher to meticulously train the children in pronunciation until they pronounce

sort of mixed bag of regional features from widely different areas but corresponds to a great extent to much later proposals of Guðfinnsson (1947) for a Standard Icelandic pronunciation.

> both these sounds and others correctly. (Ólafsson 1915: 18; my translation, emphasis in original)

Remarks on 'skewed speech' are also to be found, for instance, in the first Icelandic elementary grammar of the then newly established state publishing house for textbooks (*Ríkisútgáfa námsbóka*). The grammar states that those who 'mix these sounds together' have no option but to 'try to the best of their ability to correct their pronunciation' (Hjartar & Jónsson 1940: 15, my translation). This is indeed what the school system attempted – and apparently succeeded in.

The eradication of 'skewed speech' involved a massive, orchestrated effort to systematically train students to properly distinguish these four vowels and to impede the spread of the merger through unprecedented social stigmatisation. In fact, the Faculty of Arts of the University of Iceland in their 1951 report on plans to standardise pronunciation insisted that people with this linguistic trait should not be allowed to speak on the radio nor to perform any acting at the national theatre (see Kristinsson 2008: 37 and reference therein). Despite much discussion, especially during the first half of the century, standardised pronunciation never became a part of official legislation, but widespread negative attitudes towards 'skewed speech' appear to have sufficed (see e.g. Guðmundsdóttir 2022: 60–1, 117–23). Explicit remarks on the linguistic feature actually only made it into the 1976 national curriculum (see Óladóttir 2017: 62), although previous national curricula, the (draft) national curriculum of 1948 and the 1960 curriculum, both mention speech exercises and in particular the training of 'beautiful pronunciation' as in the 1948 curriculum, which was obviously the same thing (see Hjálmarsdóttir 2021: 24).

As can be read in the state reports on primary education from 1920 to 1966 (Statistical Bureau of Iceland 1967: 54), systematic efforts to eradicate 'skewed speech' through speech training at the elementary level were initially led by the linguist Björn Guðfinnsson (1905–1950), starting in winter 1946–1947 in Laugarnesskóli, Reykjavík. Guðfinnsson, who had documented phonological variation and dialects in the whole of Iceland in the 1940s (see e.g. Þráinsson & Árnason 1992), was eager to produce a state standard for the spoken language with selected features that he deemed worthy of preservation and to be taught in schools. These ideas even resulted in a concrete proposal (Guðfinnsson 1947).

While Guðfinnsson thus obviously had greater plans, the special speech training in elementary schools was nearly entirely devoted to 'teaching the correct pronunciation of the vowels *e*, *i*, *u* and *ö*' (Statistical Bureau of Iceland 1967: 54) and involved annual testing of all children nine to twelve years of age. Of those that needed training to alleviate this 'speech deficit' ('málgalli'), twenty-five lessons are said to be sufficient for most, although some children required more extensive training – up to forty lessons according to the report for 1956–1957. In that year, over 15% of children (117 out of 776 children) were diagnosed with 'skewed speech', of which ninety-two (78.6%) passed. A year later, 1957–1958, about 14% of children (107 out of 788) aged nine to twelve were reported to need this training, with eighty-four (78.5%) passing the exam (Statistical Bureau of Iceland 1967: 60–1). In both years, the proportion of children with 'skewed speech' in the youngest group is considerably higher than for the rest, at 26% and 19%, respectively. After Guðfinnsson's passing in 1950, the speech training programme was continued. In 1961, however, the category 'skewed speech' (or *hljóðvilla* 'sound error', used in 1960) had disappeared from the statistics and was no longer listed as such among speech deficits.[6] Over the next few decades, 'skewed speech'

6 Unless it has been subsumed under the headings 'unclear speech' ('óskýrmæli'), 'immature language'/'late language development' ('vanþroska málfar'/'seinn málþroski') or 'various/other speech difficulties' ('ýmsir/aðrir talörðugleikar'). In total, twenty-five out of ninety-nine students listed for speech therapy in 1961–1962 are found in these categories, fifty-four out of 152 students in 1962–1963 (Statistical Bureau of Iceland 1967: 54, 58). The most likely candidate for 'flámæli' being the last category, the numbers are seven (7.1%) and twenty-eight (18.4%), respectively. The remaining speech deficit categories are 'stuttering' ('stam'), 'guttural speech [i.e. uvular /r/ as opposed to alveolar trill]' ('kverkmæli'), 'lisp' ('smámæli'), 'nasality' ('nefmæli') and 'cleft palate' ('gómklofi').

was effectively eliminated, at least in the capital region (see Þráinsson & Árnason 1984).[7]

4. A glance at Icelandic elementary grammars: 1878–1944

Ottósson's (1990) hypothesis regarding effects of compulsory schooling on the shape and tone of Icelandic language policy, cited in the beginning of this chapter, is an intriguing one and merits much further research. By way of testing this hypothesis, grammars from both before and after the Education Act of 1907 have to be consulted and analysed insofar as (explicit or implicit) prescriptive remarks and negative (as well as positive?) evaluation is provided. In nineteenth-century grammars, typically intended for the secondary school level and beyond, the norm had been to state the standard norm without any negative evaluation (or even mention) of non-standard variants. Friðriksson's (1861) grammar, influential in the nineteenth century, is a textbook example of this tradition and is in that regard very different from Guðfinnsson's (1937) grammar, which frequently passes judgement on non-standard variants.

Guðfinnsson's grammar book was popular and used throughout the twentieth century, appearing in multiple editions. The claim by Ottósson (1990) is likely based at least in part on this grammar due to its harsh prescriptive tone (see also Blöndal 2001 on the lasting effect of Guðfinnsson's

7 In 2013, the Icelandic National Broadcasting Service (RÚV) aired an item on 'skewed speech' as part of a light-hearted television series on the Icelandic language (Pálmason, Skúlason & Þorgeirsdóttir 2013). The item featured an interview with an older person who had the trait in question and was proud of it. Stating that it was difficult to find speakers nowadays exhibiting 'skewed speech', RÚV jokingly advertised the interview under the heading 'the last skew-spoken Icelander?' (RÚV 2013). During her study of 'skewed speech', Hjálmarsdóttir (2021: 54) also reports that she often experienced the feeling of arriving forty years too late to the scene to study the phenomenon.

grammar on the Icelandic grammar tradition). The fifth edition, incidentally, published posthumously, is by many considered the ultimate version of the book, although it was heavily edited (and not by Guðfinnsson himself), becoming more concise and explicit, the style harsher, the commentary reduced and thus all in all having an even more austere appearance (cf. Sigtryggsson 2006: 50).

Partly similar to Ottósson (1990), Konráðsson (2004: 201) mentions remarks on language use in textbooks on Icelandic, pointing in particular to 'books for younger children' as regards explicit evaluation in terms of correct/incorrect or good/bad. While there is an overlap here with Ottósson's (1990) claim at some level, Konráðsson (2004: 203) makes rather different observations with regard to the more contemporary Icelandic elementary grammars more generally when he writes:

> On the whole, direct instruction is not common in the textbooks. The authors avoid the terms 'correct' or 'incorrect' when discussing language usage, and what little direct instruction is to be found in the books of the last three decades or so is very monotonous: 'dativitis' and the inflection of a few nouns.

With this brief introduction in place, let us now move on to explore Icelandic elementary grammars from a historical perspective. The following grammars have been consulted and analysed in terms of overt remarks and evaluation of language use – beginning with Ásmundarson's (1878, 1880) elementary grammar(s) and ending with the publication of the first Icelandic grammar by the state publishing house for textbooks in the 1940s:

- *Stuttar rjettritunarreglur …* (Ásmundarson 1878)
- *Ritreglur* (Ásmundarson 1880)
- *Stutt ágrip af íslenzkri mállýsingu* (Briem 1891)
- *Íslensk málfræði handa byrjendum* (Jónasson 1909, 1910, 1920)
- *Móðurmáls-bókin / Litla móðurmáls-bókin handa börnum og byrjendum* (Ólafsson 1911, 1915, 1920)
- *Ágrip af íslenzkri málfræði* (Briem 1910, 1918, 1921, 1929, 1932, 1936)
- *Málfræði handa börnum* (Jónsson 1912)

- *Íslensk málfræði handa alþýðuskólum* (Björnsson 1922, 1931, 1935)
- *Ágrip af íslenzkri málfræði fyrir barnaskóla* (Helgason 1935)
- *Íslenzk málfræði handa skólum og útvarpi* (Guðfinnsson 1937, 1939, 1944, 1946 …)
- *Íslenzk málfræði* (Finnbogason & Magnússon 1940, 1941, 1942)
- *Íslenzk málfræði* (Hjartar & Jónsson 1944, 1947 …)

The above grammars are not just a small random sample but a representative and complete overview of Icelandic grammar books of a general nature, intended for or used in elementary schools up until (some well beyond) the first half of the twentieth century (cf. also the overview in Þorsteinsson 2012: 95–105). By reading through these grammars in the first editions (unless otherwise stated) and noting explicit negative remarks on language use, we arrive at the overview in Table 14.1.

Table 14.1 shows considerable evidence of explicit remarks and negative evaluation for which Ottósson's hypothesis predicts an increase in grammars after 1907. Among the early post-1907 grammars, *Ágrip af íslenzkri málfræði* (1910) and *(Litla) Móðurmáls-bókin* (1911/1920) stand out in terms of the amount of prescriptive remarks, especially the latter, whereas *Íslensk málfræði handa byrjendum* (1909) and *Málfræði handa börnum* (1912) have few or, in the case of the latter, none whatsoever. The two following grammars, *Íslensk málfræði handa alþýðuskólum* (1922) and *Ágrip af íslenzkri málfræði fyrir barnaskóla* (1935), both have a relatively modest amount of such remarks, whereas Guðfinnsson's *Íslenzk málfræði handa skólum og útvarpi* (1937) raises the bar again considerably, while still somewhat below that of *(Litla) Móðurmáls-bókin* (1911/1920). Finally, the two state-printed grammars, the first edition of *Íslenzk málfræði* (1940/1942) and the revised *Íslenzk málfræði* (1944/1947) appear strikingly different. While the first edition is relatively free of explicit remarks on language use, the modified second edition changes the tone considerably. What, then, about *(Litla) Móðurmáls-bókin* (1911/1920), with its relatively great number of explicit remarks and, as we have seen above (cf. Section 3), extremely harsh tone? Its author, the editor Jón Ólafsson, was a unique character, known for his vivid writing style characterised by swagger, personal conceit and a hot temper which led him to flee the country twice

Table 14.1 Overview of overt remarks on language use using the terms 'wrong' or some other form of negative evaluation.

Title (year of publication)	'Wrong'	Other negative evaluation	Total number of pages
Stuttar rjettritunarreglur með málfræðislegum skýringum (1878)	6	3	46
Ritreglur (1880)	7	11	69
Stutt ágrip af íslenskri mállýsingu handa alþýðuskólum (1891)	10	0	79
Íslensk málfræði handa byrjendum (1909)	1	3	91
Ágrip af íslenzkri málfræði (1910)	13	3	110
(Litla) Móðurmáls-bókin (1911/1920)	29	19	180
Málfræði handa börnum (1912)	0	0	23
Íslensk málfræði handa alþýðuskólum (1922)	6	1	66
Ágrip af íslenzkri málfræði fyrir barnaskóla (1935)	3	0	43
Íslenzk málfræði handa skólum og útvarpi (1937)	8	17	162
Íslenzk málfræði (1940/1942)	1	3	84
Íslenzk málfræði (1944/1947)	10	5	96

(Aðalsteinsdóttir 2016: 181–3). As such, his grammar book may not at all be typical for post-1907 grammars in general.

As can be seen, the grammars thus differ rather greatly among themselves and there is definitely not a clear-cut pre- or post-1907 effect to be seen. This becomes even clearer if we group the grammars together and normalise the frequency of negative evaluations based on page numbers, as shown in Table 14.2.

Although normalisation by page number is a very crude measure, obviously, it still gives a better idea when comparing individual grammars

Table 14.2 Aggregated values of negative evaluation over three periods, normalised per ten pages (data from Table 14.1).

Year of publication	Negative evaluation (normalised)
before 1907	5,8
1907–1925	5,6
1925–1950	4,3

since they differ greatly in size, some being small leaflets and others book-length textbooks. What is striking is that there is, in fact, hardly any difference between the period before and after 1907. Moreover, the normalised rate after 1925 is even lower than the two previous periods.

While more research is clearly needed that not only takes into account negative remarks (proscription) but also, for instance, positive remarks and codification that is not based on general usage but instead the attempted standard norm, the overviews above at least attest to the fact that the nineteenth-century tradition of 'presupposing' rather than explicitly forbidding is slowly shifted in late nineteenth-century and early twentieth-century grammars.

Despite the fact that Table 14.2 did not yield the exact picture one would expect to find according to Ottósson's (1990) hypothesis, there are some possible pointers towards increasing explicitness. The two state-funded grammars on the list (*Íslenzk málfræði* 1940/1944), for one, are very interesting in how similar the books are in terms of their content and general appearance, but different when it comes to normative usage. The first edition, published in 1940, was actually met with rather harsh criticism, especially concerning its structure. The second edition, published in 1944 with different authors on the title page, is quite similar to the first edition in many regards but stricter in the negative evaluation of non-standard features. Some revealing examples are given below.

Common but non-standard morphological variation involving gender and case agreement features remarked upon in the first edition (Finnbogason & Magnússon 1940: 36), for instance, faces the verdict 'this ought to be evaded' ('þetta ber að varast', the whole phrase in bold), whereas the second edition (Hjartar & Jónsson 1944: 35) adds 'this is completely wrong and ought to be avoided' ('þetta er alveg rangt og ber að forðast'), both 'wrong' and 'ought to be avoided' in bold. More features are remarked upon in this way in the second edition and also comments on a specific set of irregular nouns such as *fótur* 'foot', *lækur* 'stream, brook' and *bróðir* 'brother', that 'care is needed to inflect them correctly' ('aðgæzlu þarf til þess að beygja þau rétt', 1944: 40).

Both Guðfinnsson's and the state-funded grammar of 1944 can be said to double down on most of the previous grammars. One cannot help but wonder if the popularity of the grammar of Guðfinnsson, which was much too complex for the elementary level, was still a reason for styling its harsher tone in the modified version of the state elementary grammar. This grammar was reprinted many times and in use until it was superseded by *Móðurmál: Námsbók handa barnaskólum* (Sigurðsson 1968).

While it was not included in the quantitative study, Sigurðsson's *Móðurmál* clearly marks a representational shift in the approach to non-standard linguistic features. Intended for fourth grade, prescriptive remarks are now presented repeatedly under the heading 'Do you speak correct language?' ('Talar þú rétt mál?'),[8] which are in prominent boxes throughout the book, set against a coloured background.[9] A number of sections also have written assignments labelled 'grammatical errors' ('málvillur'). Not only the more visual but also the more systematic way of presenting standard – sometimes even non-standard – forms is yet

8 The first edition was republished with minor changes in 1970 and reprinted at least until 1990 (Sigurðsson 1990: 112). The sequels for fifth and sixth grade (published in 1969 and 1970, respectively) are similar as regards the prominence of prescriptivism and have the same 'Do you speak correct language?' headings found in the fourth-grade material.

9 The first instance of this heading is preceded by the following observation, which attests to the supposed viability of this approach to language: 'In the language of children and adults, wrong inflections and various solecisms (*málleysur*) can often be heard. It is usually easy to relinquish grammatical errors (*málvillur*) by learning what is considered correct language. In the following, the correct use of some words will be pointed out from time to time' (Sigurðsson 1968: 15, 1990: 18).

another striking development and prominent feature of these state-led ideals of a (homogeneous) standard language.

5. Conclusion

This paper has explored Ottósson's (1990) hypothesis that the introduction of compulsory schooling in 1907 led to a noticeable change in the tone of Icelandic grammar books. According to Ottósson's hypothesis, remarks on language use became more direct and stringent as they were geared towards a younger audience. This shift, in turn, is claimed to have fed into the broader Icelandic language policy. Through an analysis of a dozen Icelandic grammar books from the period 1878 to 1944, offering a view before and after the Education Act of 1907, it has been argued that this hypothesis is at least partially supported by the data. The increasingly explicit prescriptive remarks of twentieth-century grammars as well as a broadening in focus towards the spoken language attests to changes that have had a lasting effect on Icelandic language policy.

However, the study also underscores previous findings in the literature suggesting that the traditional narrative of the widespread effects of standardisation in Icelandic is overstated and come about, it is argued, through (a) a misconstrued conflation of codification of a norm and its successful implementation or spread (cf. also Haugen 1987; Elspaß 2014; Viðarsson 2019), and (b) the relatively successful eradication of the highly stigmatised phonological merger involved in 'skewed speech' ('flámæli'), the systematic suppression of which is totally without precedent in the history of Icelandic and by no means suggestive of the effects of Icelandic standardisation and language-planning efforts in general.

Turning, finally, to the future outlook from a present perspective, it is worth pointing out that the latest revision of Icelandic language policy (2021–2030), issued by the Icelandic Language Council, distances itself partly from tradition. Current policy explicitly states the necessity of a shift towards a more inclusive and positive approach to language cultivation, one which presents standard norms without any condemnation of speakers' actual language usage (Íslensk málnefnd 2021). It is the hope of the author that this revised policy will make negative twentieth-century attitudes and views on language, which undermine common goals of ensuring the future preservation of the Icelandic language (as outlined in the language policy), a thing of the past.

Bibliography

Aðalsteinsdóttir, Auður. 2016. *Bókmenntagagnrýni á almannavettvangi: Vald og virkni ritdóma á íslensku bókmenntasviði*. Doctoral dissertation, University of Iceland.

Árnadóttir, Arngunnur & Gyða Dögg Einarsdóttir. 2007. 21 aðferð til að beygja orðið hönd: Könnun á beygingu orðsins hönd í íslensku nútímamáli. *Skíma: Málgagn móðurmálskennara* 30(1), 18–21.

Árnason, Kristján. 2003a. Icelandic. In Ana Deumert & Wim Vandenbussche (eds). *Germanic Standardizations: Past to Present*. Amsterdam/Philadelphia: Benjamins. 245–79.

Árnason, Kristján. 2003b. Language planning and the structure of Icelandic. In Kristján Árnason (ed.). *Útnorður*: *West Nordic Standardisation and Variation*. Reykjavík: Institute of Linguistics, University of Iceland Press. 193–218.

Arnbjörnsdóttir, Birna. 2006. *North American Icelandic: The Life of a Language*. Winnipeg: University of Manitoba Press.

Ásmundarson, Valdimar. 1878. *Stuttar rjettritunarreglur með málfræðislegum skýringum*. Reykjavík.

Ásmundarson, Valdimar. 1880. *Ritreglur*. Akureyri: Björn Jónsson.

Benediktsson, Hreinn. 1969. On the Inflection of the *ia*-Stems in Icelandic. In Jakob Benediktsson (ed.). *Afmælisrit Jóns Helgasonar 30. júní 1969*. Reykjavík: Heimskringla. 391–402.

Benediktsson, Jakob. 1985. Inngangur. In Helgi Þorláksson (ed.). *Crymogæa*: *Þættir úr sögu Íslands*. Reykjavík: Sögufélag. 13–56.

Björnsson, Benedikt. 1922. *Íslensk málfræði handa alþýðuskólum*. Akureyri: Björn Jónsson.

Blöndal, Þórunn. 2001. Er líf eftir Björn? *Íslenskt mál og almenn málfræði* 23, 203–13.

Briem, Halldór. 1891. *Stutt ágrip af íslenzkri mállýsingu*. Reykjavík.

Briem, Halldór. 1910. *Ágrip af íslenzkri málfræði*. Reykjavík.

Deumert, Ana & Wim Vandenbussche. 2003. Research directions in the study of language standardization. In Ana Deumert & Wim Vandenbussche (eds). *Germanic Standardizations: Past to Present*. Amsterdam, Philadelphia: Benjamins. 455–69.

Elspaß, Stephan. 2014. Prescriptive norms and norms of usage in nineteenth-century German. In Gijsbert Rutten, Rik Vosters & Wim Vandenbussche (eds). *Norms and Usage in Language History, 1600–1900: A Sociolinguistic and Comparative Perspective*. Amsterdam, Philadelphia: Benjamins. 303–20.

Finnbogason, Magnús & Marteinn Magnússon. 1940. *Íslenzk málfræði*. Reykjavík: Ríkisútgáfa námsbóka.

Friðriksson, Halldór Kr. 1861. *Íslenzk málmyndalýsing*. Kaupmannahöfn: Hið íslenzka bókmentafèlag.

Friðriksson, Halldór Kr. 1871. Nokkrar athugagreinir um íslenzku: verba valent usu. *Þjóðólfur* 23, 105–11.

Gíslason, Konráð. 1843. Bókafregn. *Fjölnir: Árrit handa Íslendíngum* 6, 59–74.

Gíslason, Konráð. 1844. Bókafregn. *Fjölnir: Árrit handa Íslendingum* 7, 71–104.

Guðfinnsson, Björn. 1937. *Íslenzk málfræði handa skólum og útvarpi*. Reykjavík: Ríkisútvarpið.

Guðfinnsson, Björn. 1947. *Breytingar á framburði og stafsetningu*. Reykjavík: Ísafoldarprentsmiðja.

Guðmundsdóttir, Margrét. 2022. *Mál á mannsævi: 70 ára þróun tilbrigða í framburði – einstaklingar og samfélag*. Doctoral dissertation, University of Iceland.

Gunnarsson, Gísli. 1983. *Monopoly Trade and Economic Stagnation: Studies in the Foreign Trade of Iceland* 1602–1787. Lund: Ekonomisk-historiska föreningen.

Guttormsson, Loftur. 1993. Frá kristindómslestri til móðurmáls: Hugmyndafræðileg hvörf í lestrarefni skólabarna um síðustu aldamót. *Uppeldi og menntun* 2, 9–23.

Haugen, Einar. 1987. Blessings of Babel: Bilingualism and Language Planning. Problems and Pleasures. *Contributions to the Sociology of Language* 46. Mouton de Gruyter.

Hauksdóttir, Auður. 2016. Björguðu Danir íslenskunni? Um tengsl íslensku og dönsku á átjándu öld. *Skírnir* 190(2), 420–57.

Heimisdóttir, Linda Ösp. 2008. *Yfirlit yfir beygingu karlkyns ija-stofna frá nítjándu öld til nútímamáls*. BA thesis, University of Iceland.

Helgason, Sigurður. 1935. *Ágrip af íslenzkri málfræði fyrir barnaskóla*. Reykjavík.

Hjálmarsdóttir, Emma Björk. 2021. *Hljóðrugl, hljóðsýki og framburðarfaraldur: Voru fordómar gegn flámæli dæmi um þekkingarlegt ranglæti*? MA thesis, University of Iceland.

Hjartar, Friðrik & Jónas B. Jónsson. *Íslenzk málfræði*. Reykjavík: Ríkisútgáfa námsbóka.

Ingvadóttir, Kristjana Vigdís. 2021. *Þrautseigja og mikilvægi íslenskrar tungu: Um notkun dönsku og erlend áhrif á íslensku*. Reykjavík: Sögufélag.

Íslensk málnefnd. 2021. *Íslensk málstefna* 2021–2030. <https://islenskan.is/malstefna/> (accessed 08 August 2024).

Jensson, Gottskálk. 2008. The Latin of the North: Arngrímur Jónsson's Crymogæa (1609) and the Discovery of Icelandic as a Classical Language. *Renæssanceforum* 5, 1–28. <https://www.njrs.dk/5_2008/gj.pdf> (accessed 08 August 2024).

Jónasson, Jónas. 1909. *Íslensk málfræði handa byrjendum*. Akureyri: Oddur Björnsson.

Jónsson, Arngrímur. 1985. *Crymogæa: Þættir úr sögu Íslands*. Ed. by Helgi Þorláksson. Reykjavík: Sögufélag. [Icelandic translation by Jakob Benediktsson.]

Jónsson, Björn. 1900. *Íslenzk stafsetningarorðabók*. Reykjavík: Ísafoldarprentsmiðja.

Jónsson, Finnur. 1912. *Málfræði handa börnum*. Kaupmannahöfn: Møller.

Jónsson, Jóhannes Gísli & Þórhallur Eyþórsson. 2003. Breytingar á frumlagsfalli í íslensku. *Íslenskt mál og almenn málfræði* 25, 7–40.

Konráðsson, Sigurður. 2004. Language cultivation in elementary school. In Ari Páll Kristinsson & Gauti Kristmannsson (eds). *Málstefna –Language Planning.* Reykjavík: Íslensk málnefnd. 197–210.

Kristinsson, Ari Páll. 2008. *'Í fréttum er þetta helst': Rannsókn á einkennum íslensks útvarpsmáls.* Doctoral dissertation, University of Iceland.

Kristinsson, Axel. 2018. *Hnignun, hvaða hnignun? Goðsögnin um niðurlægingartímabilið í sögu Íslands.* Reykjavík: Sögufélag.

Kristmannsson, Gauti. 2004. Will Icelandic language policy corner the Icelandic language? In Ari Páll Kristinsson & Gauti Kristmannsson (eds). *Málstefna – Language Planning.* Reykjavík: Íslensk málnefnd. 155–62.

Kusters, Wouter. 2003. *Linguistic Complexity: The Influence of Social Change on Verbal Inflection*. Utrecht: LOT.

Leonard, Stephen Pax & Kristján Árnason. 2011. Language ideology and standardisation in Iceland. In Tore Kristiansen & Nikolas Coupland (eds). *Standard Languages and Language Standards in a Changing Europe.* Oslo: Novus Press. 91–6.

Möller, Alda B. 2017. Íslenskukennsla í Bessastaðaskóla 1806–1846 og á fyrstu árum Reykjavíkurskóla. *Orð og tunga* 19, 1–40.

Ottósson, Kjartan G. 1990. *Íslensk málhreinsun: Sögulegt yfirlit.* Reykjavík: Íslensk málnefnd.

Ottósson, Kjartan. 2005. Language cultivation and language planning IV: Iceland. In Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr Jahr, Allan Karker, Hans-Peter Naumann & Ulf Teleman (eds). *The Nordic Languages: An International Handbook of the History of the North Germanic Languages.* Berlin, New York: Walter de Gruyter. 1997–2007.

Óladóttir, Hanna. 2017. *Skólamálfræði: Hver er hún og hver ætti hún að vera? Um markmið og áhrif málfræðikennslu á unglingastigi grunnskólans*. Doctoral dissertation, University of Iceland.

Ólafsson, Bragi Þ. 1999. Í þágu niðjanna – Framtíðarsýn Íslendinga á nítjándu öld. *Sagnir: Tímarit um söguleg efni* 20, 4–11.

Ólafsson, Jón. 1911. *Móðurmáls-bókin: Kenslubók. 1. (Orðflokkar – Hljóðfræði – Beygingafræði – Orðmyndunarfræði).* Reykjavík: Jón Ólafsson.

Ólafsson, Jón. 1915/1920. *Litla móðurmálsbókin: Handa börnum og byrjendum*. Reykjavík: Bókaverzlun Sigfúsar Eymundssonar.

Pálmason, Konráð, Bragi Valdimar Skúlason & Brynja Þorgeirsdóttir. 2013. *Orðbragð.* Episode 4. RÚV.

Pálsson, Gísli. 1995. *The Textual Life of Savants: Ethnography, Iceland, and the Linguistic Turn*. Chur: Harwood Academic Publishers.

Pettersson, Gertrud. 1988. Bisatsledföljden i svenska eller Varifrån kommer BIFF-regeln? *Arkiv för nordisk filologi* 103, 157–80.

Rastrick, Ólafur. 2003. Lestrarkennsla og ritþjálfun í barnaskólum og alþýðlegum framhaldsskólum 1880–1920. In Ingi Sigurðsson & Loftur Guttormsson (eds). *Alþýðumenning á Íslandi 1830–1930: Ritað mál, menntun og félagshreyfingar.* Reykjavík: Háskólaútgáfan. 93–113.

RÚV. (2013, 15 December). *Síðasti flámælti Íslendingurinn*? <https://www.ruv.is/frettir/menning-og-daegurmal/sidasti-flamaelti-islendingurinn> (accessed 08 August 2024).

Sigtryggsson, Jóhannes B. 2006. Tvær gerðir málfræðibókar: Breytingar á Íslenzkri málfræði Björns Guðfinnssonar. *Skíma* 29(1), 46–50.

Sigurðsson, Ársæll. 1968. *Móðurmál: Námsbók handa barnaskólum*. Reykjavík: Ríkisútgáfa námsbóka.

Sigurðsson, Ársæll. 1990. *Móðurmál.* Reykjavík: Námsgagnastofnun.

Sigurðsson, Gísli. 2006. Íslensk málpólitík. *Tímarit Máls og menningar* 3, 59–74; 4, 77–91.

Statistical Bureau of Iceland. 1967. *Barnafræðsluskýrslur árin 1920–1966. Primary education 1920–66. Statistics of Iceland, supplementary series, no. 1.* Reykjavík: Hagstofa Íslands/Statistical Bureau of Iceland.

Svavarsdóttir, Ásta, Gísli Pálsson & Þórólfur Þórlindsson. 1984. Fall er fararheill: Um fallnotkun með ópersónulegum sögnum. *Íslenskt mál og almenn málfræði* 6, 33–55.

Svavarsdóttir, Ásta. 2013. Þágufallshneigð í sjón og raun: Niðurstöður spurninga-kannana í samanburði við málnotkun. In Höskuldur Þráinsson, Ásgrímur Angantýsson & Einar Freyr Sigurðsson (eds). *Tilbrigði í íslenskri setningagerð: I. Markmið, aðferðir og efniviður.* Reykjavík: Málvísindastofnun Háskóla Íslands. 83–110.

Smári, Jakob Jóh. 1920. *Íslenzk setningafræði*. Reykjavík: Bókaverzlun Ársæls Árnasonar.

Teleman, Ulf. 2010. Lågtyska lånord. Om stad och land i svensk språkhistoria. In Gunilla Byrman, Aanna W. Gustafsson & Henrik Rahm (eds). *Svensson och svenskan. Med sinnen känsliga för språk*. Lunds universitet: Språk- och litteraturcentrum. 331–41.

Thomason, Sarah G. 2001. *Language Contact*. Washington: Georgetown University Press.

Vandenbussche, Wim. 2007. Shared Standardization Factors in the History of Sixteen Germanic Languages. In Christian Fandrych & Reinier Salverda (eds). *Standard, Variation und Sprachwandel in germanischen Sprachen. Standard, Variation and Language Change in Germanic Languages*. Tübingen: Gunter Narr. 25–36.

Viðarsson, Heimir F. 2018. *Corpus of Corrections: Reykjavík Grammar School (1846–1904)*. <https://github.com/heimirfreyr/RLSS> (accessed 8 June 2024).

Viðarsson, Heimir F. 2019. *Socio-Syntactic Variation and Change in Nineteenth-Century Icelandic: The Emergence and Implementation of a National Standard Language*. Doctoral dissertation, University of Iceland.

Wright, Sue. 2016. *Language Policy and Language Planning: From Nationalism to Globalisation*. 2nd edition. Basingstoke: Palgrave Macmillan.

Þorkelsson, Jón. 1870. Um nokkurar rangar orðmyndir eða orðskipanir í íslenzku. *Norðanfari* 9, 82–3, 86–7, 89–90.

Þorkelsson, Jón. 1887. *Breytingar á myndum viðtengingarháttar í fornnorsku og fornislensku*. Reykjavík: Prentsmiðja S. Eymundssonar & S. Jónssonar.

Þorkelsson, Jón. 1903. Haldór Kristján Friðriksson. *Andvari* 28, 160.

Þorkelsson, Jón. 1918. Frásögn síra Einars Hálfdanarsonar um hlaupið úr Öræfajökli 1727. *Blanda* 1, 54–9.

Þorsteinsson, Helgi Már. 2012. *'Að vanrækja móðurmálið … það er synd, sem hefnir sín'. Íslenskukennsla í barnaskólum 1880–1920*. MA thesis, University of Iceland.

Thórarinsdóttir, Hallfrídur 1999. *Purity and Power: The Policy of Purism in Icelandic Nationalism and National Identity*. Doctoral dissertation, New School for Social Research.

Þórarinsson, Jón. (1888). Lestur. *Tímarit um uppeldi og menntamál* 1, 3–15.

Þráinsson, Höskuldur & Kristján Árnason 1984. Um reykvísku. *Íslenskt mál og almenn málfræði* 6, 113–34.

Þráinsson, Höskuldur & Kristján Árnason. 1992. Phonological Variation in 20th Century Icelandic. *Íslenskt mál og almenn málfræði* 14, 89–128.

GURO NORE FLØGSTAD

15 A prescriptive source with a spoken language imprint? Tense and aspect in Argentinean language textbooks

ABSTRACT
Recent studies suggest a link between textbooks and market pressures, particularly on the Latin American continent, with exclusion of regional forms as one consequence. Still, Latin American Spanish varieties (as any other) exhibit variation due to ongoing change. How do textbooks convey this variation? I investigate a specific case: the expansion of the Preterit category, a major morphosyntactic change occurring in Argentinean Spanish, and its representation in language textbooks. Normative and prescriptive sources are typically conservative, but Argentina stands out with its strong cultural industry and colloquial literary ideal. This study shows a surprising result: the textbooks indeed reflect ongoing change and regional forms; they are prescriptive sources with imprints associated with orality and the language of young people. I approach this finding from sociolinguistic and semantic angles and discuss both the complexity of the semantic distinction involved, and notions of standardisation and conservatism. Ultimately, I link the tension between normative and textual sources to the effect on sociolinguistic differentiation on language change.

1. Introduction

Prescriptive authorities are usually conservative. In prescriptive materials, such as textbooks, grammars and dictionaries, grammatical, orthographic and lexical distinctions not necessarily representative of the surrounding spoken language prevail. Even the smallest morphological markers may play a role in such materials. A well-known example of the conservative nature of prescriptive authorities is eighteenth-century English grammarians, who were responsible for not only legitimising, but perhaps even saving inflectional morphology that was on the verge of extinction in the English language. Apparently, on the model of Latin, grammarians were eager to preserve tense, person and number distinctions as much as they could, even though such forms were on their way out (Milroy & Milroy 1985: 70–1). In addition, standardised grammars tend to represent a language where variation has no place, and standardisation is in this sense tightly linked to variation reduction and seeks to suppress optional variability.

We know that when variability does exist, the choice of one variant over another has the potential to take on social meaning. Still, the preference of linguists during the last few decades has been to account for spontaneous, authentic, spoken language, and questions of normativity have often been deemed unscientific (e.g. see discussion in Lebsanft & Tacke 2020). Nevertheless, investigating the way such variables are transmitted in prescriptive works may inform us about how these variables are viewed in the relevant linguistic community, and linguists should engage critically with prescriptive texts. After all, they impose grammatical norms on an entire speech community, in this case, on schoolchildren.

Language textbooks are standardising agents *par excellence*, intrinsically linked to the notion of a standard language or variety. The creation of a standard may be an informal, long-lasting practice, but these informal practices are typically accompanied by 'formal activities that give explicit recognition to the linguistic norm or standard' (Lebsanft & Tacke 2020: 3), such as the elaboration of 'systematically designed reference books on orthography and orthoepy, on grammar and vocabulary' (Lebsanft & Tacke 2020: 3). The standard then becomes the normal, exemplary use associated with the well-educated upper/middle class, and the process that leads to this, i.e. standardisation, can be broadly defined as a process in which one variety of a language becomes widely accepted in the speech community, a supradialectal norm rated above regional and social dialects, although these may be felt to be appropriate in some domains (see Lane et al. 2018). Despite research on textbooks as standardising agents, more research is needed to contextualise them and their role in shaping young learners' linguistic identities, especially on the Latin American continent, where the use of Spanish is intrinsically linked to colonisation, and to powerful promotors of Standard Spanish. This is the task of this paper, by means of focusing on a semantic distinction with substantial local variation: the use of Preterits[1] vs. Perfects. I seek to answer the following questions: How are Perfect/Preterit categories presented in textbooks in Argentina's elementary schools? Do textbooks convey synchronic variation (as present in the speech community), or are they influenced by one of the contiguous prestigious varieties: Peninsular Spanish, on the one hand, or a unified Latin American supranational norm, on the other?

The initial hypothesis was that the local textbooks fail to reflect the complexity of the local norm, but rather represent a standard-like, or supranational, norm in Argentine schoolchildren; that is, to use the Perfect frequently. This was assumed due to previous research, which shows that in the domain of personal pronouns, textbooks reflect a near-standard norm, as well as what we know about variation and ongoing change being notoriously difficult to communicate in normative sources. Contradicting this hypothesis, I find that the Argentinean textbooks represent a grammatical norm reminiscent of that of younger people. This finding is understood in light of sociolinguistic theories of standardisation, differentiation and ultimately linguistic change, including the role of written, normative sources in linguistic change.

2. Background

2.1 *The Argentinean sociolinguistic context*

The hegemony of Spanish saw its final consolidation through nation building in nineteenth-century Argentina. Argentina received five

1 Capital letters indicate language-specific categories.

million immigrants after its birth as a nation, and a monolingual policy was soon imposed on the newcomers, who shifted rapidly to the surrounding language. The current linguistic context includes indigenous languages, recent immigrant languages, and the languages brought to the country during the large-scale immigration to the country in the late nineteenth and early twentieth centuries. This panorama makes the local linguistic situation complex and the local linguistic variety interesting in at least two ways. First, it deviates quite drastically from both Peninsular Spanish varieties and from most Latin American varieties (Uruguayan Spanish being the prominent exception, although there are differences, particularly in prosody and lexicon). It differs most notably in terms of phonological features (*ʃeísmo*), prosody (early peak alignment), lexical loans from former immigrant languages (such as Italian), and importantly for us here, morphosyntactically (*voseo*), as well as in the preference for the Preterit category, as we shall see in the next section (for an overview, see Colantoni & Rodríguez Louro 2013). Second, as a community, it is characterised by a strong literary scene, an important cultural industry, and, according to some studies, a strong linguistic self-esteem, sometimes interpreted as nationalism (Delaney 2002). Llull & Pinardi (2014), in a study of linguistic attitudes in 400 Argentinean individuals, conclude that these speakers, when asked about attitudes toward the variety spoken in Buenos Aires, feel 'identificación y orgullo por la propia variedad lingüística' (Llull & Pinardi 2014: 56), and that this pride is visible in the reported resistance toward changing accents, and a clear preference for the local variety in media and popular culture.

Argentina today holds a particular place in the Spanish-speaking world. In the *Diccionario panhispánico de dudas* (Real Academia and Asociación de Academias de la Lengua 2015: xiv), it is noted in the introduction that the Spanish-speaking world constitutes a norm which is 'extraordinariamente homogénea en todo el ámbito hispánico, con variaciones mínimas entre las diferentes zonas, casi siempre de tipo fónico y léxico'. That is, they see variational differences as most pronounced in the phonological and lexical domains. This claim stands in stark contrast to both perceived and actual regional differences. Stepanov wrote about strong local varieties of Spanish in Latin America, derived from what he called the 'dialecto colonial', and pointed to a particular linguistic nationalism within the Spanish-speaking world – Argentina being one such case, alongside Brazil. Del Valle & Villa (2012: 47) also point to 'la peculiaridad lingüistica rioplatense', *rioplatense* being synonymous with the Spanish variety spoken in Buenos Aires and around Río de la Plata. Previous research (López García 2010, 2015, 2017) has suggested that Argentinean publishers, in line with globalisation pressures and compliance with a supranational Latin American norm, adhere to the latter rather than reflecting the local Argentine variety and its variations. Lauria & López García (2009) also argue that the supposed interest and apparent inclusive approach of the Real Academia Española versus Panhispanism merely represent discursive instalments, meant to reproduce old hierarchies, far from fomenting any linguistic equality. In any case, the pluricentricity of norms, and the role of strong local linguistic varieties, has not received the attention it deserves, Lebsanft & Tacke conclude (2020: 25).

2.2 Language textbooks as standardising agents in the local context

Textbooks' conservative nature and relation to standardisation is well known, but recent years have seen an increased focus on textbooks and their role in the 'global linguistic order' (Curdt-Christiansen 2017). Textbooks are increasingly understood not merely as selectors in the 'messy variability of spoken interaction', to paraphrase Gal & Woolard (1995), but in light of their particular social, cultural, political and economic purposes. Pennycook (2001: 81) argues that studying them may make ideological systems and representations transparent. It is implicit in this line of research that textbooks have the power to define what constitutes legitimate knowledge, and hence contribute to shaping young learners' linguistic identities. In Argentina, much research on textbooks has occurred within the field of *glotopolítica*, a tradition of critical discourse analysis. Narvaja de Arnoux (2016: 1), for example, notes that prescriptive grammars participate in the 'reproducción o transformación de sociedades'.

2.3 *The grammatical categories in question*

Language-specific studies on textbooks support the notion that such normative sources are likely to ignore local linguistic varieties and regional linguistic features. López García (2017: 164–5) argues that Argentinean textbooks ignore local morphosyntactic features such as the possessive in constructional position, e.g. *de ella,* in favour of the more standard-like *suya*, which is rare in Argentinean, and asks pupils to choose between the two ('¿La ....[suya/de ella] tendrá créditos?'), while ignoring other local forms (e.g. using the local *detrás mío* is explicitly deemed an error; the 'correct' – that is, standardised – form being *detrás de mi*). In a similar vein, López López et al. (2019) argue that the pronoun *vos*, typically associated with Argentinean and Uruguayan Spanish, is excluded from the classroom in Spanish teaching, where the *tú* is preferred. Amorós-Negre (2020: 596–7) also provides examples of morphosyntax that does not reflect the Latin American reality.

Do Argentinean textbooks reflect such a linguistic reality? In order to approach this question, a grammatical distinction, whose description in prescriptive works provides a particularly productive point of departure, is the distribution of the Preterit (a past action, such as *Leí* 'I read') vs. the Perfect (a past action with current relevance, such as *He leído* 'I have read') in varieties of Latin American Spanish. Just as orthography may carry social meaning, so, too, can morphosyntactic variables. The prototypical distribution, where the perfect expresses past actions with relevance in the present (see e.g. Comrie 1976) does not exist in many languages, but notably does in English and most Peninsular Spanish varieties. On the Latin American continent, the distribution of the two categories has proven to be particularly complex (Howe 2013; Fløgstad 2016; Rodríguez Louro 2009, 2016). As a general tendency, the Preterit is preferred over the Perfect in a variety of contexts in Spanish spoken on the Latin American continent, with Buenos Aires Spanish being the most pronounced in that direction, making a speaker of such a variety likely to use the Preterit to express meanings such as experientiality, as in *Estuve en Nueva York* (*Estuve* is Preterit) for 'I have been to New York' (as opposed to *He estado en Nueva York*). I return to the details of this distribution below, after I outline this paper's approach to tense and aspect.

2.4 Tense and aspect

The approach to tense and aspect I employ here corresponds to the functional-typological approach of Comrie (1976, 1985) and Bybee, Perkins & Pagliuca (1994). Using a post-structuralist, substantialist approach to tense and aspect based on large typological samples, the authors identify *gram-types*, identifiable by virtue of their prototypical characteristics. The proto-meaning of a perfect, then, is a past action with *current relevance*, a notion which essentially means that a past action is connected with the present time in some way. Following this approach, I define the Romance Perfect a 'past with current relevance'. The Romance Preterit, on the other hand, would correspond to a bound action that occurred before the speech event. Crucially, perfect meaning may be sub-divided into four categories, listed here from the more to the less prototypical (refer to the proto-meaning of current relevance above): resultative, experiential, recent past, and ongoing situation. These subfunctions will be employed in the detailed analysis of the textbooks below.

2.5 The Perfect/Preterit distinction in Argentinean Spanish

The prototypical distribution of Perfects and Preterits is one in which Perfects express pasts with current relevance, and Preterits past actions, as in English. As for Romance languages, the entire Latin American continent differs from the Peninsular Romance varieties in the way their Perfect has evolved. Schwenter and Cacoullos (2008) compare frequencies of Perfect use in contexts where this form is expected (e.g. current relevance contexts). They find that in Peninsular Spanish, the Perfect is used in 54%, Mexican 15%, Salvadorian 22% and Peruvian 21.9%. However, the most extreme of the Spanish varieties in terms of avoiding the Perfect is Argentinean (Fløgstad & Rodríguez Louro 2021; Rodríguez Louro 2009), although it is unclear whether the Perfect is now restricted to certain semantic contexts (e.g. experiential settings to express generic reference and indefinite past, as argued by Rodríguez Louro 2009: 249), or if the Preterit is taking over the Perfect's entire terrain and becoming a generalised past tense. In any case, Fløgstad & Rodríguez Louro (2021) show that there has been a steep decline in Perfect usage in Argentinean Spanish from the late nineteenth century. Argentinean adults currently exhibit considerable synchronic variation in their use of these verbal categories to express the same semantic context, while children remarkably less so. In their cross-linguistic study of children's narratives, Berman & Slobin (1994) note the complete absence of Perfects in Argentineans children's narratives (while children acquiring Spanish in Madrid, on the other hand, prefer the Perfect to express, e.g. resultant states, and use this form widely), which leads the authors to the conclusion that the Perfect is 'marginal' and restricted to written language in the Argentinean variety (see also discussion in Rodríguez Louro 2009: 48). Other studies find children to use the Perfect in only 6% of the relevant contexts (Fløgstad 2016), and they show a clear preference for the Preterit (94%).

Argentinean Spanish is not, however, homogeneous, and it should be noted that in this article, I refer to the variety spoken in and around the capital, Buenos Aires, often referred to as *porteño*. Certain Northern provinces are, on the other hand, reported to show a preference for the Perfect, not the Preterit (Fløgstad 2016: 73). Despite this variation, the research referred to above concentrates on the variety of the capital, and the textbooks analysed are published there, which motivates this geographical focus.

3. Method

The data collection involved gathering and analysing different text and document types, primarily textbooks, but also other government documents, such as curricula and webpages that provide information about the selection and distribution of textbooks. The approach is reminiscent of multisited document ethnography in that these sources are understood both in terms of their textual properties and history, as well as their practical meaning and simultaneity.

3.1 Selection of textbooks

The nine textbooks chosen here belong to established Argentinean publishers and are frequently used in both private and public schools. The books are freely distributed to the public schools through *Dirección Nacional de Políticas Socioeducativas del Ministerio de Educación*, and books are chosen through regional committees based on a list of criteria related to text type, presentation of content, etc. Established as a means of securing teaching materials to all pupils, this practice has been problematised as a regulatory axis of school language policy (López García 2017: 158) with possible consequences for the language representations that are being transferred to pupils at schools.

The textbooks chosen are grammar and activity books for grades four, five and six in primary school. These levels constitute the so-called *segundo ciclo* within the local school system. The first *ciclo*, grades one – three, involves learning basic skills such as reading, writing and mathematics, whereas the *segundo ciclo* (grades four to six) involves learning grammar in greater depth.

Table 15.1 List of textbooks analysed.

Name	Year of publication	Grade	Publisher (name and abbreviation)	Place of print	Distributed nationally?	Public and private schools?
Va con vos Prácticas de lenguaje 4; 5; 6	2017; 2017; 2018	4; 5; 6	Santillana (Sant)	Buenos Aires, Argentina	Yes	Yes
Avanza Prácticas de lenguaje 4; 5; 6	2017; 2016; 2017	4; 5; 6	Kapelusz (Kap)	Buenos Aires, Argentina	Yes	Yes
En órbita Prácticas de lenguaje	2018; 2018; 2018	4; 5; 6	Puerto de Palos (PP)	Buenos Aires, Argentina	Yes	Yes

3.2 Authorship

All textbooks list teams of editors and authors as responsible for the editions. Puerto de Palos (2018a, 2018b, 2018c) lists different editors with different responsibilities. Both Kapelusz (2017a, 2017b, 2017c) and Santillana (2017a, 2017b, 2017c) explicitly refer to the textbooks as collective works. For this reason, the books are referred to by the name of their publishing house, and not their editorial team members. It should be noted that given the collective authorships of these textbooks, we may not make assumptions about the varieties spoken by the authors. We may, however, assume that the textbooks have been written by highly educated speakers, with a clear pedagogical purpose, curated by editors and reviewers, whose age and education level suggest a frequent use of Perfects (Kubarth 1992; but see Rodríguez Louro 2009: 138).

3.3 Curricula

Before embarking on the analyses of the textbooks, it is crucial to consider whether local curricula explicitly mention teaching of verbal categories. As it turns out, verbal morphology is indeed referred to in the curricula, not surprising given the complexity of Spanish verbal morphosyntax. Having finished sixth grade, Argentine youngsters must distinguish 'Relaciones entre tiempos, modos y aspectos verbales', 'perfecto imperfecto', 'perfecto simple, presente, futuro' (Dirección de Curricula 2012: 754), as well as have knowledge of '[t]iempos de la narración (uso del pluscuamperfecto, perfecto simple, imperfecto, condicional, por ejemplo, en cuentos, novela, relatos de experiencias)' (Dirección de Curricula 2012: 754), and should be familiar with relations between verbal categories, e.g. '*perfecto imperfecto, perfecto simple presente-futuro; pluscuamperfecto perfecto simple;imperfecto; presente*' (Dirección de Curricula 2012: 745).

4. Analysis of the relevant categories and their usage

The linguistic text material in the textbooks is manifold, and in order to assess information about the nature of the linguistic distinction in question, a number of text types were considered. Each text type required a suitable methodology. The types of data were broadly divided into explicit descriptions of the relevant categories, on the one hand, and their actual usage in texts, on the other.

The obvious point of departure was the *descriptions of the verbal* categories in question. These descriptions provide the basis of the analyses in this paper. In addition, the appendices' representations of verbal paradigms were analysed. The question of explicit mentioning involved analysing the texts in order to identify their stance towards the relevant linguistic categories. I use two types of qualitative content analysis to interpret these documents: descriptive (providing a systematic analysis of the object of study) and thematic (synthesising the descriptive findings in a critical, thematic manner).

The second type of data is *examples of the relevant categories* extracted from the prose in the textbooks. This category includes examples provided to accompany the descriptions of the verbal categories. In addition, this category involves the metaprose provided by the authors, for example, introducing a new topic, or guiding the pupil in the text. This additional datasource was added because the aim was not only to assess the explicit teaching of the relevant categories, but the totality of the language to which pupils were exposed. Therefore, the *entire* textbook was considered a source (in line with, e.g. the methodological approach of Requena & Tissera (2018) with regard to clitic placement represented in textbooks). Note that excerpts from novels, newspapers and short stories also present in the textbooks were later excluded, as their origin and authorship were unclear.

The examples were assessed by identifying the relevant forms in the textbooks. This approach means, of course, that although the entire textbook was viewed as a source, only relevant forms were extracted and analysed. This task called for a systematic assessment. Because the categories in question are undergoing semantic change, this assessment was particularly complex and called for both formal and semantic analyses. The formal analysis involved identifying the relevant forms through morphological criteria:

- the analytical perfect: *He estado* 'I have been'
- the synthetic preterit: *Estuve* 'I was'

The semantic analysis involved identifying contexts which prototypically call for Perfects, where those exist, such as those outlined in Table 15.2, adapted from Comrie (1976: 56–61).

The semantic analysis was crucial. Languages that have a perfect (such as English, exemplified in Table 15.2) will employ that form to express the aforementioned subfunctions. If, on the other hand, the Preterit was found instead of the Perfect in these subfunctions, this would indicate an expansion of the Preterit to context typically reserved for the Perfect form. As we shall soon see, this was indeed the case.

Table 15.2 Overview of subfunctions of the Perfect.

Subfunction	Definition	Example
Experiential	Refers to a past action without referring to a specific occurrence, characterised by non-specific past time reference	*I have been to New York*
Ongoing situation	Asserts that the event holds throughout an interval, delimited by the reference time and a certain time prior to it	*I have lived five years in New York*
Recent past	Is used where there is a relation of temporal closeness between the present and past situations	*I have just seen it*
Resultative	Indicates that a present state is the result of some past situation, and asserts that the result state of the event holds at the reference time	*I have made a cake*

Table 15.3 Overview of categories taught explicitly.

Publisher		4th grade			5th grade			6th grade		
		Sant	*Kap*	*PP*	*Sant*	*Kap*	*PP*	*Sant*	*Kap*	*PP*
Tense	*Past*	x	x	x	x	x	x	x	x	x
	Present	x	x	x	x	x	x	x	x	x
	Future	x	x	x	x	x	x	x	x	x
Aspect	*Present perfect*	-	-	-	-	-	-	-	-	-
	Past perfect		-	-	x	x	-	x	x	-
	Progressive	-	-	-	-	-	-	-	x	-
	Perfective/ Imperfective	x	-	x	x	x	x	x	x	x
Mood	*Indicative*	x	-	-	x	-	-	x	x	-
	Subjunctive	-	-	-	x	-	-	x	x	-
	Imperative	-	-	-	x	-	-	x	x	

4.1 Overview of findings

I begin the analysis by providing an overview of the relevant grammatical categories. Table 15.3 indicates whether the categories in question are explicitly taught in the textbooks.

Two findings emerge from this overview. First and foremost, it indicates the absence of the Perfect as a category. Second, it illustrates a clear progression from tense to other verbal categories. Progressive aspect, for example, is not mentioned until sixth grade. Mood is not mentioned at all in fourth grade and only by one manual in fifth grade. This prevalence of tense over aspect is expected, as I return to in the discussion.

Table 15.4 Descriptions of grammatical categories in the different textbooks by grade.

		Past	Past perfective	Past imperfective
Santillana	4	Si ocurrió antes del 8momento en que se habla el tiempo del verbo es pasado o pretérito (ganó) (p. 87)	El pretérito perfecto simple se usa cuando una acción ha empezado y finalizado en el pasado (Los demás habitantes quedaron tan impresionados con la hazaña que la nombraron reina) (p. 105)	Se emplea para las acciones del pasado que tuvieron cierta duración o que ocurrían de manera habitual (El pez se burlaba de ellos con carcajadas estrepitosas). También es el verbo que se usa en las descripciones (En un principio la isla no tenía gobernantes) (p. 105)
	5	Se relatan hechos ya ocurridos (p. 105)	Pretérito perfecto simple expresa acciones puntuales, es decir, con un comienzo y un final precisos (corrió, notó). (p. 105)	Pretérito imperfecto indica acciones que duran en el tiempo (estaba, movía). Se emplea para actividades habituales o descripciones. (p. 105)
	6	(not mentioned)	Se utiliza para acciones pasadas y puntuales, es decir que empezaron y terminaron en un momento concreto (perdí la pulsera; la saqué del cajón) (p. 105)	Se usa para acciones pasadas no puntuales, que se extienden en el tiempo (tenía la pulsera); también para hábitos (se la usaba) y descripciones (era plateada y traía piedritas de colores) (p. 105)
Kapelusz	4	Nombra una acción estado o proceso en qué sucedió con anterioridad al momento en que se habla o escribe (nadaba muy bien; fui al cine) (p. 59)	(not mentioned)	(not mentioned)

Table 15.4 (*Continued*)

		Past	Past perfective	Past imperfective
	5	Expresa acciones, estado o procesos que ocurrieron antes del momento en que se habla (ayer comí dos pastelitos) (p. 47)	Para expresar acciones que ocurrieron una vez y terminaron (repetí, caminé, hice) (p. 47)	Para expresar una acción con cierta duración, un hábito o una descripción en el pasado (caminaba, hacía, repetía) (p. 47)
	6	Antes del momento en que se emite el mensaje (p. 47)	Cuando la forma verbal se refiere a una acción que ocurrió una vez y ya concluyó (cantó) (p. 47)	Cuando la forma verbal se refiere a una acción que se producía habitualmente o que tiene cierta duración (cantaba) (p. 47)
Puerto de Palos	4	Ocurrió en el pasado (Ayer recorrimos el lago Mascardí) (p. 51)	Hacer avanzar las narraciones (p. 51)	Para hacer descripciones (El grano de maíz era más pequeño y parecía una gota de agua) (p. 51)
	5	Ocurrió en el pasado (Ayer pulgarcito estudió mucho) (p. 51)	En las narraciones se suele utilizar el pretérito perfecto simple para hacer avanzar la acción (El rey invitó al marqués a subir a su carroza) (p. 51)	El pretérito imperfecto, para hacer descripciones (El joven era apuesto) (p. 51)
	6	Pretérito o pasado; lo expresado por el verbo es anterior al momento en que se emite el mensaje. (Ayer tuvimos una prueba de inglés) (p. 51)	Se utiliza para narrar las acciones (Las trillizas bucearon) (p. 51)	Se usa para hacer descripciones (Tenían piel perlada) o contar acciones de fondo (Bucearon mientras los padres tomaban mate) (p. 50)

4.2 Descriptions of grammatical categories

I now take a closer look at the descriptions of temporal and aspectual categories in the textbooks for each grade, by looking at the explicit teaching of the linguistic categories in question. I focus particularly on the relevant categories: past perfective, perfect and their definitions.

These detailed descriptions confirm the overall findings from Table 15.1: the Perfect is absent. When we look in detail at the relevant categories, some patterns emerge. When it comes to the category past, all textbooks provide simple and cognitively graspable definitions of the past as an action located in the past, that is, prior to the moment of speech. The past perfective is more complicated, because the authors are forced to make explicit the function of the aspectual component of perfectivity – a notoriously complex task. Puerto de Palos (2018a, 2018b, 2018c) attempts to do this through focusing on stylistics and narration. The other textbooks focus on the concluded aspect of the verb's meaning.

Ultimately, with regard to the past imperfective, both Santillana (2017a, 2017b, 2017c) and Kapelusz (2017a, 2017b, 2017c) focus on the durational aspect, whereas Puerto de Palos (2018a, 2018b, 2018c) has a description focused on the stylistic or narrative aspect and defines the imperfective as a descriptive aspect.

Overall, these semantic descriptions vary in specificity and in linguistic terminology. They do, however, all attempt to give detailed accounts of subtle aspectual/functional and temporal differences. This makes it all the more noteworthy that no such attempt is given in the case of the Perfect – it is simply absent from the descriptions.

4.3 Examples from text

This section lists cases of use of the Perfect, or the Preterit in perfect function, which appear in examples and running text in the different textbooks. All subfunctions associated with the Perfect were expressed through Preterits, as seen in Table 15.5, confirming the Preterit's possible use in those functions. The examples were analysed semantically, following the semantic analysis based on the distinction between subfunctions outlined in Table 15.2.

4.4 Variation

Do the textbooks involve any variation between the two forms – Perfect and Preterit – in terms of current relevance function? Recall that analyses of spontaneous speech have shown that adult speakers of Argentinean Spanish vary between the Perfect and the Preterit to express current relevance. This variation is particularly pronounced in the specific subfunction *ongoing situation* (see also Fløgstad 2016). As observed above, there is no mention of synchronic variation or ongoing change in the descriptions of past categories in the textbooks. We have also pointed out that all subfunctions associated with the perfect are expressed through Preterits in the descriptions. However, one domain stands out where both Perfects and Preterits are used in its expression in the metaprose of the textbooks. This is the subfunction *ongoing situation*, as seen in examples (1) and (2), which refer to activities (publishing) that are still occurring.

(1) *Publicó varios libros de poesía para niños* (Kapelusz, 2017a: 143)
'He has published several poetry books for children'

(2) *Ha publicado, entre otras obras, Maxi Marote, en la línea recta, XVZ: Archivos Ultrasecretos, Cinco Problemas para don Caracol y El misterio de la fuente* (Kapelusz, 2017a: 142)
'He has published, among other works, Maxi Marote en la línea recta, XVZ: Archivos Ultrasecretos, Cinco Problemas para don Caracol and El misterio de la fuente'

This is the only function in which the Perfect appears, and this variation is not unexpected. In spontaneous speech, ongoing situation is the last stance of the Perfect (Fløgstad 2016). In any case, it is noteworthy that this variation is not reflected in the descriptions of the categories.

4.5 Adherence to other standards

There is no sign of adhering to the more conservative standards either in the descriptions of grammatical categories or in their actual use. However, despite a spoken imprint in the presentation of linguistic categories in practice, the textbooks choose to include the same categories that they ignore in the running text, in the *appendices* listing grammatical categories, and there is an explicit mention of standard conjugation in the appendices. Pupils are directed to extra material, appendices and conjugation tables in order to assess the correct conjugation of regular verbs. Kapelusz has an adjacent *Wikibloc*, with conjugation patterns and exercises. In Santillana's (2017b) textbook for fifth graders, we find the following instructions:

(3) *Encontrá como conjugar los tiempos verbales en las páginas 172 y 173*
(Santillana 2017b: 104)
'Find how to conjugate the verbal tenses on pages 172 and 173'

Puerto de Palos (2018c) for sixth graders explicitly refers to the Real Academia Española's web pages in order to access the correct conjugation:

(4) *Tip: Pueden consultar la conjugación de los verbos en el diccionario en línea dle.rae.es* (Puerto de Palos 2018c: 51)
'Tip: You can consult the conjugation of verbs in the online dictionary dle.rae.es'

The use of regional forms in the textbooks vis-à-vis the reliance on standard grammar is noteworthy. Although RAE's webpages include the imperative forms ending in á (such as *encontrá* 'find'), this source does not explicitly acknowledge categorical differences (such as the non-use of Perfects). This observation may be understood in light of a formal vs.

Table 15.5 Examples of semantic subfunctions in textbooks.

	4th grade			5th grade			6th grade		
	Santillana	Kapelusz	Puerto de Palos	Santillana	Kapelusz	Puerto de Palos	Santillana	Kapelusz	Puerto de Palos
Experiential	¿Usaste alguna vez las palabras emoticono o emoticón? (p. 80)	¿Cuántas veces besaste un sapo? (p. 17)	Leí que en el hipermercado hay muchas ofertas (p. 22)	El dios del lago se enfureció por lo que hicimos (p. 7)	¿Fuiste al teatro una vez? ¿Qué viste? (p. 121)	¿Alguna vez vieron un tapir? (p. 28)	Perdí la pulsera de la suerte (p. 104)	¿Alguna vez leíste una novela policial? (p. 64)	Escribió obras teatrales (p. 118)
Resultative	La señorita Eugenia pidió materiales sobre el invierno para una cartelera (p. 88)	Llegó el cartero (p. 54)	¿Incluyeron todas? ¿Porqué? (p. 9) Alguien se tomó el agua (pg. 49) ¿Hicieron la torta? (p. 40)	¿Conseguiste el mapa? (p. 18)	El telégrafo ha sido reemplazado por otros medios de comunicación (p. 45)	¿Siguen los días de lluvia o ya mejoró el tiempo? (p. 22)	Ganó numerosos premios (p. 31)	¿Qué diferencias hay entre las publicidades que leíste en este libro y las que ves en la televisión? (p. 122)	Llegó el momento, Zeus (p. 7)

(*Continued*)

Table 15.5 (*Continued*)

	4th grade			5th grade			6th grade		
	Santillana	Kapelusz	Puerto de Palos	Santillana	Kapelusz	Puerto de Palos	Santillana	Kapelusz	Puerto de Palos
Ongoing situation	Buenos aires se llenó de mosquitos resistentes al frio (p. 90)	Desde tiempos muy antiguos, para hilar con mayor facilidad se utilizó un instrumento llamado huso (p. 16)	Durante el día tome ocho vasos de agua (p. 33)	Este animal fue perdiendo mucho territorio (p. 47)	Público, entre otras cosas, las obras de poesía (p. 96)	Recibió varios premios, entre ellos en Premio Barco de Vapor Argentina (pg. 117)	Publicó más de 50 libros para chicos (pg. 26) También publicó la novela flor de loto (p. 27)	Además, escribió algunas adaptaciones de obras clásicas (p. 57)	Recibió varios premios (p. 121) Se transformó en una ciudad muy importante (p. 27)
Recent past	Me enteré de que su maestra es X, como nosotras (p. 57)	¿Trabajaste bien en equipo? (p. 124)	¿Hiciste un regalo? (p. 40)	¿Ya viste lo que hay en la página 30 de tu anotatodo? (p. 93)	No pude (p. 48)	¿Hicieron la tarea? (pg. 40)	¿Tuviste el examen de inglés? ya? (p. 154)	Ya llegamos al Calafate (p. 29)	¿Escuchaste ese ruido? Sí lo escuché (p. 79)

a functional approach to language teaching, i.e. as the remains of a prescriptive grammar vs. a spoken imprint.

5. Discussion

In this section, I discuss the findings from the previous sections, regarding the presentation of the Perfect/Preterit categories presented in textbooks for Argentina's elementary schools, whether textbooks convey synchronic variation (as present in the speech community), and whether they convey traditional forms associated with a standard Peninsular variety, or a unified, Latin American norm.

The Perfect is absent in the formal presentations of the categories and has been replaced by the Preterit. Synchronic variation is not conveyed, except in the subfunction of ongoing situation. Moreover, there is no indication of adhering to either a standard-like Peninsular, or a supranational Latin American norm. Regional features associated with the speech of children and adolescents are indeed included in the analysed prescriptive material. The findings were surprising and warrant discussions rooted in both semantics and sociolinguistics. Since both research on normative sources per se, as well as case studies of the inclusion of specific linguistic features that deviate from the standard indicate their absence in such sources, the obvious question arises: why is this not the case for tense/aspect, and why not in Argentinean? First, I look to semantic explanations, before turning to broader sociolinguistic questions, eventually relating the findings to the topics of prescriptive sources and language change.

5.1 Semantic explanations

The prevalence of one temporal category – Preterit – over aspectual categories may be contextualised through Comrie's (1976, 1985) classical works on tense and aspect. Particularly the notion of time understood as a simple timeline – a diagrammatic representation adequate for illustrating time or tense in human languages, which intentionally leaves vague a number of elements – is relevant. Comrie saw a tendency to use notions of tense to describe notions that are aspectual.

These remarks may contextualise the *spoken imprint* visible in the prescriptive source. The school textbooks represent a new norm, that of younger people. There is no attempt at reflecting either the variation typical of the system of adults, or the exogeneous Peninsular Spanish norm. This is exhibited both in the explicit explanations of the verbal categories introduced and in the examples given. In a current relevance context, prototypically expressed by means of the perfect in, for example, English or Peninsular Spanish, the textbooks use the Preterit. There is also a clear *progression* when it comes to teaching semantic categories. In the aspectual domain regarding imperfective and perfective, the textbooks gradually introduce those after they introduce tense. The timeline presented is predictable: *presente*, *pasado*, *future* – present, past, future. The authors go on to gradually introduce aspect and mode. According to Comrie's timeline (1985: 5–7), this is expected: tense precedes aspect and mood – the latter less easily transmitted than the former.

In this sense, the textbooks demonstrate simplification, both in terms of the categories represented (tense over aspect/mood) and the chronology of their appearance (tense first). The temporal timeline is easier to grasp, explain, and it is cognitively more intuitive than both aspect and mood. In fact, this timeline coincides with most (Western) people's conceptualisations of time, and it coincides with the representations of tense in the textbooks analysed here. Finally, this avoidance of synchronic variation may in fact aid the transmission, diffusion and standardisation of an ongoing change.

The Perfect's absence stands in contrast to other morphosyntactic features: the conjugation associated with the pronoun *vos*, the local equivalent of *tú*. Both Kapelusz (2017a, 2017b, 2017c) and Santillana (2017a, 2017b, 2017c) explicitly acknowledge variation in the pronominal/morphosyntactic paradigm (*vos/tú*). There is thus explicit mentioning of variation of the *voseo*-conjugation – i.e. morphosyntactic variation – but not of the categorical variation found in the verbal domain. The explanation for this finding may lay in the formal vs. semantic properties of these features. *Vos/tú* is easily observable and its geographical distribution easy to grasp. Subtle semantic differences, such as those involved in the Perfect/Preterit distribution, are far less easy to grasp. Meaning is simply less salient than form. The details involved also require a complex semantic metalanguage. As Comrie noted (see above), aspectual differences are more difficult to convey – and likely, easier to leave out.

5.2 Sociolinguistic explanations

The semantic explanation outlined above fails to provide us with an answer to the obvious question: why would the authors, who are likely to use both Perfects and Preterits (recall that adults use both), choose the children's preferred form and not their own? If simplification was the driving force, why not decide on the Perfect, in line with the available prestigious standard variety? In order to contextualise this finding, we must look to the speech community as a whole and to attitudes towards non-standard speech. Historically, Argentina has a strong linguistic tradition of embracing orality and colloquial speech, dating back to the publication of the country's national epos, *Martín Fierro*, by José Hernandez (1872), famous for its deviating orthography, aimed at representing the then local variety and creating a national variety (which, some may argue, has also found its way to contemporary fiction, see, e.g. Cabezón Cámara 2017). The country's renowned authors, among them Jorge Luis Borges and Roberto Arlt, almost a century ago explicitly pointed to the discrepancy between the actual spoken vernacular and the language found in literature and dictionaries, as discussed by e.g. Thon (2010). Borges & Clemente (1968: 21), in their early description of Argentinean Spanish, even explicitly noted that dictionaries are replete with words that would never find their way into a speaker's mouth, an observation that resonates well with the consensus of linguists. Moreover, Lunfardo – a lexical repertoire that emerged in the late nineteenth associated with immigration – was frequently represented in literature, song lyrics and even dictionaries (see, e.g. Conde 2011). Summing up, Argentina has an important cultural industry, including a notable literary tradition relevant to our case.

Since the development of literary language is related to standardisation (Lane et al. 2018), the insistence on the oral and colloquial forms from the most prominent voices of the literary scene may have had an impact on the speech community and, more specifically, on the choices made by authors of language textbooks. These choices do not, as might be expected, reflect a formalised reminiscence of middle-class writing culture, but rather a move towards the colloquial and vernacular. This move, now emblematic of Argentinean linguistic and literary consciousness, makes it contiguous to understand it considering Gal & Irvine's (2019) notion of differentiation. Linguistic differentiation involves the well-known process of increased divergence of linguistic form, but Gal and Irvine focus on how such differences are framed and interpreted and subsequently made significant as conceptual schemas or ideologies. Importantly for our case, clearer boundaries and increased divergence may contribute to linguistic change.

In Gal & Irvine (2019), differentiation is explicitly linked to the advancement of linguistic change with Nguni languages' acquisition of clicks as one example. But where the Nguni languages became more similar to surrounding languages in certain phonological features, Argentinean Spanish has become more different from the surrounding Spanish varieties.

Previous research has suggested that the adherence to a *supranational* norm would impose conformity on sources of Spanish – e.g. language textbooks – due to a pressure to adopt a Eurocentric linguistic identity and a unified norm. Panhispanism is relevant here; an attempt at unifying the Spanish-speaking world linguistically, an attempt which actually saw its roots in the late nineteenth century, but whose effects are subject to discussion, for example, in light of theories of globalisation and commodification (see Del Valle 2009). While recent research has pointed to the absence of regional features in language textbooks, this study suggests a more nuanced picture, with normative sources opting for the morphosyntactic category that is leading the change. This warrants a broader discussion: may precisely this inclusion of regional forms in normative sources impact their legitimacy in the Argentinean speech community?

5.3 Differentiation and linguistic change

We have seen that there is a generational pattern in the usage of the Perfect in Argentinean Spanish, with elderly speakers using it far more frequently than young and adolescent speakers. This tendency suggests that the Perfect is indeed disappearing in this variety, at a much faster rate than in any other Latin American variety. Given this observation, it is pertinent to discuss the impact prescriptive authorities have on the Argentinean linguistic context. We know that lack of institutional support and acquisition may equal high variability of a given linguistic entity. Argentina stands out with its institutional support of local forms, as observed in this study, and pointed out by literary scholars. May this institutional support have influenced the legitimisation of local forms, such as the Preterit? What role do institutionally supported linguistic norms play in the spread of a linguistic entity, and where does the spread of grammar occur? Irvine and Gal note that the imagery involved in differentiation may influence or even generate linguistic differences (2019: 39).

Dictionaries and grammars are two means of gatekeeping, aimed at providing clear-cut answers, void of variability (Milroy & Milroy 1999: 19; see also Kristiansen & Coupland 2011 for discussions on standards and variation). As such, ongoing changes do not typically find themselves reflected in the standardised, prescriptive grammars. Once they do, this inclusion is likely to aid their acceptance and transmission. The exclusion of the Perfect in the Argentinean textbooks analysed may illustrate how authors of normative sources faced with synchronic variation in the aspectual domain opt to avoid it. Elspaß (2020) argues that written sources should constitute a legitimate source of sociolinguistic inquiry, and that written sources must be included to understand language history. Could tense and aspect in Argentinean textbooks not only reflect a linguistic change, but also aid the spread in the speech community?

6. Conclusion

In this article, I have shown that when faced with variation in the speech community, authors of Argentinean language textbooks opt for a morphosyntactic norm close to the spoken language – they chose the verbal category associated with children and adolescents. I have contextualised this finding of a double standard in two ways. First, I have pointed out how this finding's nature is likely related to semantic simplification, particularly how aspect is less salient and typically ignored in descriptions of verbal semantics. Second, I have contextualised the Argentinean case in Panhispanism and standardisation literature. The case in question has proven to be a particular one, with verbal morphosyntax conforming neither to a Peninsular, nor to a supranational Latin American norm. The Argentinean case has pointed to an overlooked phenomenon: the spoken imprint of a prescriptive source and its relation to an ongoing linguistic change, possibly as a result of increased divergence.

This finding illustrates the need for a more nuanced understanding of the relations between prescriptivism, regional varieties and sources that constitute legitimate knowledge. The Argentinean case shows that enhancing the regional aspect may be crucial in this legitimation process. Whether Argentina, with its particular cultural history, literary tradition with an oral imprint and loyalty to a local, differentiated, norm, constitutes an outlier, remains an open question.

Bibliography

Amorós-Negre, Carla. 2020. Normative grammars. In Franz Lebsanft & Felix Tacke (eds), *Manual of Standardization in the Romance Languages*. Berlin, Boston: de Gruyter. 581–604.

Berman, Ruth A. & Dan Isaac Slobin. 1994. *Relating Events in Narrative. A Cross-linguistic Developmental Study*. New Jersey: Lawrence Erlbaum Associates.

Borges, Jorge Luis & José E. Clemente. 1968. *El Lenguaje de Buenos Aires*. Buenos Aires: Emecé Editores.

Bybee, Joan, Revere Perkins & William Pagliuca 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Cabezón Cámara, Gabriela. 2017. *Las aventuras de la China Iron*. Buenos Aires: Random House.

Colantoni, Laura & Celeste Rodríguez Louro (eds). 2013. *Perspectivas teóricas y experimentales sobre el español de la Argentina*. Madrid: Iberoamericana.

Comrie, Bernard. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.

Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press.

Conde, Oscar. 2011. *Lunfardo. Un estudio sobre el habla popular de los argentinos*. Buenos Aires: Taurus.

Coupland, Nikolas & Tore Kristiansen. 2011. *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus Press.

Curdt-Christiansen, Xiao Lan. 2017. Language socialization through textbooks. In Patricia A. Duff & Stephen May (eds), *Language Socialization*. Cham: Springer. 195–210.

Del Valle, José. 2009. Total Spanish: The politics of a pan-hispanic grammar. *PMLA: Publications of the Modern Language Association of America* 124(3), 880–6.

Del Valle, José & Laura Villa. 2012. La disputada autoridad de las Academias: Debate lingüístico-ideológico en torno a la 'Ortografía' de 2010. *Revista Internacional de Lingüística Iberoamericana: RILI* 10 1(19), 29–53.

Delaney, Jean H. 2002. Imagining el ser argentino: Cultural nationalism and romantic concepts of nationhood in early twentieth-century Argentina. *Journal of Latin American Studies* 34(3), 625–58.

Dirección de Curricula. 2012. *Diseño curricular para la escuela primaria. Segundo ciclo, educación general básica*. Gobierno de la Ciudad Autónoma de Buenos Aires, Ministerio de Educación.

Drinka, Bridget. 2017. *Language Contact in Europe: The Periphrastic Perfect through History*. Cambridge: Cambridge University Press.

Elspaß, Stephan. 2020. Alternative sources of data for alternative histories of standardisation. *Language Policy* 19(2), 281–99.

Elspaß, Stephan, Nils Langer, Joachim Scharloth & Wim Vandenbussche (eds). 2007. *Germanic Language Histories 'from Below' (1700–2000)*. Berlin, New York: de Gruyter.

Fløgstad, Guro Nore. 2016. *Preterit Expansion and Perfect Demise in Porteño Spanish and beyond. A Critical Approach to Cognitive Grammaticalization Theory*. Leiden: Brill.

Fløgstad, Guro Nore & Celeste Rodríguez Louro. 2021. Gauging expansion in synchrony: The periphrastic perfect in nineteenth century Rioplatense Spanish. In Kristin Melum Eide & Marc Fryd (eds), *The Perfect Volume: Papers on the Perfect*. Amsterdam: Benjamins. 241–58.

Gal, Susan & Judith Irvine. 2019. *Signs of Difference: Language and Ideology in Social Life*. Cambridge: Cambridge University Press.

Gal, Susan & Kathryn A. Woolard. 1995. Constructing languages and publics: Authority and representation. *Pragmatics: Quarterly Publication of the International Pragmatics Association* 5(2), 129–38.

Howe, Chad. 2013. *The Spanish Perfects: Pathways of Emergent Meaning*. Basingstoke: Palgrave Macmillan.

Kapelusz. 2017a. *Avanza. Prácticas del lenguaje* 4. Buenos Aires: Kapelusz Editora.

Kapelusz. 2017b. *Avanza. Prácticas del lenguaje* 5. Buenos Aires: Kapelusz Editora.

Kapelusz. 2017c. *Avanza. Prácticas del lenguaje* 6. Buenos Aires: Kapelusz Editora.

Kristiansen, Tore & Nikolas Coupland (eds). 2011. *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus Press.

Kubarth, Hugo. 1992. El uso del Pretérito Simple y Compuesto en el español hablado en Buenos Aires. In *Scripta Philologica II*. México: Instituto de Investigaciones Filológicas, Universidad Nacional Autónoma de México. 553–66.

Lane, Pia et al. 2018. *Standardizing Minority Languages: Competing Ideologies of Authority and Authenticity in the Global Periphery*. London: Routledge.

Lauria, Daniela & María López García. 2009. Instrumentos lingüísticos académicos y norma estándar del español: La nueva política lingüística panhispánica. *Lexis* 33(1), 49–89.

Lebsanft, Franz & Felix Tacke. 2020 (eds). *Manual of Standardization in the Romance Languages*. Berlin, Boston: de Gruyter.

Llull, Gabriela & Carolina Pinardi. 2014. Actitudes lingüísticas en la Argentina. El español en Buenos Aires: Una aproximación a las representaciones de sus hablantes. In Ana Beatriz Chiquito & Miguel Ángel Quesada Pacheco (eds), *Actitudes lingüísticas de los hispanohablantes hacia el idioma español y sus variantes*. University of Bergen. 1–62.

López García, María. 2010. Norma estándar y variedad rioplatense en instrumentos de gramatización. *Bolívar* 11(14), 163–78.

López García, María. 2015. *Nosotros, vosotros, ellos. La variedad rioplatense en los manuales escolares*. Buenos Aires: Miño y Dávila.

López García, María. 2017. Las variedades lingüísticas del español en los libros escolares argentinos: Entre las políticas del estado y el mercado. *Estudios de Lingüística Aplicada* 65, 157.

López López, Giovani et al. 2019. The exclusion of *vos* from the Spanish as a foreign language classroom discourse: A critical examination through the lens of the language management theory. *Critical Inquiry in Language Studies* 16(4), 229–48.

Milroy, James & Lesley Milroy. 1985. *Authority in Language: Investigating Language Prescription and Standardisation*. London: Routledge.

Milroy, James & Lesley Milroy. 1999. *Authority in Language: Investigating Standard English*. London: Routledge.

Narvaja de Arnoux, Elvira Beatriz. 2016. La perspectiva glotopolítica en el estudio de los instrumentos lingüísticos: Aspectos teóricos y metodológicos. *Matraga* 23(38), 18–42.

Palos, Puerto de. 2018a. *Activa XXI. En órbita. Prácticas del lenguaje* 4. Buenos Aires: Editorial Puerto de Palos.

Palos, Puerto de. 2018b. *Activa XXI. En órbita. Prácticas del lenguaje* 5. Buenos Aires: Editorial Puerto de Palos.

Palos, Puerto de. 2018c. *Activa XXI. En órbita. Prácticas del lenguaje* 6. Buenos Aires: Editorial Puerto de Palos.

Pennycook, Alastair. 2001. *Critical Applied Linguistics. A Critical Introduction*. London: Routledge.

Real Academia Española and Asociación de Academias de la Lengua Española. 2005. *Diccionario panhispánico de dudas*. Madrid: Santillana.

Requena, Pablo E. & María Victoria Tissera. 2018. Variation in second language Spanish textbooks: A study of variable clitic placement. *Hispania* 101(1), 55–68.

Rodríguez Louro, Celeste. 2009. *Perfect Evolution and Change: A Sociolinguistic Study of Preterit and Present Perfect Usage in Contemporary and Earlier Argentina*. Doctoral dissertation, School of Languages and Linguistics, Faculty of Arts, The University of Melbourne.

Rodríguez Louro, Celeste. 2016. Indefinite past reference and the present perfect in Argentinian Spanish. *Studies in Language* 40(3), 622.

Santillana. 2017a. *Va con vos. Prácticas del lenguaje* 4. Buenos Aires: Ediciones Santillana.

Santillana. 2017b. *Va con vos. Prácticas del lenguaje* 5. Buenos Aires: Ediciones Santillana.

Santillana. 2017c. *Va con vos. Prácticas del lenguaje* 6. Buenos Aires: Ediciones Santillana.

Schwenter, Scott A. & Rena Torres Cacoullos. 2008. Defaults and indeterminacy in temporal grammaticalization: The 'perfect' road to perfective. *Language Variation and Change* 20(1), 1–39.

Stepanov, G. V. 2004. *La lengua española de España y América Latina: Sobre la variabilidad lingüística*. Vol. 42. LINCOM Europa. *Lincom Studies in Romance Linguistics*.

Thon, Sonia. 2010. La identidad lingüística argentina a través de Borges y Puig. *Arbor* 186(741), 117–27.

Notes on contributors

WENDY AYRES-BENNETT is Emerita Professor of French Philology and Linguistics, University of Cambridge (UK). She has published widely on the history of French and the history of linguistic thought. Her current research interests include language policy, standardisation and codification, and she is co-editor of the *Cambridge Handbook of Language Standardization* (2021). She was Principal Investigator on the multidisciplinary research project, *Multilingualism: Empowering Individuals, Transforming Societies*, funded by the AHRC 2016–2021 (<www.meits.org>). The project worked closely with policymakers and practitioners to promote the value of languages for key issues of our time and the benefits of language learning for individuals and societies.

SILVIA BALLARÈ is Junior Assistant Professor in Linguistics at the University of Bologna (Italy), where she teaches applied linguistics and sociolinguistics. Her main research interests are in the field of Italian sociolinguistics, with a focus on the relationship between intra- and interlinguistic variation.

ROSA MARIA CALAFAT VILA is Senior Lecturer in the Department of Catalan Studies and General Linguistics as well as the Language Policy Coordinator at the University of the Balearic Islands in Mallorca (Spain). Her research areas are pragmatics and historical sociolinguistics. Her recent publications on these two research areas focus on the role of language ideology in the discourse of the history of language.

STEPHAN ELSPAß is Professor of German Linguistics at the University of Salzburg (Austria). He has worked on various topics in sociolinguistics, variational linguistics and the history and grammar of German. His publications include a monograph on *Language History from Below* as well as edited volumes and articles on topics in historical sociolinguistics, in recent years with a focus on standardisation.

ALENA A. FIDLEROVÁ is a researcher in the Department of Czech Language and Communication Theory at the Faculty of Arts of Charles University in Prague (Czech Republic). She received her PhD from Charles University in 2008 and has published extensively on the (early) modern history of Czech, particularly on graphematics and orthography and the role of text genres.

GURO NORE FLØGSTAD is Professor of Linguistics and Norwegian at the University of South-Eastern Norway in Drammen (Norway) and an elected member of The Young Academy of Norway. Her research focuses on variation and change in Romance, Romani and Scandinavian languages.

JULIA HÜBNER is Postdoctoral Researcher at the University of Hamburg (Germany). She has published on norm and variation in the early modern period as well as on the interface between morphology and pragmatics. She received her PhD from the Free University of Berlin in 2021 for her thesis on *Norm und Variation. Paradigmenwechsel anhand frühneuzeitlicher Fremdsprachenlehrwerke* ('Norm and Variation. Paradigm shifts in foreign language textbooks of the early modern period').

ALENKA JELOVŠEK is a researcher at the Research Centre of the Slovenian Academy of Sciences and Arts, Fran Ramovš Institute of the Slovenian Language in Ljubljana (Slovenia). In her research she focusses on the Slovenian literary language of the sixteenth century, especially on the development of orthographical norms and personal pronouns; additionally, she researches the development of forms of address in Slovenian through history. She is a co-author and co-editor of the *Dictionary of the 16th-Century Slovenian Literary Language* (Vol. 1 A–D published in 2021) and member of the editorial board of the journal *Jezikoslovni zapiski*.

VIKTORIJA KOSTADINOVA is Lecturer in English Sociolinguistics at the University of Amsterdam (Netherlands), a position she has held since 2017. Prior to this, she completed her PhD on prescriptive ideology in American English at Leiden University (2018). Her main research interests include language prescriptivism, language variation and change, language use in online communities and American English.

IMKE MENDOZA is Professor of Slavic Linguistics at the University of Salzburg (Austria). She has worked on the history of Slavic and Slavic historical sociolinguistics with a special focus on Russian and Polish. Her research interests also include Russian and Polish syntax, semantics and morphosyntax.

KONSTANTIN NIEHAUS is Senior Scientist in German Linguistics at the University of Salzburg (Austria). Previously, he worked as a Research Assistant in Augsburg and Salzburg and as a Postdoctoral Assistant at the University of Innsbruck. He received his PhD in German Linguistics from the University of Augsburg in 2015. His research interests include standardisation, language variation and change, language and identity, enregisterment, geolinguistics and Digital Humanities. He is currently one of the Principal Investigators on a project on the geolinguistic investigation of diatopic variation in colloquial German.

BERNHARD PÖLL is Professor of Romance Linguistics at the University of Salzburg (Austria). His main research focus is sociolinguistics, in particular the diversification of linguistic norms. Besides his interest in sociolinguistics, he has also worked extensively on lexicology/(meta-)lexicography as well as grammar theory.

SIMON PRÖLL is Professor of German Linguistics at the Albert Ludwig University of Freiburg i. Br. (Germany). When he is not pondering about the socio-historical and grammatical evolution of Standard German, he works on quantitative methods for analysing data on language variation and change. In addition to a geostatistical approach to historical and contemporary dialect variation, he has developed a strong interest in typological differences between Germanic standard and non-standard varieties.

ERIK SCHLEEF is Professor of English Linguistics at the University of Salzburg (Austria). His research focuses on language variation in the UK and the acquisition of variation; language attitudes, perception and discrimination; language, gender and sexuality; cross-cultural communication; and language standardisation in contemporary English and its history. He is co-editor of the *Routledge Sociolinguistics Reader* and co-author of *Doing Sociolinguistics: A Practical Guide to Data Collection and Analysis*.

Felix Tacke is Professor of Spanish Linguistics at the Philipps University of Marburg (Germany). His primary research focus lies in the study of pluricentric tendencies in Romance standard language cultures, language planning and the promotion of minority languages. He co-edited the *Manual of Standardization in the Romance Languages* (2020), a reference work focusing on processes of formal standardisation. His other research areas involve the study of historical syntax and pragmatics in Romance languages and on demonstrative constructions, in particular.

Marco Tamburelli is Professor of Linguistics (Bilingualism) at Bangor University (UK). His research focuses on bi- and multilingualism, both from a representational perspective and from a sociolinguistic/comparative perspective. He has worked on Bilingual First Language Acquisition, particularly in relation to syntactic and phonological transfer effects in simultaneous bilinguals, and on lexical and phonological development in typically developing children and in children with SLI. He also works on social and comparative aspects of bilingual development, particularly on regional and minority languages, the measurement of linguistic attitudes in bilingual communities and on the measurement of phonetic distance and intelligibility rates in linguistic continua.

Judit Vári is currently a Postdoctoral Researcher and Academic Coordinator at Munich Centre of Linguistics as well as the Graduate School of Language & Literature at the Ludwig Maximilians University in Munich (Germany). Judit completed her PhD at Bangor University (UK) before working at the Excellence-Cluster *The Politics of Inequality* at the University of Konstanz (Germany). In her PhD thesis, she investigated attitudes towards speakers of standard and non-standard language varieties in the German-speaking community of Belgium and the Grand Duchy of Luxembourg. Her research interests lie at the intersection between sociolinguistics and social psychology of language including implicit and explicit perceptions of minority and minoritised speakers in general as well as in the political domain.

HEIMIR F. VIÐARSSON is Adjunct Lecturer and Assistant Professor at the School of Education of the University of Iceland in Reykjavík (Iceland). He was a guest researcher in Leiden, Utrecht and Cologne and taught Icelandic language courses in Amsterdam and Ghent. In 2019, he received his PhD from the University of Iceland for his thesis on *Socio-Syntactic Variation and Change in Nineteenth-Century Icelandic: The Emergence and Implementation of a National Standard Language*. He has participated in several research projects, including *Icelandic Dialect Syntax* (IceDiaSyn), a part of the ScanDiaSyn research network and the *Nordic Centre of Excellence Norms, Language Change and Language Variation in the 19th Century* (LCLV19), and he currently works in a comparative research project on Icelandic grammar books and grammar writing at the School of Education, University of Iceland.

MONIKA WINGENDER is Professor of Slavic Linguistics in the Department of Language, Literature and Culture of the Justus Liebig University in Gießen (Germany). She has researched and published widely on the grammar and semantics of Russian and on the sociolinguistics, contact linguistics, language politics and conflicts in the Slavic language area and its multilingual societies as well as on standardisation of Slavic languages. She is currently leading two research projects on language ideologies and language policies in multilingual Russia.

Index

HISTORICAL SOCIOLINGUISTICS

Studies on Language and Society in the Past

SERIES EDITORS

Nils Langer
Stephan Elspaß
Joseph Salmons
Wim Vandenbussche

The interdisciplinary field of Historical Sociolinguistics seeks to reveal the impact of language development on society and the role of individuals and society in the changing forms and usage of language. This book series is aimed at sociolinguists and social historians who are keen to publish studies on the social history of languages, the interaction of linguistic practices and society, and the sociological significance of linguistic variation with a historical dimension. The purpose of the series is to provide empirically supported studies that will challenge and advance current language historiographies, which often continue to present the history of particular languages as necessarily leading to the creation of a standard or prestige variety. Of particular interest are topics such as the following: language myths and language ideology, historical multilingualism and the formation of nation-states, the sociolinguistics of minority and regional languages, the rise of urban vernaculars, immigrants and their languages, the role of prescriptive grammarians, and the social history of pidgins and creoles.

Book proposals from historians and linguists working on any language in any period are welcome, in particular those that include a comparative dimension as well as those with a strong empirical foundation. The language of publication is primarily English, though other languages may be considered. The editors guarantee that all publications in this series have been subjected to external and anonymous peer review. The four series editors and twenty-six members of the advisory board are all members of the Historical Sociolinguistics Network (HiSoN).

Vol. 1 Vladislav Rjéoutski, Gesine Argent and Derek Offord (eds):
European Francophonie: The Social, Political and Cultural History of an International Prestige Language
2014. ISBN 978-3-0343-1703-0

Vol. 2 Anna Havinga and Nils Langer (eds):
Invisible Languages in the Nineteenth Century
2015. ISBN 978-3-0343-1968-3

Vol. 3 Luis Andrade Ciudad:
The Spanish of the Northern Peruvian Andes: A Sociohistorical and Dialectological Account
2016. ISBN 978-3-0343-1790-0

Vol. 4 Gijsbert Rutten and Kristine Horner (eds):
Metalinguistic Perspectives on Germanic Languages: European Case Studies from Past to Present
2016. ISBN 978-3-0343-1853-2

Vol. 5 Markus Schiegg and Judith Huber (eds):
Intra-Writer Variation in Historical Sociolinguistics
2023. ISBN 978-1-80079-703-1

Vol. 6 Sabrina Freund:
Lexikalische Variation in Patientenbriefen des 19. und frühen 20. Jahrhunderts
2024. ISBN 978-1-80374-271-7

Vol. 7 Samantha M. Litty and Nils Langer (eds):
Language Ideology, Policy, and Practice: Focus on Minoritized Languages Past and Present
2025. ISBN 978-1-80374-556-5

Vol. 8 Sarah van Eyndhoven:
'Ane end of an auld song?': Macro and Micro Perspectives on Written Scots in Correspondence during the Union of Parliaments Debates
2025. ISBN 978-1-80374-501-5

Vol. 9 Stephan Elspaß, Imke Mendoza, Bernhard Pöll and Erik Schleef (eds):
Double Standards: Codified Norms and Norms of Usage in European Languages (1600–2020)
2025. ISBN 978-1-80374-940-2