

Global and local perspectives on language contact

Edited by

Katrin Pfadenhauer

Sofia Rüdiger

Valentina Serreli

Contact and Multilingualism 7



Contact and Multilingualism

Editors: Isabelle Léglise (CNRS SeDyL), Stefano Manfredi (CNRS SeDyL)

In this series:

1. Lucas, Christopher & Stefano Manfredi (eds.). Arabic and contact-induced change.
2. Pinto, Jorge & Nélia Alexandre (eds.). Multilingualism and third language acquisition: Learning and teaching trends.
3. Hakimov, Nikolay. Explaining Russian-German code-mixing: A usage-based approach.
4. Sánchez Moreano, Santiago & Élodie Blestel. Prácticas lingüísticas heterogéneas: Nuevas perspectivas para el estudio del español en contacto con lenguas amerindias.
5. Migge, Bettina & Shelome Gooden. Social and structural aspects of language contact and change.
6. Coler, Matt & Andrew Nevins (eds.). Contemporary research in minoritized and diaspora languages of Europe.
7. Pfadenhauer, Katrin, Sofia Rüdiger & Valentina Serreli (eds.). Global and local perspectives on language contact.

ISSN (print): 2700-8541

ISSN (electronic): 2700-855X

Global and local perspectives on language contact

Edited by

Katrin Pfadenhauer

Sofia Rüdiger

Valentina Serreli

Katrin Pfadenhauer, Sofia Rüdiger & Valentina Serreli (eds.). 2024. *Global and local perspectives on language contact* (Contact and Multilingualism 7). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/345>

© 2024, the authors

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/> 

ISBN: 978-3-96110-431-4 (Digital)

978-3-98554-089-1 (Hardcover)

ISSN (print): 2700-8541

ISSN (electronic): 2700-855X

DOI: 10.5281/zenodo.10438503

Source code available from www.github.com/langsci/345

Errata: paperhive.org/documents/remote?type=langsci&id=345

Cover and concept of design: Ulrike Harbort

Typesetting: Sebastian Nordhoff

Proofreading: Agnes Kim, Amir Ghorbanpour, Axel Fanego Palat, Alexandru Craevschi, Eliane Lorenz, Elliott Pearl, Emilia Slavova, Fabio Gasparini, Georgios Vardakis, Giorgia Troiani, Inga Hennecke, Jacopo Falchetta, Janina Rado, Jean Nitzke, Jeroen van de Weijer, Kate Bellamy, Klaudia Hahn, Linda Bäuml, Lisa Schäfer, Mary Ann Walter Miriam Neuhausen, N. Sokolnik, Rebecca Madlener, Tihomir Rangelov, Tom Bossuyt, Troy E. Spier

Fonts: Libertinus, Arimo, DejaVu Sans Mono, Arabtype

Typesetting software: Xe_{La}T_EX

Language Science Press

xHain

Grünberger Str. 16

10243 Berlin, Germany

<http://langsci-press.org>

Storage and cataloguing done by FU Berlin

Freie Universität  Berlin

Contents

Introduction

Katrin Pfadenhauer, Sofia Rüdiger & Valentina Serreli **iii**

I Globalization and power relations

- 1 Translanguaging in Bulgarian street signs**
Emilia Slavova **3**
- 2 The “unnecessary” use of French in Moroccan Arabic: Social
discriminant or collaborative enterprise?**
Jacopo Falchetta **37**
- 3 The discursive construction of code-switching in Yanito among the
young population of Gibraltar**
Marta Rodríguez García **63**

II Language contact and group identity

- 4 Language change and stance in a remote Mennonite community in
Canada**
Miriam Neuhausen **99**
- 5 Why a language dies: The case of Bəṭaḥrēt in Oman**
Fabio Gasparini **129**
- 6 Creating a class of elite Palestinian multilinguals in Israel: Reflections
on research in late capitalism**
Nancy Hawker **151**

III Migration and language contact

- 7 **Mobile and complex: A West African linguistic repertoire**
Klaudia Dombrowsky-Hahn & Axel Fanego Palat 183
- 8 **Asymmetrical power relations between languages of Equatorial Guinea: Views from the migration context in Madrid**
Sandra Schlumpf 215

IV Focus on microlinguistic outcomes of language contact

- 9 **Systems of pragmatic markers in contact: Processes and outcomes**
Inga Hennecke 245
- 10 **Phonic adaptation of Spanish Anglicisms in Mexico and Spain: A corpus data analysis**
Linda Bäumlér 273
- 11 **Imitating the Arabic model: The case of valency-increasing operations in the Old Spanish translation of the Arabic *Kalīla wa-Dimna***
Hans-Jörg Döhla 301

Discussion

- 12 **Reconciling the global and local in language contact**
Yaron Matras 333
- Index** 348

Introduction

Katrin Pfadenhauer, Sofia Rüdiger & Valentina Serreli

University of Bayreuth

This introductory chapter presents a short summary of the field of contact linguistics, setting the stage for the following chapters and establishing the interdisciplinary and multi-methodological mindset of the authors in this edited volume.

1 Previous research on language contact: An overview

Language contact is a major force in the evolution of language and language change; it presupposes the coexistence and interaction of different languages at the individual, community, and societal level. In essence, language contact depends on the existence and extent of contact between speakers. It is also fundamentally inter-idiolectal, meaning that groups are only in contact to the extent that their individuals and their idiolects are in contact with each other (Mufwene 2001; Mufwene & Vigouroux 2012).

Since the 1950s, linguistic studies on language contact have attempted to individuate the processes and outcomes of contact in linguistic systems, producing models to explain contact-induced change; that is, the effect of contact in language structures (see Weinreich 1953; Thomason & Kaufman 1988; Thomason 2001; van Coetsem 2000; Heine & Kuteva 2005; Bakker & Matras 2013; for a general overview, see Winford 2003; Matras 2009; Hickey 2010). Among the subfields of interest within the broad area of language contact are the study of code-switching and code-mixing (e.g., Muysken 2000; Myers-Scotton 2002; Gardner-Chloros 2009), dialect contact (e.g., Trudgill 1986; Kerswill & Williams 2000), and pidginization and creolization (e.g., Arends et al. 1994; Siegel 2008; Aboh 2015; Velupillai 2015; Knörr & Filho 2018).

While most of the studies mentioned above share a conception of languages or varieties as discrete and bounded systems, a more recent approach to language contact, starting from the conceptualization of language as a social practice, has focused on how linguistic resources – meaning either language systems or single



features – are used by speakers to create social meaning and to negotiate identities (e.g., Sánchez Moreano & Blestel 2021). Contact-induced variation is thus observed in language repertoires rather than languages as discrete systems, and new frameworks have been devised to help describe and explain the fluidity of language practices, including *crossing* (Rampton 1995), *linguaging* (Canagarajah 2007; Makoni & Pennycook 2007; Jørgensen & Juffermans 2011), *polylingualism* (Jørgensen 2008), and *translinguaging* (García & Wei 2014). Finally, a broader view of contact in context is found in studies that assume language evolution takes place within local linguistic ecologies and is influenced by the complex interrelation between demographic, social, economic, and intrinsically linguistic factors (Mufwene 2001; Pennycook 2010).

Languages are brought into contact in a number of ways, the most obvious of which is migration – voluntary or forced, individual or large-scale, temporary or permanent – whereby the interaction of speakers with different histories and repertoires results in the reinterpretation of existing affiliations and identities and the negotiation of new ones (Aalberse et al. 2019; Canagarajah 2019). Yet in the era of globalization, characterized by easier communication and circulation of information, heterogeneity and diversity can also be found within local populations, both in metropolises at the center of power and in smaller communities on the periphery (Blommaert 2010; Pietikäinen et al. 2016).

2 Global and local perspectives on language contact

This edited volume recognizes both traditional and innovative language contact research and brings together contributors whose expertise covers various languages, with the goal of examining both general phenomena of language contact and specific features in a comparative approach. The mixture of research on Romance languages, Semitic languages, and Germanic languages is unique in its scope and leads to fruitful synergistic effects. While English, so often ubiquitous, is indeed at the center of some of the studies, a particular asset of this volume is not only the variety of contact settings included, but also the range of subject languages under investigation. Furthermore, the inclusion of Arabic studies enriches the volume due to the increasing relevance of Arabic in contact settings, for instance in the context of migration.

A particular focus is on past and present contact scenarios between languages of unbalanced political and symbolic power. Insights from diachronic and synchronic language-contact research have important linguistic and societal implications, especially considering current global migration streams. The contributions

in this volume describe phenomena of language contact between and with Romance languages, Semitic languages, and various forms of English. This sees a diverse range of contact constellations and settings covered in the volume (these are grouped together here but still roughly follow the order of the chapters): English and Bulgarian (Slavova); English and Spanish (in Mexico and Spain: Bäumlér; in Gibraltar: Rodríguez García); Moroccan Arabic and French (Falchetta); Pennsylvania German English and Canadian English (Neuhausen); Arabic and Hebrew (Hawker), Baṭṭhari and Arabic (Gasparini); German, French, and West African languages (Dombrowsky-Hahn & Fanego Palat); Spanish, French, Bantu languages, the Portuguese-based creole Fá d'Ambô, and the English-based pidgin Pichi (Schlumpf); Manitoban French and English (Hennecke); and Old Spanish and Arabic (Döhla).

3 Structure and scope of this volume

The contributions in this volume are grouped into four thematic sections, depending on the respective main focus of each chapter: I) globalization and power relations; II) language contact and group identity; III) migration and language contact; and IV) microlinguistic outcomes of contact. The volume concludes with a final discussion chapter by Yaron Matras, who brings together the different perspectives from the individual chapters and consolidates them in his integrated theory of language contact (Matras 2009/2020).

The first thematic section of the volume is concerned with language contact scenarios involving asymmetrical power relations due to globalization. Many factors, including transnational flows, migration (see Section III of this volume), displacement, tourism, and the internet, have contributed to an increase in language contact scenarios around the world – scenarios that have been extensively discussed in the research community (see, for example, the contributions in Collins et al. 2009 and Coupland 2010 for an overview; also see Blommaert 2010). The three contributions in this section focus specifically on settings involving languages in uneven power relations, such as local and global linguistic forces in the urban landscape of Sofia, Bulgaria (Slavova), the language of the former colonizer – namely, French – in the Moroccan setting (Falchetta), and a diagglossic situation involving various languages – British English, Gibraltar English, Andalusian Spanish, and Yanito – in Gibraltar (Rodríguez García).

The section begins with Emilia Slavova's survey on translanguaging in the linguistic landscape of Sofia. This mainly involves Bulgarian and English, but also includes other languages, and becomes all the more salient in the script

choices employed by venues in the inner city. Next, Jacopo Falchetta examines the use of French lexical items in colloquial Arabic in Morocco, showing how they fulfil prestige-related functions while also (potentially) contributing to language learning. These forms of code-switching and code-mixing thus constitute part of the habitual and indexical linguistic practices of the surveyed speakers. Code-switching practices are also at the center of Marta Rodríguez García's contribution on Yanito in Gibraltar. Rodríguez García focuses on young adults – a sociodemographic group whose sociolinguistic realities have been neglected in previous linguistic research – and shows how thematic factors and topics, discursive elements, and idiosyncratic factors influence switching between the various languages present in the linguistic ecology of Gibraltar.

Section II focuses on the connection between language contact and group identities and ideologies. The three chapters in this section show that the local evolution of linguistic repertoires is subject to changes related to different factors – demographic, social, economic, political, and intrinsically linguistic. Such changes affect individual linguistic repertoires or idiolects, through the filter of the speaker's attitudes, and dominant ideologies, and the extent to which such ideologies circulate within and across groups. In turn, regular individual patterns of behavior converge toward a norm or practice routine. Miriam Neuhausen investigates the performative use of Canadian Raising patterns by Old Order Mennonites in southern Ontario. She combines quantitative and qualitative analyses and shows that members of the community manipulate the use of linguistic features to foreground their Mennonite identity or show identification with the Canadian mainstream (it being understood that an individual's behavior is related to the extent of their exposure to English). The dialect contact situation described involves a closed community of Pennsylvania German English speakers and the wider community of Canadian English speakers, affording Neuhausen a unique look at the local and translocal forces at play. The community under investigation in Fabio Gasparini's chapter is also closed, and in fact even smaller, as his informants make up the last ten speakers of Baḥari in Oman. Gasparini shows how vernacular and standard varieties of Arabic have replaced Baḥari in every domain within the local community and presents this linguistic shift as part of a process of identity reshaping and a shift toward the local Bedouin Arabic culture. In other words, language change is related to demographic, economic, and social changes in the Baḥāhira community and the nation-state ideologies that accompany them. Community boundaries and nation-state ideologies also play an essential role in Nancy Hawker's chapter on the use of Arabic and Hebrew by Palestinian and Jewish Israeli politicians. Hawker analyses the pragmatic functions related to the use of Arabic by Palestinian politicians in the Knesset (the

Israeli parliament) and aptly shows how the extent to which existing norms are manifested, contested, and renegotiated by speakers ultimately depends on the speakers' relative power. However, this power is related to their identities, and their linguistic behavior depends on which of these identities the speakers intend to foreground in any given interaction. Altogether, the three chapters in Section II deal with language ideologies, identification processes, and the construction of the speakers' identities in relation to their in-group and out-group relations. They show that language, as a means through which boundaries are constructed and communicated, allows for the erasure of difference – through accommodation (Neuhausen) and language shift (Gasparini) – or the foregrounding of difference – through maintenance of diverging features (Neuhausen) and the display of multilingualism (Hawker).

Section III highlights two studies investigating matters of language contact and migration. Migration and diaspora have been widely theorized and investigated in linguistic work on contact settings (see, e.g., Piller 2016 for a compilation of influential publications). Ever more important roles in settings involving language contact and migration are played here by the notions of transnational labor migration (e.g., Lorente 2018), neoliberalism and postmodern societies (e.g., Allan & McElhinny 2017), and superdiversity (e.g., Creese & Blackledge 2018). Both chapters in this section focus on African migrant groups in Europe, but adopt very different methodological approaches and analytical perspectives. Klaudia Dombrowsky-Hahn and Axel Fanego Palat analyze the history of language acquisition processes in the migration trajectories of people from West Africa who settled in the German Rhine-Main region, and show that linguistic interferences in their use of German and French can be attributed to their wide language repertoires as mobile speakers rather than their L1 only. Sandra Schlumpf then explores the attitudes of Equatoguinean migrants in Madrid toward the different languages spoken in Equatorial Guinea, explaining these attitudes on the basis of language ideologies and the sociohistorical and cultural background of the speakers. She also looks at the questions of power, prestige, and social inclusion/exclusion implied in such attitudes, with the aim of questioning traditional colonial and Eurocentric hierarchies, ideologies, and approaches.

Section IV features three chapters, each of which discusses individual structural phenomena that are the result of specific language contact situations. The contributions are concerned with different aspects of linguistic systems, from (morpho)syntax to phonetics. Inga Hennecke discusses processes of contact-induced language change in pragmatic markers, using data from Franco-Manitoban, a contact variety of Canadian French isolated for a long time from other varieties of French and in strong language contact with English. Hennecke explains why

English and French markers show different outcomes and why certain markers undergo contact-induced language change, while others of the same system do not. Linda Bäumlér investigates the variation in integration of Anglicisms among Spanish-speaking natives from Mexico and Spain. By analyzing the phonetic realization of a given grapheme, and whether speakers imitate the English model or follow the Spanish equivalent, Bäumlér demonstrates that the urban versus rural variable is mainly responsible for the variation, while regional differences are neutralized in the era of globalization. The typologically driven chapter by Hans-Jörg Döhla on TMA-marking in Spanish closes the section. He considers the grammaticalization pathways of the morphological devices observed in the tense–aspect intersection on the basis of Old Spanish, as translated from Arabic literature and the Hebrew bible in the Middle Ages, and additional data from Spanish creole languages.

The volume concludes with a final discussion chapter by Yaron Matras, who draws together the different strands of research presented in the individual contributions and describes the emerging meta themes. Besides providing a comprehensive synthesis of the presented approaches to language contact, the discussion chapter offers an outlook on where the field is headed and incentives for follow-up studies. This synthesis by Matras, who has broad expertise in the fields of contact linguistics and multilingualism, not only consolidates the findings but also adds yet another layer of insight to the exploration of cross-language contact dynamics, which is particularly important in light of the plethora of contact settings covered in the volume.

Acknowledgements

We are very grateful for the financial support provided by the *WiN-Academy of the University of Bayreuth*, which enabled us to host the *Language Contact through Time and Space* conference in 2021, on which the contributions in this edited volume are based. In particular, we would like to thank Mabel Braun for her patience and encouragement in our endeavors.

We would also like to express our gratitude to all our colleagues who were involved in the peer review process and who provided invaluable feedback and input. Tobias Berner and Davide Frapporti provided invaluable help with LaTeX. Our thanks also go to the series editors at Language Science Press, Isabelle Léglise and Stefano Manfredi, who have been most helpful in the process of putting this volume together. And, of course, *thank you, Danke, merci, gracias*, and *شكراً* to our contributors for filling this volume with their insights and perspectives on language contact around the world.

The volume editors and authors of the introductory chapter are listed in alphabetical order, which is indicative of equal contribution.

References

- Aalberse, Suzanne, Ad Backus & Pieter Muysken. 2019. *Heritage languages: A language contact approach*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Aboh, Enoch Oladé. 2015. *The emergence of hybrid grammars: Language contact and change*. Cambridge: Cambridge University Press.
- Allan, Kori & Bonnie McElhinny. 2017. Neoliberalism, language, and migration. In Suresh Canagarajah (ed.), *The Routledge handbook of migration and language*, 79–101. London/New York: Routledge.
- Arends, Jacques, Pieter Muysken & Norval Smith. 1994. *Pidgins and creoles: An introduction*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Bakker, Peter & Yaron Matras. 2013. *Contact languages: A comprehensive guide*. Berlin/Boston: De Gruyter Mouton.
- Blommaert, Jan. 2010. *Sociolinguistics of globalization*. Cambridge: Cambridge University Press.
- Canagarajah, A. Suresh. 2007. The ecology of global English. *International Multilingual Research Journal* 1(2). 89–100.
- Canagarajah, A. Suresh. 2019. Changing orientations to heritage language: The practice-based ideology of Sri Lankan Tamil diaspora families. *International Journal of the Sociology of Language* 2019(255). 9–44.
- Collins, James, Mike Baynham & Stef Slembrouck. 2009. *Globalization and language in contact: Scale, migration, and communicative practices*. London: Bloomsbury Publishing.
- Coupland, Nikolas (ed.). 2010. *The handbook of language and globalization*. Chichester: Wiley-Blackwell.
- Creese, Angela & Adrian Blackledge (eds.). 2018. *The Routledge handbook of language and superdiversity*. London: Routledge.
- García, Ofelia & Li Wei. 2014. *Translanguaging: Language, bilingualism and education*. London: Palgrave Macmillan UK.
- Gardner-Chloros, Penelope. 2009. *Code-switching*. Cambridge: Cambridge University Press.
- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Hickey, Raymond. 2010. *The handbook of language contact*. Hoboken: Wiley.

- Jørgensen, J. Normann. 2008. Polylingual languaging around and among children and adolescents. *International Journal of Multilingualism* 5(3). 161–176.
- Jørgensen, J. Normann & Kasper Juffermans. 2011. Languaging. *Toolkit for Transnational Communication in Europe*. <http://www.toolkit-online.eu/docs/languaging.html>.
- Kerswilll, Paul & Ann Williams. 2000. Creating a new town koine: Children and language change in Milton Keynes. *Language in Society* 29(1). 65–115.
- Knörr, Jacqueline & Wilson Trajano Filho. 2018. *Creolization and pidginization in contexts of postcolonial diversity: Language, culture, identity*. Leiden: Brill.
- Lorente, Beatriz P. 2018. *Scripts of servitude: Language, labor migration and transnational domestic work*. Bristol: Multilingual Matters.
- Makoni, Sinfree & Alastair Pennycook (eds.). 2007. *Disinventing and reconstituting languages*. Clevedon; Buffalo: Multilingual Matters.
- Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.
- Matras, Yaron. 2020. *Language contact*. 2nd edn. Cambridge: Cambridge University Press.
- Mufwene, Salikoko & Cécile Vigouroux. 2012. Individuals, populations, and timespace: Perspectives on the ecology of language. *Cahiers de Linguistique* 38(2). 111–137.
- Mufwene, Salikoko S. 2001. *The ecology of language evolution*. Cambridge: Cambridge University Press.
- Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.
- Myers-Scotton, Carol. 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Pennycook, Alastair. 2010. *Language as a local practice*. Milton Park: Taylor & Francis.
- Pietikäinen, Sari, Helen Kelly-Holmes, Alexandra Jaffe & Nikolas Coupland. 2016. *Sociolinguistics from the periphery: Small languages in new circumstances*. Cambridge: Cambridge University Press.
- Piller, Ingrid. 2016. *Language and migration: Critical concepts in linguistics*. London/New York: Routledge.
- Rampton, Ben. 1995. *Crossing: Language and ethnicity among adolescents*. London: Longman.
- Sánchez Moreano, Santiago & Élodie Blestel (eds.). 2021. *Prácticas lingüísticas heterogéneas. Nuevas perspectivas para el estudio del español en contacto con lenguas amerindias*. Berlin: Language Science Press.
- Siegel, Jeff. 2008. *The emergence of pidgin and creole languages*. Cambridge: Cambridge University Press.

- Thomason, Sarah Grey. 2001. *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Blackwell.
- van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Velupillai, Viveka. 2015. *Pidgins, creoles and mixed languages: An introduction*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Weinreich, Uriel. 1953. *Languages in contact: Findings and problems*. The Hague: Mouton.
- Winford, Donald. 2003. *An introduction to contact linguistics*. Hoboken: Wiley.

Part I

Globalization and power relations

Chapter 1

Translanguaging in Bulgarian street signs

Emilia Slavova

St. Kliment Ohridski University of Sofia

Translanguaging has become a popular concept in sociolinguistic studies. It marks a shift from thinking of languages as bounded and homogeneous systems, closely linked to national identities, towards perceiving them as mobile resources which are in constant contact with other languages and are combined creatively to create new meanings. One area where translanguaging can easily be observed is urban spaces, where people from diverse linguistic and cultural backgrounds meet. Bulgarian street culture is an interesting case in point. After a long period of isolation and repression behind the Iron Curtain, Bulgarian street culture has flourished in recent years. The monolingual paradigm, strictly observed in the past, has been replaced by the multilingual and translanguaging trend. This has given rise to street signs (shop windows, café, and restaurant names, etc.) in multiple languages and combinations of languages, used in creative ways not only to address an international customer base, but also to appeal to local citizens with an open mindset, multilingual repertoires, and a global, transcultural identity. Using a linguistic landscape approach, this chapter explores the street culture and naming practices in central Sofia as the city negotiates between a local and a global identity.

1 Introduction

In the context of globalization, weakened national borders, and superdiversity (Blommaert & Rampton 2011), languages are also losing their strict boundaries, creating new, hybrid forms. These linguistic phenomena clearly challenge the traditional view of languages as bounded, homogeneous systems that are used in isolation from other languages and are closely linked to their speakers' national



Emilia Slavova. 2024. Translanguaging in Bulgarian street signs. In Katrin Pfadenhauer, Sofia Rüdiger & Valentina Serreli (eds.), *Global and local perspectives on language contact*, 3–35. Berlin: Language Science Press. DOI: 10.5281/zenodo.10497369



identity and sense of belonging. Translanguaging is a fairly recent sociolinguistic term which captures these new ways of meaning-making and identity formation and allows researchers to explore these phenomena from a novel perspective beyond the traditional categories inherited from earlier generations.

One area where translanguaging is particularly salient in late modernity is urban linguistic spaces. Large, cosmopolitan capital cities are places where people of diverse linguistic and cultural backgrounds congregate. This diversity is encouraged through various semiotic signs, including the languages displayed in public spaces and the way these languages mix to index new identities (Blommaert 2010, Blommaert & Maly 2014, Gorter 2013).

Bulgarian street culture is a good example. As a result of the country's opening up to the Western world after the fall of the Iron Curtain, its accession to the EU in 2007, and the overall effects of globalization and increased mobility, there has been considerable transformation of urban spaces in Sofia. Both central (as the capital city) and peripheral (with regard to Western European city culture), the city strives to define itself as cosmopolitan, open, and part of a larger European context. This trend is countered by a growing sense of nationalism and tribalism, often fuelled (or at least benevolently supported) by the media and populist politicians.

In this context, this chapter explores the dominant linguistic norms and the linguistic landscape in the centre of Sofia; specifically, the naming practices employed by shop, café, and restaurant owners. These are considered against the background of dominant ideologies and naming practices during socialist times. The analysis seeks to establish how a post-communist country processes its difficult history and current social upheavals, negotiates its relationship with the outside world, and redefines its own identity through the linguistic choices in its urban landscapes.

2 Translanguaging: Theoretical background

Mobility has become a central factor in current sociolinguistic studies, rethinking language beyond narrow national confines and looking at it in its wider social context and deep historical embeddedness (Blommaert & Rampton 2011). Languages are no longer seen as normative systems that exert their power over the individual, who has to follow rules and observe external standards; instead, they are regarded as resources on which individuals can draw, exploring them creatively in order to express themselves freely. Dominant language norms are shifting from a focus on monolingualism and idealized linguistic purity towards a

preference for hybridity and diversity (Jørgensen et al. 2011). A burgeoning field of research has appeared urging sociolinguistics “to unthink its classic distinctions and biases and to rethink itself as a sociolinguistics of mobile resources” (Blommaert 2010: 1). A more detailed analysis of these shifts follows.

2.1 ‘Language’: A reconceptualization

The conceptualization of language as a self-contained, homogeneous system with clear boundaries has been seriously challenged in recent decades. After the emergence of sociolinguistics in the 1960s, “language” was no longer regarded in isolation, devoid of any social context. Contemporary approaches view it as a social practice rather than an abstract, rule-governed system. The very concept of a “language” as a discrete unit is under scrutiny by critical discourse analysts, and there is a growing recognition that separate “languages” are a social construct, ideological creations asserting linguistic boundaries to reinforce political ones (Blommaert & Rampton 2011, Saraceni 2015, Jørgensen et al. 2011).

Drawing on recent studies, Blommaert & Rampton argue that “named languages” (such as “English”, “German”, and “Bengali”) are ideological constructions related to the emergence of the nation-state, and that in “differentiating, codifying and linking a “language” with a “people”, linguistic scholarship itself played a major role in the development of the European nation-state” (Blommaert & Rampton 2011: 4). While conceding that the traditional idea of a “language” still has immense ideological power, they observe that there has been a major shift in fundamental ideas about language, language groups, and communication, which leads to a very different approach in the study of languages.

Rather than working with homogeneity, stability and boundedness as the starting assumptions, mobility, mixing, political dynamics and historical embedding are now central concerns in the study of languages (Blommaert & Rampton 2011: 3).

Similarly, Jørgensen et al. (2011: 27) argue that traditional concepts of different “languages” as bounded systems are sociocultural constructs, ill-suited to capturing the reality of real language use in late-modern superdiverse societies. They point out that drawing borders between closely related languages (for example, German and Dutch) using purely linguistic criteria is not possible, since the idea of individual languages is based on ideology rather than real-life language usage.

Languages, Mario Saraceni argues, “don’t just exist alongside each other, but merge, blend, mesh, coalesce into a symbiosis where traditional labels struggle to

find a place” (2015: xi). In places where monolingualism is the official language policy, there may be strong institutional support for the national language at the price of suppressing neighbouring minority languages, so that a language may exist as a separate entity. This, however, is hardly the case in contemporary societies characterized by mobility, interconnectedness, and great linguistic and cultural diversity. A new theoretical approach is needed.

2.2 Linguaging and translanguaging

As it has recently been argued, instead of theorizing language as a static, monolithic, rule-governed system, it may be more appropriate to look at actual language usage, which some scholars refer to as languaging (Swain 2006, Lankiewicz & Wąsikiewicz-Firlej 2014). The concept was popularized by Merrill Swain, who sees languaging as a dynamic process of using language to make meaning and to shape knowledge and experience, of “coming-to-know-while-speaking” (2006: 97).

This definition is extended by Lankiewicz & Wąsikiewicz-Firlej (2014: 4), who describe languaging not only as a way of knowing and meaning-making, but of identity formation and the “unbridled, natural way of using language beyond the normative constraints of a language”.

Other related terms, with slight differences in meaning, have been put forward to capture this dynamic, constraint-free language usage observed in present-day superdiverse societies: plurilingualism, polylinguaging, poly-lingual languaging, metrolingualism, translingualism, transidiomaticity, translingual writing, translingual practices, multiliteracies, pluriliteracy, fused lects, heterography, and ludic Englishes (Canagarajah 2013: 9).

The translingual paradigm, Canagarajah (2013: 6) contends, highlights two key significant concepts: firstly, communication transcends individual languages; secondly, communication transcends words and involves diverse semiotic resources and ecological affordances. While *multilingual* may suggest an additive relationship between languages, *translingual* moves away from the *mono/multi-dichotomy* which he sees as reductive (Canagarajah 2013: 7).

In unifying these definitions, translanguaging can be understood as a dynamic process of creative language use, playful meaning-making, and identity formation, in which languages, unconstrained by notions of homogeneity, boundedness, and isolation, mix and mesh freely, transcending language boundaries and involving various semiotic resources.

2.3 Language norms: From monolingualism to translanguaging

In spite of these recent theories, the monolingual paradigm is so pervasive that people still perceive languages as monolithic, homogeneous, bounded systems, isolated from other languages (Jørgensen et al. 2011). Until not long ago, such attitudes were part of the mainstream language ideology. Before the rise of sociolinguistics in the 1960s, code-switching was typically considered deviant behaviour, and bilinguals were thought of as imperfect language users. The monolingualism norm stipulates that only one language should be spoken at a time; persons who command two or more languages should speak each language “purely”, without mixing it with other languages (Jørgensen et al. 2011: 33).

This attitude has changed in recent years, and the monolingualism norm has been replaced by the bilingualism/multilingualism norm, according to which people who have a good command of two (or more) languages will use their full linguistic repertoire, switching between languages when necessary. Jørgensen et al. (2011) contrast this with the polylinguaging norm, which does not require full command of two (or more) languages, but instead allows language users to employ whatever linguistic features are at their disposal, no matter how well they know the languages involved and whether they can make claims to “possess” the languages (Jørgensen et al. 2011: 34).

I would argue that the polylinguaging norm is not a “norm” per se. While the monolingualism norm has established itself as the unmarked, default norm, explicitly upheld and imposed by institutions such as the media and the educational system, the bilingualism/multilingualism norm has recently become acceptable and widespread, for example, in road signs which have to be bilingual. The polylinguaging norm, however, can hardly claim to be an institutionalized norm. It is rather the absence, or relaxation, of rigid language norms which require linguistic purity and norm-observance. So I would suggest treating it as a trend rather than a norm, and in line with the previous discussion of translanguaging, I would replace “polylinguaging” with “translanguaging”, as the “trans” prefix works better than “multi” and “poly” at capturing the fluidity of languages and the transcendence of linguistic boundaries.

3 Linguistic landscaping

The phenomena discussed above could be observed not only in the speech of individual speakers, but also in contexts such as the written texts in shop windows, and the names of cafés and restaurants found at a given location. I decided to explore the streets of central Sofia and observe how language visibility, social

attitudes, and linguistic norms have changed in the relatively short time span of a few decades. The method chosen is described as linguistic landscape studies in Blommaert & Maly (2014: 1) – an attempt to create “accurate and detailed inventories of urban multilingualism” through investigating “the presence of publicly visible bits of written language: billboards, road and safety signs, shop signs, graffiti and all sorts of other inscriptions in the public space, both professionally produced and grassroots”. An overview of the field and methodological approaches used in recent years can be found in Gorter (2013).

Linguistic landscape studies in Bulgaria are still rare, but there are exceptions, such as Politov & Lozanova’s (2015) study of multilingualism in Sofia’s urban spaces, Krusteva’s (2017) exploration of the linguistic landscape in the Sofia subway, and Atanassova-Divitakova’s (2017) survey of the linguistic landscape of Veliko Turnovo.

This study limits its scope to the examination of the naming practices and linguistic choices in shop windows and the names of cafés and restaurants (collectively labelled “street signs” or just “signs” for the purposes of the study), as observed in the centre of Sofia in the summer of 2021. They have been contextualized and historicized against the social changes which have taken place in recent years (discussed in Section 4). An exploration of peripheral neighbourhoods or locations outside the capital would undoubtedly have produced different results, but this could be the subject of a different study.

The signs have been divided into three large groups with three subgroups each, with somewhat fuzzy boundaries between them. They are discussed in Sections 5, 6, and 7 (as presented in Table 1).

Table 1: Classification of signs

5 Monolingual signs	5.1 Monolingual signs in Bulgarian
	5.2 Monolingual signs in English
	5.3 Monolingual signs in other languages
6 Bilingual signs	6.1 Bilingual signs in Bulgarian and English
	6.2 Bilingual signs in Bulgarian and another language
	6.3 Bilingual/multilingual signs in English and other languages
7 Translingual signs	7.1 Transliterated words and names
	7.2 Mixing languages and scripts
	7.3 Word play and double meaning

The data consist of a total of 240 signs, 160 of which were randomly collected in the central streets of Sofia. They provide illustrative examples of the different types of naming practices employed, and the various linguistic strategies of code-mixing, script-mixing, and translanguaging. The other 80 signs were collected systematically in a single street, Ivan Shishman Street, a trendy shopping destination with predominantly small shops, cafés, and restaurants. The systematic approach allows for a quantitative analysis of the naming practices employed (discussed in Section 8).

A separate corpus of 120 images was collected for the sociohistorical analysis in Section 4. It was gathered from a vast online collection of images from the socialist period (Danov & Galabov n.d.) and a selection of images borrowed from the same database (HighViewArt 2014), as well as a recent photographic collection of old street signs surviving to this day (Dnevnik 2011).

The analysis attempts to answer the following questions:

1. What is the overall pattern in terms of linguistic norms, languages, and language-mixing strategies?
2. What motivates those linguistic choices?
3. How are these practices culturally and historically embedded?

The following sections seek to answer these questions.

4 Sociolinguistic context

In order to appreciate the change which has taken place in Sofia's linguistic landscape in recent decades, a look at the linguistic, social, cultural, and historical context is in order. Three major factors are outlined below, namely the importance of the Cyrillic alphabet in the identity construction of Bulgarians, the Soviet influence after 1944 and Bulgaria's isolation from the Western world, and the nation's opening up to the outside world in 1989 after the fall of the Iron Curtain, which coincided with the spread of globalization and English being recognized as a global language.

4.1 The Cyrillic script and Bulgarian identity

Even though Bulgarians tend to think of the Cyrillic script as the Bulgarian alphabet, they actually share it with some 250 million other people. Russia accounts

for the greatest number of users, and for this reason, it is often wrongly considered to be “the Russian alphabet” (Nikolova 2021). Serbia, North Macedonia, Montenegro, Belarus, and Ukraine also use it, albeit with some modifications (the Latin alphabet is also used in Serbia, Montenegro, and Bosnia). While it is mainly associated with Slavic languages, not all Slavic languages use it. At the same time, a number of Finno-Ugric, Turkic, Iranian, and Caucasian languages have adopted it. Some of them shifted from Cyrillic to Latin after the disintegration of the Soviet Union in 1991, or are currently in the process of doing so (Iliev 2013).

For foreign visitors who are used to Roman letters, however, Cyrillic can pose a serious challenge. Deciphering street signs and other texts is largely impossible, even when they hide familiar borrowings. Having some familiarity with the Cyrillic letters may help. In Table 2, they have been divided into three groups: letters which are identical to Latin letters, so-called false friends which look the same but have different values in Cyrillic and Latin, and unique letters which have no counterpart in the Latin script.

Table 2: The Bulgarian Cyrillic alphabet

Identical letters	False friends	Unique letters
<i>Aa</i> – a	<i>Bb</i> – v	<i>Бб</i> – b
<i>Ee</i> – e	<i>Cc</i> – s	<i>Гг</i> – g
<i>Kk</i> – k	<i>Hh</i> – n	<i>Дд</i> – d
<i>Mm</i> – m	<i>Pp</i> – r	<i>Жж</i> – zh
<i>Oo</i> – o	<i>Ii</i> – i	<i>Зз</i> – z
<i>Tt</i> – t	<i>Xx</i> – h	<i>Лл</i> – l
	<i>Yy</i> – u	<i>Пп</i> – p
		<i>Фф</i> – f
		<i>Цц</i> – ts
		<i>Чч</i> – ch
		<i>Шш</i> – sh
		<i>Щщ</i> – sht
		<i>Ъъ</i> – ʌ as in cut
		<i>Юю</i> – yu
		<i>Яя</i> – ya

It is worth noting that the upper-case and lower-case forms of Cyrillic letters may not coincide with the respective Latin letters; for example, the upper-case

form of <M> is identical, but the lower-case <m> in Bulgarian is spelled <М>, while the Bulgarian <м> corresponds to the cursive form of lower-case <T>.

Commissioned in the late ninth century by the First Bulgarian Empire during what was considered The Golden Century in Bulgarian history, the Cyrillic alphabet was developed for the purposes of Old Church Slavonic in parallel with the introduction of Christianity. It replaced the official Bulgarian Glagolitic script, created by Constantine the Philosopher (Saint Cyril), but was later (somewhat misleadingly) named after him (Iliev 2013).

The episode is ingrained in the memories of countless students through the educational system, which has glorified the Cyrillic alphabet and turned it into an indelible part of Bulgarian national identity. There is even a national holiday celebrating the Day of Bulgarian culture and the Slavonic alphabet, complete with its own hymn. For this reason, any attempt to replace the Cyrillic script with the Latin one, or even to use Latin in informal communications, is met with resistance. Bulgarians feel intimately related to the Cyrillic script, which has developed as a strong marker of national ideology and identity construction (Norman 2019).

4.2 The Soviet influence

Bulgaria was liberated from Ottoman rule with the help of Russia at the end of the nineteenth century and was occupied by the Red Army in 1944. Isolated from the West behind a symbolic Iron Curtain, it lived under a totalitarian regime for 45 years. The Soviet influence was palpable: a centralized economy; a dominant socialist ideology which permeated every aspect of public life; no consumerist culture, no brand names, almost no advertising; long queues for deficit goods; strong political, economic, and cultural ties with the Soviet state and the other countries in the Eastern bloc; and very limited contact with the rest of the world.

Russian functioned as a *de facto* second language in Bulgaria. It was studied from an early age in state schools. A pen-pal exchange system with Soviet children was established for schoolchildren. On Fridays, the national television network broadcast the news in Russian, without translation, straight from the Ostankino Tower in Moscow. The Friday news was followed by the obligatory Soviet film. Russian was used as a *lingua franca* for international travel, conferences, and academic exchanges within the Eastern bloc.

One of the most well-stocked bookshops was the Russian bookshop named *Lenin* in central Sofia, as found in the sociohistorical corpus:

- (1) *Съветски книги КНИЖАРНИЦА ЛЕНИН*
savetski knigi knizharnitsa lenin
'Soviet books BOOKSHOP LENIN'

Next to it was another landmark location indexing the symbolic closeness to the Soviet capital:

- (2) *Ресторант Москва*
restaurant moskva
'Restaurant Moscow'

Both signs were written in Bulgarian, but since the Russian alphabet is similar and the two languages are from the same language family, sharing multiple cognate words, the signs were completely understandable to Russian speakers. A large number of other words were imported from Russian during that period. The influence of other foreign languages (mostly French, English, and German), which were popular in the country before the Second World War, was suppressed during the Soviet regime. Even though monolingualism was the dominant norm at the time, imposed by a nationalist state ideology, a rigid school system, and controlled state media, Russian had symbolically permeated the Bulgarian language and street signs through many borrowings and cultural references.

Instead of advertisements of brand names and luxury goods, cityscapes were dominated by ideological slogans, such as the large imperative neon slogan on a wall facing the Central Departmental Store:

- (3) *Гледайте новите български и съветски филми!*
gledaite novite balgarski i savetski filmi
'Watch the new Bulgarian and Soviet films!'

The sociohistorical corpus of archival images from Sofia in the 1960s, 1970s, and 1980s reveals many no-name shop signs simply stating what they sell: *Хляб* 'bread'; *Плодове и зеленчуци* 'fruit and vegetables'; *Месо и колбаси* 'meat and meat products'; *Яйца и птици* 'eggs and birds'; *Бързи закуски* 'fast food'; *Шапки* 'hats'; *Обувки* 'shoes'; *Очила* 'glasses'; *Парфюмерия* 'perfumery'; *Платове* 'fabrics'; *Трикотаж* 'knitwear'; *Аптека* 'pharmacy'; *Кафе* 'café'.

Most shops sold the same limited range of goods at fixed prices. Only a few had names, for example the women's clothing store *Валентина* 'Valentina' (appropriately, a female name).

With almost every shop and street sign written in Cyrillic, a traveller from outside the Eastern bloc would have found it difficult to navigate the city unless

they knew Cyrillic. Even taxis (*такси*) and hotels (*хотел*) had their signs written in Bulgarian. One of the few shops with a name in Latin script was *Corecom*, a chain of tax-free shops operating between the 1960s and 1990s. Deriving its name from the French *Co(mptoir de) re(présentation et de) com(merce)*, ‘a direction of representation and commerce’, *Corecom* traded in a foreign currency (US dollars), sold luxury Western goods to foreign visitors and privileged locals, and gave Bulgarians a taste of the coveted Western consumerist culture.

In some cases, bilingual signs did appear. One pharmacy in the corpus had a sign in Bulgarian and a smaller sign in French: *Аптека | Pharmacie*. For ideological reasons, French rather than English was the preferred language for addressing foreigners. Air France had an office in central Sofia and a sign in both Bulgarian and English. Bulgarian Airlines also had a sign in both languages. Some hotels followed a similar pattern (*Grand Hotel Balkan | Гранд хотел Балкан*).

Apart from a few exceptions, however, the Cyrillic script was omnipresent. When, by some curious twist of international trade relations with the West, Coca-Cola started producing its iconic beverage in Bulgaria in 1965 (Bulgaria was the first country in the Eastern bloc to produce the drink and to export it to other Eastern bloc countries), it designed its bottles with Cyrillic text on them: *Кока-Кола*, ‘Coca-Cola’.

The geographical references found in the streets of Sofia provided self-references (*Grand Hotel Sofia*, *Grand Hotel Bulgaria*, *Hotel Rodina* [‘motherland’], *Grand Hotel Balkan*), referred to Soviet realia (*Park Hotel Moskva*, surprisingly spelled in Latin script), or represented capitals from the Eastern bloc written in Cyrillic (*Prague café*, *Berlin ice-cream parlour*, *Budapest Hungarian restaurant*). Another popular location in central Sofia was the elegant Russian restaurant called *Crimea*. International influences were limited by the invisible Iron Curtain; Soviet references dominated the cityscape. *Moscow* was not only the name of a central restaurant, but also a hotel and a cinema; *Lenin* was the name of a bookshop, a boulevard, a neighbourhood in the city, and a university.

4.3 Opening up to the outside world

The Soviet influence and Bulgaria’s isolation from the Western world came to a rather abrupt end in 1989, following the fall of the Berlin Wall. As the country gradually transitioned to a democratic form of government and a free-market economy, the USSR disintegrated, and multinational companies entered the market, Russian was also replaced as the dominant second language, with English quickly taking its place. A new era had begun, not only in political and economic terms, but also linguistically.

This coincided with a global phenomenon: the spread of American-driven globalization in the 1990s, and the assertion of English as the global language of international communication. The first McDonald's restaurant opened in Sofia in 1995. Other global Western brands followed suit. The liberation of the market meant that everybody could now start a business and open a shop, and multiple new outlets appeared in the streets.

English was seen as the language of freedom and democracy, of free trade and global culture; as a result, many people rushed to learn it. Organizations such as the British Council and the US Peace Corps facilitated the process by bringing in language-teaching expertise, training materials, teachers, and teacher trainers. English language teaching schools flourished. English became the main foreign language taught in schools, quickly replacing Russian. Other Western languages (French, German, and Spanish) were also in high demand, though they did not see the explosive success English enjoyed.

As Bulgaria transitioned politically and economically and opened up to the Western world, it applied for membership of the European Union. The country became a member in 2007, giving Bulgarian citizens new opportunities to travel, study, and work abroad. Many EU citizens also moved to Bulgaria permanently. English became the main foreign language. Bulgarian, on the other hand, became one of the 24 official languages of the European Union, and the Cyrillic script joined the Latin and Greek scripts as one of now three officially recognized scripts in the EU.

However, in spite of the significant cultural shift and Bulgaria's opening up to the outside world, the country is still split between its pro-Western orientation, liberal-democratic views, and EU membership, on the one hand, and strong nationalistic, xenophobic, anti-EU, anti-democratic undercurrents, often coupled with nostalgia for the Soviet past and its symbols, on the other. How much the two opposing trends have influenced Sofia's linguistic landscape is the object of study in what follows.

5 Monolingual signs

The first group of street signs includes signs written in a single language. The signs have been divided into three groups: signs in Bulgarian, signs in English, and signs in other languages. Contrary to the monolingual norm discussed earlier, the monolingual signs include not only signs in Bulgarian, but also signs in English and signs in other languages, as well as languages that have not been identified.

5.1 Monolingual signs in Bulgarian

Some of the signs in Bulgarian follow the well-known pattern of simply stating what they sell, without any naming and branding, as in Example (4), Figure 1. These signs seem to be relics from socialist times. The new shop owners must have taken a conscious decision to preserve the original design and to keep the spirit of the past.

- (4) Meco
meso
'Meat'



Figure 1: *Meco* 'Meat'. Photo: Emilia Slavova.

While Example (5) seems to follow a similar pattern, it uses very different stylistics. In contrast to the main text that simply states a job description (baker), the accompanying text points to a recent trend for craft foods. It follows a Western pattern by giving the start of the enterprise ("2013"). The contemporary design of the sign clearly indexes a hipster outlet for young, middle-class, outward-looking people.

- (5) ХЛЕБАР. Занаятчийски хляб София 2013
hlebar zanayatchiiski hlyab sofia 2013
'Baker. Craft bread Sofia 2013'

Магазин №10 (Example 6) follows a similar pattern: it simply calls itself "shop". The sign has a distinctly vintage style and could either be a relic salvaged from

socialist times or a newly designed sign in a retro style. The artistic décor and eccentric handmade clothes clearly show that the name is intentionally bland and seemingly dated, while the shop itself is anything but.

- (6) Магазин № 10
magazin nomer deset
'Shop No. 10'

Бутик №2 (Example 7), a trendy clothes shop, has a similarly bland name with an arbitrary number, but the word choice follows a different logic. It is a transliteration of a European word, 'boutique', which entered the Bulgarian language in the 1990s. Most likely borrowed from English ('trendy fashion shop'), the word is actually French in origin, with cognates in other European languages: *bottega* (Italian); *apotheca* ('storehouse' in Latin); and even *ἀποθήκη* (*apothékē* in Ancient Greek). In Cyrillic, however, these etymological connections are completely lost. Boutiques entered the linguistic market in the 1990s, alongside a range of small, independent shops, often selling handmade goods, in contrast to the centrally produced socialist garments of previous times.

- (7) Бутик № 2
Butik nomer dve
'Boutique No. 2'

Ревю (Example 8) also uses an English word, 'review', which in Bulgarian has a narrower meaning and refers mostly to fashion. Likewise, *Бустро* (Example 9) is a French borrowing with presumably Russian roots, hardly recognizable in its Cyrillic orthography.

- (8) Ревю
revyu
'Review'
- (9) Бустро
bistro
'Bistro'

The same can be said for *Doner Miami* (Example 10), selling Turkish doners but evoking a faraway destination in the USA, and *The Master and Margarita* (Example 11), a flower shop bearing the name of the Russian literary classic by the dissident (Kyiv-born) writer Mikhail Bulgakov. The latter points to a distinction many Bulgarians make between Russian culture, to which they feel deeply connected, and Soviet ideology, which they despise.

- (10) Дюнер Маями
dyuner mayami
'Doner Miami'
- (11) Майстора и Маргарита
maistora i margarita
'The Master and Margarita'

Finally, *Къоше* (Example 12), a trendy souvenir corner shop, has a name which sounds distinctly Bulgarian, but is actually Turkish in origin (*köşe* 'corner'). Due to the many years of Ottoman rule, Turkish borrowings have often blended with Bulgarian, or are stylistically marked as more intimate, informal, and "folksy" compared with their more neutral Bulgarian counterparts.

- (12) Къоше
kyoshe
'Corner'

5.2 Monolingual signs in English

Some of the monolingual signs in English found in the streets of Sofia have a distinctly global orientation. *Simple: Taste the world* (restaurant) is a case in point. Others have English names without any specific cultural reference, such as *Social Café*, *Farmers Soups and Sandwiches*, *The Little Things* (restaurant), *The Gourmet House* (fine china), *LightHouse* (candles), and *Orange. Books Music Stationery* (a multistorey bookshop). The only example of a shop with a distinctly English orientation is *Elephant Bookstore* (Example 13).

- (13) ELEPHANT Bookstore. Vintage & English Gifts

Tourists in Bulgaria no longer need to know Cyrillic in order to find their hotel. Hotel names, such as *Grand Hotel Sofia*, are usually spelled in Latin and are easy to read. Many traditional gift shops (Example 14) also have names spelled in Latin. In a trendy tourist shop, *GIFTED* (Example 15), Figure 2, you can not only buy souvenirs, but also leave your luggage, book a visit to the *Red Flat* to explore life under socialism, or book a *Free Sofia Tour* trip, all in English.

- (14) SOUVENIRS of Bulgaria
- (15) GIFTED. Gallery, Luggage lockers, Urban culture hub, The Red Flat, Free Sofia Tour



Figure 2: GIFTED. Photo: Emilia Slavova.

5.3 Monolingual signs in other languages

It would be easy to assume that the global status of English would allow it to dominate the linguistic landscape. But a closer look reveals that many of the signs in Sofia's streets are in other languages, including German (Example 16), Spanish (Example 17), French (Example 18), and Italian (Examples 19, 20, and 21).

- (16) BISMARCK. Frischer Fisch und Deutsches Bier seit 2020
'Bismarck. Fresh fish and German beer since 2020'
- (17) La Casa del Habano
'The Havana House'
- (18) Bistrot L'Etranger. Maison fondée en 2001
'Bistro The Stranger. House founded in 2001'
- (19) Gelateria Caffeteria CONFETTI
'Ice-cream parlour Café Confetti'
- (20) Trattoria Neapolitana. Pizza & Aperitivo Est. 1867
'Neapolitan eatery. Pizza and Aperitif Established 1867'
- (21) La Bottega Due Piani
'The two floors shop'

Italian seems to be the most popular language related to cuisine in central Sofia, and even though many Italian words have become sufficiently transparent

in Bulgarian (*pizza, caffeteria*), others are not (*trattoria, bottega, piani*). They are used to convey foreignness and – by association with Italian – to promise tasty food, even if the meaning may be opaque.

This is not the case with *Банкович* (Example 22), a popular Serbian restaurant. The sign is spelled in Cyrillic, but the Serbian-sounding family name *Brankovich* and a slight change in the spelling of *restaurant*, which is missing the final <m> (*ресторан* versus *ресторант*) signifies that it is in Serbian.

- (22) Ресторан Банкович
restoran bankovich
'Bankovich restaurant'

Another Balkan cuisine establishment in the corpus is *Beyzade* (Example 23), Figure 3. The name, meaning 'noble, aristocratic, from a good family' in Turkish, is opaque in Bulgarian, but clearly indexes Oriental foreignness, further supported by the familiar Turkish desserts and pastry *baklava, kunefe*, and *borek*.

- (23) BEYZADE. Baklava, kunefe & borek



Figure 3: *Beyzade*. Photo: Emilia Slavova.

A similar effect is achieved by the distinctly Japanese-sounding name *Tanoshi* (Example 24). A Romanized variant of the Japanese word (meaning 'fun, pleasant, delightful'), the name carries no specific meaning to most Bulgarians, apart from indexing Japanese language and culture, which is its main function.

- (24) TANOSHI

6 Bilingual signs

The next group in the corpus consists of signs in two languages: Bulgarian and English, Bulgarian and another language, or English and another language (and, occasionally, more than one other language). While in some instances the bilingual signs have identical meanings in both languages, this is not the case most of the time. The two languages usually complement each other, with part of the information only available in one language. Since the graphic design features of the signs are lost in the examples below, a dividing line | is used to separate the two languages for the sake of clarity.

6.1 Bilingual signs in Bulgarian and English

Some outlets choose to have their names spelled both in English and Bulgarian, with one script (English in this case) considerably larger than the other (Example 25).

- (25) Souvenirs from Bulgaria | Сувенири от България
suveniri ot balgariya
'Souvenirs from Bulgaria | Souvenirs from Bulgaria'

In other cases, such as Example (26), only one of the words is translated, and the other is not.

- (26) Книги Хеликон | Books
knigi helikon
'Helikon Books | Books'

Elsewhere, the two languages function together as two parts of a whole, code-mixing between Bulgarian and English, as in Examples (27), (28), and (29).

- (27) Магазини | KINKY
magazini
'Kinky Shops'
- (28) Химическо чистене | Clean the world
himichesko chistene
'Dry cleaning | Clean the world'
- (29) Cozy маркет | Кафе алкохол цигари
market kafe alkohol tsigari
'Cozy market | Coffee, alcohol, cigarettes'

A number of outlets choose a mixture of Bulgarian and English for their signs, code-switching between the two languages, while the cultural reference points elsewhere. Examples include *Royal Thai* (Example 30), *Lokah* (Example 31), where a Sanskrit word is spelled in Latin script, and *JOVAN The Dutch Bakery* (Example 32), followed by descriptions in Bulgarian which clarify what they offer: Thai food, Indian goods, and Dutch baked goods.

- (30) Royal Thai | Тайландски ресторант
tailandski restorant
'Royal Thai | Thai Restaurant'
- (31) Lokah | Индийски стоки
indiiski stoki
'Lokah | Indian Goods'
- (32) JOVAN. The Dutch Bakery | Холандската Фурна
holandskata furna
'Jovan. The Dutch Bakery | The Dutch Bakery'

Another example of a mixture of Bulgarian and English, indexing a third culture (Arabian) by means of the fictional character Ali Baba from the popular book *Arabian Nights*, is shown in Example (33).

- (33) Бърза закуска | Ali Baba Fast food
barza zakuska
'Fast snack | Ali Baba Fast food'

A curious example in this group is Example (34), Figure 4, a shop selling a traditional Bulgarian snack: *banitsa* (a kind of pastry, also known as *borek* in Turkish). It is considered typically Bulgarian, even though variants can be found in other Balkan cuisines. The "traditional" in the name clearly evokes nationalistic feelings, also discretely signalled by a thin line at the top representing the Bulgarian national flag (white, green, and red). In spite of this, the shop cannot resist the temptation to add an English word (*handmade*) to its very traditional Bulgarian name, making it much less traditional than it purports to be.

- (34) Handmade | Традиционна БАНИЦА
traditsionna banitsa
'Handmade | Traditional Pastry'



Figure 4: Traditional *banitsa*. Photo: Emilia Slavova.

The next example also uses Bulgarian and English to refer to a third culture: Russian. The Russian restaurant *Arbat* (Example 35) is in the same location as the *Moscow* restaurant from Example (2) above. Having previously been an Italian restaurant (*Corso*), it is now Russian once again, but under a new name, *Arbat*, after a fashionable street in Moscow. The name of the street is the same in Russian and Bulgarian, but the spelling of *Russian* (single <c> rather than double) and *restaurant* (with a final <m>) clearly index that the sign is in Bulgarian rather than Russian (*руски ресторант* versus *русский ресторан*). The presence of Bulgarian and English, but no Russian, in the name of a Russian restaurant in Sofia, is a rather bold decision on the part of the owners.

- (35) АРБАТ. Руски ресторант | Russian restaurant
arbat ruski restorant

‘Arbat. Russian restaurant | Russian restaurant’

6.2 Bilingual signs in Bulgarian and another language

The next examples mix Bulgarian and another language: French (Example 36), Italian (Example 37), and Russian (Example 38), Figure 5. Interestingly, in the final example, the store name is in Russian (*Берёзка*), recognizable by the diacritics over the second “e”, but the descriptor (‘Russian food store’) is spelled in Bulgarian – *пysкy*, rather than in Russian, *русский* – even though the inverted word order (the adjective following the noun) follows the Russian syntax rather than

the Bulgarian. Another interesting detail concerning the name is that *Берёзка* was the name of the Western goods shops in the USSR, similar to the Bulgarian *Corecom* discussed earlier.

- (36) BONJOUR JULIETTE | Пекарна
pekarna
'Bonjour Juliette | Bakery'
- (37) Gelato di Natura Sofia | Натурални фермерски продукти
naturalni fermerski produkti
'Gelato di Natura Sofia | Natural farmer's products'
- (38) Берёзка | Гастроном руски
biryozka gastronom ruski
'Birch tree | Russian food store'



Figure 5: *Beryozka*. Photo: Emilia Slavova.

The *Remedium* pharmacy (Example 39) has its descriptor (*pharmacy*) written both in Bulgarian and English, and its name in Cyrillic. However, the word *remedium* does not exist in Bulgarian. It is a Cyrillicized form of a Latin word meaning 'cure' with multiple cognates in European languages: English: *remedy*; French: *remède*; Italian: *rimedio*; Portuguese: *remédio*; Romanian: *remediu*; Spanish: *remedio*, and so on. Spelled in Cyrillic, the meaning is lost on foreign visitors, while Bulgarians could only understand it through another language they are familiar with.

- (39) Аптека | Pharmacy | Ремедиум 2
apteka remedium 2
'Pharmacy | Pharmacy | Remedium 2'

6.3 Bilingual/multilingual signs in English and another language

The next group of signs mix English and another language, without any recourse to Bulgarian: a combination of Italian and English (Example 40); Spanish and English (Example 41); French and English (Example 42); Turkish and English (Example 43); Arabic and English (Example 44); and Portuguese and English (Example 45). The final example, *Garafa*, is a misspelling of the Portuguese/Spanish word *garrafa*, of Arabic origin (meaning 'bottle'), but functions as a transliteration of the Bulgarian word *zapaḡa* (/garafa/), meaning an open-top flask for pouring wine (similar to the English *carafe*).

- (40) MAMMA MIA | Restaurant and Pizzeria
(41) EL GRADO | Jewellery
(42) Bijoux | Trendy
(43) Djanam | Duner & Burger
(44) HAMAM | Home of textile
(45) Garafa | Wine Shop

El Shada (Example 46), Figure 6, is rather difficult to analyse. Seemingly in Italian (also supported by the food served), the definite article *El* appears to be in Spanish. The name may be a variation of *El Shaddai* (Hebrew for 'God Almighty'; there is a Gospel song with the name *El-Shada*), or an incorrect spelling of *L'sciadà* ('rolling pin' in Ladin), or it could have some other etymology. It could also be emulating foreignness without having any particular meaning. The other words in the name, however, are clear enough. *Pizza* is now a well-established Bulgarian word. *Pasta* has also become popular as meaning 'Italian dish made of dough', although until the 1990s, it was used with a different meaning in Bulgarian: either 'cake', 'pastry', or 'paste'. *Vino* was borrowed so long ago that it is now an authentic Bulgarian word with no alternative.

- (46) EL SHADA | Pizza. Pasta. Vino

The next Example (47) mixes three languages. The Italian name *Ottimo* ('optimal, excellent', a superlative form of *buono* 'good') is coupled with a Bulgarian and an English reminder to call and pick up the food.



Figure 6: *El Shada*. Photo: Emilia Slavova.

- (47) OTTIMO pizza & pasta | Обади се и вземи за вкъщи | Pick up
 obadi se i vzemi za vkashti
 ‘Ottimo pizza & pasta | Call and take away | Pick up’

The final example in this group, *NAMOOΣ* (Example 48), also uses more than two languages. *Namoos* (the alternative spelling) is the only word in the corpus with a Greek letter in it (<Σ>). It creates the impression that it is a Greek word, but it is actually Arabic (*nāmūs* ‘law’, ‘custom’, or ‘honour’). The reason it was chosen was probably the connection with the Ancient Greek word *nómos* (*NOMOS*) ‘law, custom’. The next word refers to the Greek island of Mykonos. The slogan is in English, and the booking details are in Bulgarian.

- (48) NAMOOΣ Mykono! | Mediterranean love & food |
 телефон за резервации
 telefon za rezervatsii
 ‘Namoos Mykono! | Mediterranean love & food |
 phone number for bookings’

7 Translingual signs

The final group in the corpus includes signs which mix languages and scripts in unexpected ways, such as transliterated words and names, code-mixing and script-mixing, and word play and double meaning. Linguistic signs are often mixed with graphic signs to enhance understanding and augment the meaning.

7.1 Transliterated words and names

The first example is *Pileto* (Example 49), Figure 7. It represents a transliteration of the Bulgarian word for ‘chicken’, with an image of a chicken incorporated into the O sign of the logo, so that the meaning can be derived from the graphic sign, even if the word is unclear. The rest of the sign is in English. The word may be opaque to non-Bulgarian speakers, but it serves as an attention-getter and adds a Bulgarian “vibe” to a name that is easy to read, transcending not only the boundary between separate languages, but also linguistic and visual signs.

- (49) PILETO | Hand made
‘The bird | Handmade’



Figure 7: *Pileto*. Photo: Emilia Slavova.

Bilkova (Example 50), a trendy bar popular with locals and foreigners alike, uses a similar strategy. It has retained its Bulgarian name and identity, but has mixed it with Latin script, a global identity, and a visual symbol: a leaf intertwined with the letter <V>.

- (50) BAR BILKOVA est. 1991
‘Herbal bar est. 1991’

Another example of the same strategy is the optician’s shop called *Ochila* (Example 51), the Bulgarian word for ‘glasses’, transliterated in Latin but completely opaque from the perspective of English. The Bulgarian word for ‘optician’ does

not contribute to making the meaning any clearer, but a diacritic line connecting the O and C letters of *OChila*, imitating a graphic image of glasses, disambiguates the transliterated word.

- (51) ОПТИКА OCHILA
optika ochila
'Optician's glasses'

The reverse strategy is used in another optician's shop, *New Vision* (Example 52). This time, the English name is written in Cyrillic, requiring not only familiarity with the Cyrillic script from foreigners, but also familiarity with the English language from Bulgarians, who would not be able to make sense of the name otherwise.

- (52) ОПТИКА Ню Вижън
optika new vision
'New Vision Optician's'

7.2 Mixing languages and scripts

In the next group of examples, we can see how the mixing of languages and visual signs works in practice. In *bebeshore* (Example 53), Figure 8, a diminutive, informal, endearing term meaning a baby/toddler, there is a rather interesting case of language- and script-mixing. On the one hand, the word *бебе* in Bulgarian means 'baby'; *shore* is spelled in English but functions as the Bulgarian diminutive suffix, *-уоп*. However, the English meaning of *shore* is then activated through the visual symbol of waves at the end of the word. *Bebster* has no particular meaning in Bulgarian and is a blend between *baby* and *hipster*.

- (53) Бебешore | bebster style
bebeshor
'Little baby | baby style'

In *Zelena Art Gallery* (Example 54), there is code- and script-mixing between Bulgarian and English. The main word (*Zelena*) is spelled almost in Cyrillic, with the exception of the first letter, <Z>, which is spelled in Latin script. The green colour of the sign plays a part in visually conveying the message.

- (54) ZEJEHA | ART GALLERY
Zelena
'Green | Art Gallery'



Figure 8: *Bebe shore*. Photo: Emilia Slavova.

Code- and script-mixing are also used in a dairy shop, *МЛЕКАРНИЦА Раково* (Example 55). Everything is written in Bulgarian apart from one key word: *milk*, spelled *милк*. The Bulgarian word is *мляко* (*mlyako*), not *milk*. As the English word is used instead, it is transliterated in Cyrillic, with the exception of the first letter, <m>. While the upper-case letters <M> coincide in both languages, the lower-case ones differ, and <m> in Bulgarian should be read as /t/ (*tilk*). However, the graphic association with the popular chocolate brand *milka* makes the code-mixing transparent and ensures that <m> is read as the Latin letter /m/, even though the rest is in Cyrillic.

- (55) МЛЕКАРНИЦА Раково. Вкусът на млякото | милк
mlekarnitsa rakovo vkusat na mliakoto milk
'Dairy Rakovo. The taste of milk | milk'

7.3 Word play, double meaning

The final group of signs use word play and double meaning as an attention-grabbing strategy. Such is the case with *Ave New* (Example 56), a pun on the English word of French origin *avenue*, where the word is intentionally misspelled in order to evoke an association with something new – quite appropriate for a fashion shop.

- (56) AVE NEW

Bar Maze (Example 57) and *Mazetto* (Example 58) may seem to be in English/Italian, but both evoke a Bulgarian word: *мазе* /maze/, meaning ‘basement’ and ‘the basement’ (Bulgarian adds the definite article at the end of words as a suffix, and *-to* is the suffix for neuter gender). In spite of the maze drawn on the graphic sign of *Bar Maze*, the Bulgarian word easily comes to mind for Bulgarians. The *Mazetto* sign leads to a bar in an actual basement, so it is even easier to interpret it as a Romanized Bulgarian word. Few people would know the Italian word *mazzetto* ‘bunch’, and the word itself was probably selected because of the phonetic similarity with a Bulgarian word, while referencing a (misspelled) Italian word.

(57) BAR MAZE

(58) MAZETTO | Vintage concept

The next sign, *SupaStar* (Example 59), Figure 9, uses a Bulgarian word, *supa*, a borrowing meaning ‘soup’, and makes it a part of an English compound, *superstar*. Knowledge of Bulgarian is needed in order to work out what this establishment offers; otherwise, customers have to rely on the image of a bowl of soup in the middle of the round sign.

(59) SupaStar | Домашни Супи и Сандвичи
domashni supi i sandvichi

‘Soup Star | Homemade Soups and Sandwiches’

The final example is *Mandzha Stantsia: Zona Mexicana* (Example 60), Figure 10. *Mandzha* is a very colloquial word for ‘cooked food’ in Bulgarian. While it feels as intimate and folksy as a Turkish borrowing, it is actually an Italian borrowing: *mangia* is the imperative form of the Italian verb *mangiare* ‘eat’. The combination of a noun modifying another noun is a typical English construction which has entered Bulgarian in recent years. The Bulgarian language requires an adjective rather than a modifying noun in this initial position. *Stantsia* derives from *station* and is a well-established word in Bulgarian, no longer perceived as a borrowing. *Zona* ‘zone’ is another borrowing which is now firmly established and no longer feels foreign. The text is spelled in Cyrillic, but the message is a mixture of languages. The Bulgarian transliteration *мехикана* /mehikana/ reflects the Spanish pronunciation. There is a reversal of noun and adjective which is rather untypical of Bulgarian syntax but is becoming more popular due to interference from other languages.



Figure 9: *SupaStar*. Photo: Emilia Slavova.

- (60) Манджа станция. Зона мекхикана
mandzha stantsia zona mehikana
'Food station. Mexican zone'

The result of all this is a charming linguistic and cultural mash-up, perhaps serving as a metaphor for everything that has happened in the Bulgarian language after the firm grip of nationalist/communist ideology was shaken off, and the free linguistic market took its place.

8 Quantitative analysis

The quantitative analysis of the street signs on one particular street, Ivan Shishman Street (Table 3), reveals that of the 80 signs observed and analysed, multilingual signs prevail at 46%. Combined with the bilingual signs in Bulgarian and English, the percentage of bilingual/multilingual signs increases to 54%. Monolingual signs in Bulgarian make up only 17% of the signs, closely followed by signs in English, at 15%. The share of translanguing signs is similar, at 14%.

It must be said, however, that the classification is rather arbitrary, as not all the signs can be unambiguously assigned to a single category.

In many cases, it is only the different script that distinguishes a Bulgarian word from a foreign borrowing.

The main conclusions that can be drawn from quantifying the signs on a central street in Sofia are that



Figure 10: *Mandza Station. Zona Mexicana*. Photo: Emilia Slavova.

Table 3: Street signs on Ivan Shishman Street

Type	Number	Percentage
Monolingual signs in Bulgarian	14	17%
Monolingual signs in English	12	15%
Bilingual signs in Bulgarian and English	6	8%
Multilingual signs	37	46%
Translingual signs	11	14%

1. bilingual and multilingual signs prevail, making up more than half of the signs;
2. under a fifth of the signs are monolingual signs in Bulgarian;
3. English monolingual signs almost equal the number of Bulgarian monolingual signs;
4. English is widely used as an international language but has not displaced many other foreign languages, which are also visible; and
5. compared to the pre-1989 period, the translingual trend is steadily gaining ground.

However, as discussed earlier, these results are related to the central location and trendiness of the selected street and may differ considerably in more marginal areas of the city or country.

9 Discussion

The linguistic landscape study of Sofia's central streets reflects the rapid transition from a closed, centralized economy and state-controlled, heavily ideologized naming practices to an open, free-market economy, where individual shop owners use a wide range of linguistic resources to reach a wider customer base.

The data show a clear shift from a monolingual norm, dominant under socialism and part of the state's ideology of national homogeneity, towards multilingual and translingual trends, characteristic of the new openness and global outlook.

It would be easy to assume that the global status of English would allow it to dominate the linguistic landscape and to suppress all other languages, but this is not the case. The languages identified in the corpus include Italian, French, Spanish, Turkish, Russian, German, and Arabic, as well as some unidentified languages. English is used widely and functions as a neutral language, a *lingua franca*, indexing other cultures (Greek, Turkish, Indian, Thai, Russian, etc.), without necessarily being connected to either British or American culture. This shows that in spite of the explosive success of English in Bulgaria in recent decades, the country has opened up to many other languages and cultures, and Sofia's streets are communicating this to locals and visitors alike.

Borrowings, code-switching, language- and script-mixing, blending, double meaning, and word play are typical strategies used in addressing Bulgarian and international customers and evoking a mixture of local and global identities. In spite of the relatively small share of translingual signs found in the quantitative corpus, the pattern is clear: once languages have been allowed to coexist, mix, and mesh in the streets freely, they do so in different ways and to varying degrees of entanglement.

The reasons to use the above-mentioned languages and mixing strategies can be summarized as follows:

1. Addressing English-speaking travellers and visitors to the city;
2. Addressing outward-oriented Bulgarian customers;
3. Attention-grabbing via some unusual choice of language or word;

4. Signalling a global orientation and a cosmopolitan identity;
5. Retaining a local Bulgarian character within this global identity.

The analysis of the corpus shows that central Sofia today is a vibrant, multilingual, and cosmopolitan destination like many other cities in Europe; a far cry from the isolated, ideologized, monolingual place it used to be under communism. Most of the Soviet references are gone; with the exception of *Park Hotel Moskva*, all the other establishments mentioned above have changed their names and purpose. *Cinema Moscow* is a shopping centre. The former *Lenin* bookshop was a *Happy Sushi* restaurant until recently; now it is an investment management company called *Expat Capital*.

In spite of the obvious influence of American-dominated globalization in recent years, there are few explicit references to American culture and multiple references to other cultures through the medium of the English language and the Latin script. The share of monolingual signs in Bulgarian is small, but in most multilingual signs, there is a tendency to index Bulgarian identity within a larger global identity. Cyrillic script is blended with foreign words and vice versa, in an effort to transcend national and linguistic boundaries and play with languages and meanings in creative ways.

And even though there has been a disturbing nationalistic and anti-globalist wave in recent years, this has not affected the linguistic landscape of central Sofia in any significant way. Instead of the expected overt displays of national identity, there is a clear cosmopolitan identity with a distinctly Bulgarian flavour.

10 Conclusion

This study confirms that political and historical upheavals have serious linguistic consequences. This is clearly seen in the linguistic landscape of Sofia during the transition from a totalitarian communist state to a free-market economy.

Monolingualism is only viable under strict rules and a nation-state ideology which enforces it and suppresses other languages. If left alone, languages will mix and mesh, transform, and become hardly distinguishable from each other. Under a free linguistic economy, the monolingual paradigm is displaced by the multilingual and translanguaging paradigm in a celebration of open borders, free movement, and diversity.

As Russia invaded Ukraine at the time of revising this chapter, what seemed like a distant past no longer feels so distant. The traumatic history of an unjustified foreign invasion, the memories of the suppression of cultural and linguistic

diversity, and the fear of the return of a highly ideologized, repressive political and linguistic regime have all become particularly salient. Sociolinguists have a significant role to play in witnessing and recording the linguistic expressions of social and political upheavals – and a duty to issue a stark warning about the dangers of history repeating itself and totalitarianism reawakening.

References

- Atanassova-Divitakova, Svetlana. 2017. Talking space: A linguistic portrait of a popular tourist location in Veliko Turnovo. In Emilia Slavova, Bozhil Hristov, Jonathan McCreedy, Rositsa Ishpekova, Alexandra Glavanakova, Nelly Yankimova & Angel Igov (eds.), *New paradigms in English studies. Language, linguistics, literature and culture in higher education*, 35–53. Sofia: St. Kliment Ohridski University Press.
- Blommaert, Jan. 2010. *The sociolinguistics of globalization*. Cambridge: Cambridge University Press.
- Blommaert, Jan & Ico Maly. 2014. *Ethnographic linguistic landscape analysis and social change: A case study* (Tilburg Papers in Culture Studies 100). <https://research.tilburguniversity.edu/en/publications/ethnographic-linguistic-landscape-analysis-and-social-change-a-ca-4>.
- Blommaert, Jan & Ben Rampton. 2011. Language and superdiversities. *Language and Superdiversity* 13(2). 1–22. https://newdiversities.mmg.mpg.de/?page_id=2056.
- Canagarajah, A. Suresh. 2013. *Translingual practice: Global Englishes and cosmopolitan relations*. Abingdon, Oxon & New York: Routledge.
- Danov, Milen & Stoyan Galabov. n.d. *Спомени от народната република (Memories from the People's Republic)*.
- Dnevnik. 2011. *Фотогалерия: Рекламните надписи от времето на социализма в София (Photogallery: Advertising signs from the time of socialism in Sofia)*. https://www.dnevnik.bg/razvlechenie/2011/06/14/1105840_fotogaleriia_reklamnite_nadpisi_ot_vremeto_na/.
- Gorter, Durk. 2013. Linguistic landscapes in a multilingual world. *Annual Review of Applied Linguistics* 33. 190–212.
- HighViewArt. 2014. *По софийските улици през 60-те, 70-те и 80-те (In the streets of Sofia in the 60s, 70s and 80s)*. <https://www.highviewart.com/cvetno-po-sofiyskite-ulitsi-prez-60-te-70-te-i-80-te-2233.html>.
- Iliev, Ivan. 2013. Short history of the Cyrillic alphabet. *International Journal of Russian Studies* 2(2). 221–285.

- Jørgensen, J. N., M. S. Karrebæk, L. M. Madsen & J. S. Møller. 2011. Polylinguaging in superdiversity. *Language and Superdiversity* 13(2). 23–38.
- Krusteva, Asya. 2017. Exploring the linguistic landscape in the Sofia subway. Public signs and language policy. *Diggit Magazine*. <https://www.diggitmagazine.com/papers/exploring-linguistic-landscape-sofia-subway>.
- Lankiewicz, Hadrian & Emilia Wąsikiewicz-Firlej (eds.). 2014. *Languaging experiences: Learning and teaching revisited*. New Castle upon Tyne: Cambridge Scholars Publishing.
- Nikolova, Milana. 2021. It's the Cyrillic alphabet, not the Russian alphabet. *Emerging Europe*. <https://emerging-europe.com/after-hours/its-the-cyrillic-alphabet-not-the-russian-alphabet>.
- Norman, Chelsey. 2019. Linguistic ideologies in the performance of Bulgarian identity. In *Purdue Linguistics, Literature, & Second Language Studies Conference*. <https://docs.lib.purdue.edu/plcc/2019/papers/4/>.
- Politov, Alexander & Denitza Lozanova. 2015. Multilingualism in Sofia. In *Lucide city report*. LSE Academic Publishing. https://www.sofia-da.eu/images/projects/LUCIDE/Reports/14_0058%20Sofia%20LUCIDE%20Report%20V6_HR_ONLINE_1.pdf.
- Saraceni, Mario. 2015. *World Englishes: A critical analysis*. London, New York: Bloomsbury.
- Swain, Merrill. 2006. Languaging, agency and collaboration in advanced second-language proficiency. In Heidi Byrnes (ed.), *Advanced language learning: The contribution of Halliday and Vygotsky*, 95–108. London, New York: Continuum.

Chapter 2

The “unnecessary” use of French in Moroccan Arabic: Social discriminant or collaborative enterprise?

Jacopo Falchetta


University of Bergamo and IREMAM, Aix-en-Provence

The practice of code-switching between Moroccan Arabic and French is well documented among speakers of the former variety. Previous scientific stances on Moroccan Arabic-French code-switching (especially expressed in works almost entirely concerned with formal structure rather than sociopragmatic functions) merely saw bilinguals’ recurring to French as an effort to upgrade their social status by using a “prestigious” language. Based on a sample of a corpus collected by the author, this contribution presents the first steps of a study which, by adopting an interactionalist approach, aims to identify the sociopragmatic functions that speakers fulfil through code-switching. The results show that, in communities where not all speakers are highly proficient, code-switching can also be used to fill the existing gaps between the speakers’ linguistic repertoires. Such gaps are notably due to different levels of exposure to non-native French or being taught in private schools where French is the teaching language. It is also argued that, even in those cases in which code-switching is in fact due to the positive connotations of French, it is more fruitful to go beyond the “prestige versus stigmatization” dichotomy by looking at the indexical meanings associated with the forms employed. Examining the contexts of exposure to, learning, and use of these forms is suggested as a possible method to address such indexicalities.

1 Introduction

The study of code-switching (CS) between French and Moroccan Arabic (MA) has generally been focused on the formal aspects of the alternation between the



Jacopo Falchetta. 2024. The “unnecessary” use of French in Moroccan Arabic: Social discriminant or collaborative enterprise? In Katrin Pfadenhauer, Sofia Rüdiger & Valentina Serreli (eds.), *Global and local perspectives on language contact*, 37–62. Berlin: Language Science Press. DOI: 10.5281/zenodo.10497371 

two varieties, trying to explain *how* they co-exist in a single string of speech (most notably Bentahila & Davies 1983, 1995, Lahlou 1991, Ziamari 2007, 2009, 2018). Conversely, *why* speakers alternate between the two codes (i.e., the sociopragmatic functions of French-MA CS in the interactional context) has been questioned less frequently and, when it has, it has rarely been supported through in-depth interactional analyses of the socio-pragmatic reasons leading speakers to engage in CS. In most of the previous works on French and MA mixing, this practice has been attributed to the ambition to elevate one's own social status, although other studies have seen it as motivated by participation in an urban bilingual culture, or by aesthetic and expressive intent. Nevertheless, it is argued here that there can also be more urgent motives pushing speakers to engage in this practice, as will be shown through an analysis of verbal interactions occurring in the framework of a guided production test. It is also argued that, even in those cases in which the prestigious status of French seems relevant to what is happening, it is useful to attempt to qualify the notion of "prestige" – which has long been employed in the explanation of sociolinguistic phenomena as a sort of black box – by taking other aspects of the problem into consideration, such as contexts of language learning and indexicalities.

This chapter presents the partial results of a work in progress which aims to show what functions are fulfilled through CS by young, male speakers of MA, and how diverging lexical repertoires, possibly due to unequal learning opportunities, are connected to the particular practices observed and to the functions identified. First, however, a general diachronic picture of the contact between French and MA varieties will be given in order to provide a historical background to the present sociolinguistic status of French and MA in Morocco. Following this, a brief review of the previous literature on MA-French CS will be sketched. An operative definition of CS will be formulated before describing the data corpus, and then a provisional classification of sociopragmatic functions will be attempted based on the analysis of the CS occurrences. Finally, some observations will be made on the implications of the results and the steps that further research should take.

2 The status of French and Moroccan Arabic in Morocco

French has been enjoying a privileged status in Morocco since the beginning of the French Protectorate (1912–1956). During this crucial period, it not only functioned as the official language of administration and public affairs, but also as the main teaching language of the school system established by the colonizers. It

also became associated with the transmission of technical and all sorts of specialized knowledge. In this context, the introduction of new French terms connected to technological advancements must have been the main motive that pushed MA speakers to start borrowing massively from French vocabulary (cf. Brunot 1949: 355), although borrowings were and are by no means limited to technical domains.¹

After independence from France (1956), a much greater part of the Moroccan society became exposed to French – in spite of the government’s more or less immediate adoption of Standard Arabic (SA) as the only official language (Moatasime 2006, Sayahi 2014). This is because, while during the Protectorate access to education was limited to the colonizers’ and the members of a few Moroccan élite families (Redouane 1998: 197–198), “the spread of education, the sociodemographic changes (population growth and movement)” that came with the political independence, “and finally the role of mass media” contributed to increasing occasions of contact with French to an unprecedented degree (Sayahi 2014: 42).

The role played by education was significantly curbed between the 1970s and the 1980s, when the teaching of all school subjects was switched from French to Arabic with what are known as the Arabization policies. In the Moroccan context, these did not affect scientific and other university faculties (e.g., Economics, Medicine, etc.), which maintain the exclusive use of French to this day. As a consequence, the side effect of Arabization was a problematic language gap, as most students who attended state school used to find themselves abruptly switching from Arabic to French when starting their BA programmes. Of course, the same did not apply to those students whose families could afford to send them to private primary and secondary school institutes where the teaching language is French for all subjects (Ennaji 2005: 210–212, Pellegrini 2019: 81).²

Today, French competes with SA as the variety employed in formal and learned contexts, but maintains a dominant role in the scientific arena as well as in the private business sector. Ideologically, French is also the language associated with Western acculturation and/or modernity in public discourse and questionnaire-based interviews about individuals’ language attitudes.³

¹Besides, SA or MA words were also used in several cases to translate some of those innovative terms (Brunot 1949: 356).

²This issue was re-addressed in 2014, when the teaching of scientific subjects at secondary-school level was reconverted to French (Pellegrini 2019: 81–82).

³On language attitudes in Morocco towards French in general, see Ennaji (2005: 193–195) and included bibliography; for an overview of Moroccan civil society’s positions vis-à-vis the use of SA, French, and English in education and the media, see Pellegrini (2019: 109–129); on the association of French with sciences and modernity as it emerges from individual interviews, see Tamer (2003, 2006), Ennaji (2005), Boutieri (2012), Chakrani & Huang (2014).

As for MA, the other variety involved in the CS practices analyzed here, it is native to most of the Moroccan population.⁴ It has traditionally been seen as standing in a diglossic (Ferguson 1959) relationship as the L variety against SA (the H variety), although now it is generally admitted that this strictly dichotomous view, which has been revisited by the author himself (Ferguson 1991), is inadequate to represent the more fluid linguistic reality of Morocco and the other Arabic-speaking countries. Since the complex relationship and interplay between SA and Arabic colloquial varieties is not at issue in this work, it will suffice here to say very briefly that “in linguistic practice, vernacular Arabic exists in a *symbiotic* relationship with Standard Arabic”, whereby “at the level of ideology [...] vernacular Arabic exists in a *subordinate* position vis-à-vis Standard Arabic” (Hachimi et al. 2022: 4; my emphasis). However, even though MA still enjoys no official status in Morocco, it is now enlarging its domains of use at the expense of SA and French, especially in formal and learned contexts and youth cultural expressions. Ideologically, *darija* (/da:ri:ʒa/)⁵ is now reported to be positively evaluated, especially by young people, and play a role in processes of identity construction (cf. de Ruiter 2006, de Ruiter et al. 2014, Caubet 2017).

3 Previous research on MA-French code-switching

CS between MA and French in spontaneous verbal exchanges has been studied since at least the 1970s. It has most often been observed among educated speakers, especially university students and graduates (e.g., Abbassi 1977, Bentahila & Davies 1983, 1995, Lahlou 1991, Ennaji 2005, Ziamari 2007, 2009, 2018, Post 2015), as people with a certain degree of bilingualism are assumed to alternate codes more frequently. In line with an interest in the syntactic and morphological constraints of code alternation, the vast majority of these works have mainly focused on the structural features of CS, such as allowed switching points, inter- versus intra-sentential CS, and similar formal issues. Studies adopting this perspective have undoubtedly shed light on important formal dynamics of CS and serve as a basis for the type of research proposed here. However, few of them have addressed the sociopragmatic functions that the use of CS fulfils in a given

⁴According to the 2014 national census, about 92 per cent of Moroccans speak (non-Standard) Arabic as their mother tongue, while 26.1 per cent were classified as Berber speakers (RGPH 2014). While it is possible that a number of Arabic-Berber bilinguals declared that they spoke MA only, the number of Berber speakers is known to be declining, at least in Morocco.

⁵A MA word indicating the variety of Arabic spoken in ordinary daily situations. It is generally used in opposition to *fusha* (/fusˤha:/), a term which brings together both the classical and modern versions of SA.

verbal interaction. In the majority of cases, the main motive mentioned for engaging in MA-French CS is the speaker’s desire to enhance their speech style by demonstrating their knowledge of a socially prestigious and valuable language. The excerpt below exemplifies this.

The technique of studding Arabic discourse with French lexical items is a means whereby [bilinguals educated under the Arabisation policies] manage to *distinguish themselves* from their less educated contemporaries who would not be in a position to switch to French at all; the presence of French vocabulary signals that the speakers, although they are basically speaking Arabic, *do have the requisite knowledge of French to be able to call upon it when they feel like it*. (Bentahila & Davies 1995: 84; my emphasis)

Nonetheless, a few exceptions to this trend are found in later works. Ziamari, for example, makes the association between French-MA CS and urban youth practices. Her community of study is made up of students of Francophone faculties at the University of Meknes, Morocco.

Codeswitching is mainly associated with an *urban environment*. [...] It is this urban environment, where different languages come into closest contact and *where there is the greatest incidence of bilingualism*, which favors the emergence of the practice. Codeswitching in Morocco [...], while it does occur in various social categories, is essentially *a feature of the speech of young bilinguals*. (Ziamari 2007: 276; my emphasis)

More recent studies (e.g., Chakrani 2010, Chakrani & Huang 2014, Khoumssi 2020) have taken a greater interest in speakers’ attitudes towards CS. Rather than analyzing samples of code-switched utterances, the authors made use of questionnaires in which consultants were explicitly asked about their opinion of CS and other language varieties. These works provide more information about contexts of use than social functions and meanings of MA-French CS, observing that it is usually employed in informal communication in the classroom or peer groups. Conversely, Post (2015) joins the study of verbal productions and language attitudes through a systematic analysis of the relations between extralinguistic factors and CS structure and frequency of use; however, motivations for CS at the interactional level are not touched upon. To my knowledge, Ziamari’s (2009) is the only study that addresses CS in spontaneous communicative contexts from an entirely pragmatic perspective by finding correspondences between use of French and the information structure of MA utterances.⁶ Other

⁶More precisely, Ziamari finds that French constituents embedded in MA speech are frequent in focus and topic positions.

studies focusing on Maghrebi artistic texts have been more sensitive to the issue of the socio-pragmatic potential of French-Maghrebi Arabic CS, e.g. in the expression of humor (cf. Caubet 1998 for Algerian Arabic) or in the aesthetic search for expressivity in song lyrics (cf. Bentahila & Davies 2002, Davies & Bentahila 2006, 2008 for MA, and Caubet 2002 for Algerian Arabic). This chapter will therefore explore the possibility of enlarging knowledge of the sociopragmatic functions of MA-French CS in spontaneous exchanges by analyzing verbal interactions involving members of a group of young male Moroccan peers, in order to reveal their purposes for engaging in code alternation.

4 The data

The corpus that will be exploited for this analysis originally served as a basis for the study of three linguistic variables in the language use of MA speakers in the town of Temara, Morocco.⁷ The interest in studying linguistic variation in Temara, a former rural suburb of the capital, Rabat, lies in its peculiar sociodemographic situation. Since the years of the French protectorate, Temara has been attracting huge numbers of migrants from other areas of the country, especially after Rabat's residential areas became saturated, and this has led to a steep demographic growth of the town's population, particularly from the 1970s onwards (from circa 20,000 in 1972 to over 300,000 in 2014). Linguistically, the convergence of immigrants from different regions has led to the encounter of different regional varieties of MA, which has given way to phenomena of dialect contact such as levelling and reallocation of diatopically distinctive features.

What is of interest for this study is the social background of the interviewees, all of whom are children of rural migrant parents and were living in Temara at the time the data were collected. While the parents had, in most cases, received little or no education, they had managed to raise their children in more than decent material conditions and fund their education in state or (less frequently) private schools. As a result, most of the interviewees had been educated to at least the final year of secondary school, although a minority of them had not gone past middle school. In addition, given the proximity of Rabat, routes between the latter and Temara are well-served, and Temari youth are used to commuting to the capital city frequently to attend university, work in skilled or specialized positions, or engage in leisure activities that are not available in the town where they live (such as clubbing or skating).

⁷For more details on the variables and results of the analysis, see Falchetta (2019).

The data are taken from a corpus of speech samples obtained by means of an experiment which I designed specifically to elicit one of the variables and test the speakers’ tendency to employ the two variants involved (more details in Falchetta 2019). This experiment, to which I shall henceforth refer as “the test”, required (at least) two participants⁸ to engage in narratives and verbal exchanges for the fulfilment of the following tasks: one of them was asked to recount a speechless video (a hidden-camera prank) to the other, who had not watched it; the latter was subsequently asked some questions, allegedly with the aim of ensuring the first participant had been clear and exhaustive in summarizing the prank.⁹ The choice of having the consultants summarize practical jokes that were shown in routines was motivated by the exigency to minimize the observer’s paradox (Labov 1973: 208–209): this verbal task closely resembled a type of activity in which MA speakers had been seen engaging spontaneously (telling a peer about a funny hidden-camera prank), and the routine nature of the video made it easier for the recounter to memorize and verbalize the amusing situation. If there is obviously no way that one can be assured the consultants would have used the same wording and register in the researcher’s absence, the test was still effective in triggering loose exchanges between the participants, and in eliciting the desired variable.

The participants were 17 pairs of male residents of Temara; three videos were verbalized by each of the participants, meaning each of the 17 sessions included summaries of and answers to questions on six videos. Each session lasted between 30 and 60 minutes, including the time required for the participants to watch the respective videos.

The test was conducted between mid-July and mid-August 2017. Two criteria were used for the recruitment of the participants: they had to be prior acquaintances and used to speaking MA between one another.¹⁰ Their ages ranged between 18 and 38, with an average of 25.6. The data have not been transcribed; excerpts from the sessions were written down for the phonological analysis using IPA characters, and the same will be done here for the analytical purposes of this chapter.

The key advantage of this type of test is that it makes it possible to collect samples of speech from an unlimited number of speakers while keeping the communicative context and purposes constant, to the benefit of the comparability of

⁸All the test sessions analyzed in this chapter involved two participants each.

⁹In fact, the real purpose of the questions to the second participant was to elicit words containing the targeted variable.

¹⁰The latter criterion was relevant for speakers whose first native language was a variety of Berber.

language use. While the test was administered with the purpose of collecting data on variation *within* MA, the recorded exchanges between the participants in the test sessions often involved the use of French forms in mainly MA utterances, which thus makes them useful for an analysis of MA-French CS in interactions between MA native speakers. Since the prank videos at issue took place in Canada, they were checked one by one and it was made sure none of them contained any English or French signage that could have triggered CS in the speech of the viewing participants.¹¹

One of the aims of this preliminary analysis of French-MA CS in my data was to assess whether the two factors of French proficiency and connectedness to an urban environment were indeed relevant to the frequency and extent of this practice in an individual's speech, as has been contended in previous literature.¹² For this reason, 16 out of 34 consultants were selected for further analysis, so that the sample would cover different degrees of exposure to French and different types of personal social networks (urbanized versus non-urbanized). The 8 sessions thus selected yielded approximately 166 minutes of native MA speech, or an average of 20 minutes per session. The distribution of the participants' sample according to each of these two variables is reported in Table 1 and Table 2 respectively. As shown in Table 1, most interviewees did not receive a Francophone education for most of their schooling years, and presumably few of them used French on a daily basis at work. As for attending a university faculty with French as the teaching language, this did not seem to improve their ease in speaking the language; on the contrary, during interviews conducted with some participants before the test, many of those who attended university courses in French reported not being competent enough in the language to fully understand classes. Therefore, this study presents an important difference with respect to previous works (Abbassi 1977, Bentahila & Davies 1983, 1995, Lahlou 1991, Ziamari 2007, 2009, 2018): while these were explicitly focused on analyzing CS among a population of fluent French speakers, my young Temari consultants' proficiency was not a matter of concern (as the test originally had purposes unrelated to CS analysis) and ranged from low to high according to the participant.¹³ Nevertheless, even those less competent in French did engage in CS during their test performance, as will be shown in the next section.

¹¹The only potential trigger, the writing *Location* (French for 'renting') on a lorry, was not exploited by any participant, as no one used this term.

¹²However, for this study, the discussion will be limited to the recurring sociopragmatic functions that CS appeared to fulfil across the interviewees' sample, regardless of the influence of these social factors.

¹³However, it should be noted that all works, including mine, base their considerations on the participants' proficiency either on assumptions or on self-evaluations.

Table 1: Distribution of the participants according to sources of exposure to French.

Number of participants	Higher education in French	Pre-higher education in French	French as a working language
3			
7	✓		
1		✓	
3	✓		✓
2	✓	✓	✓
Total: 16			

Table 2: Distribution of the participants according to type of social network.

Number of participants	Urbanized social networks ¹⁴	Non-urbanized social networks
9		✓
7	✓	
Total: 16		

5 Formal features of MA-French CS

According to Manfredi et al. (2015: 285), “a distinction” between CS and the closely related phenomenon of borrowing “is necessary and possible”. They define CS as “the presence of lexical or sentential material belonging to different linguistic systems, provided that its *different origin is still transparent in the speaker’s output* in one or more grammatical domains” (Manfredi et al. 2015: 286; my emphasis). Therefore, in order to distinguish CS from borrowing, it is crucial to define the criteria by which the linguist can state whether “the material in the speaker’s output” enjoys “native or foreign status” (Manfredi et al. 2015: 286). Keeping this in mind, in this preliminary analysis an item was classified as ‘code-switched’ if at least one semantically equivalent lexical item was employed in the recipient language (i.e., MA) with equal or greater frequency. In this sense, it

¹⁴By this, I mean social networks that include work- or leisure-related contacts with speakers from larger urban centres, such as Rabat or Casablanca.

is assumed that a given item enjoys ‘foreign status’ insofar as it is seemingly ‘unnecessary’ to the speakers of the recipient language, as they do have synonyms of that item in their native repertoire. Manfredi et al. (2015: 300, 306) also occasionally classify an item of “foreign” origin as CS (rather than borrowing) by using a similar criterion, although the main point of their study is that prosodic and intonational regularities can be important cues for the identification of CS.¹⁵ Using the criterion of the absence of an equally or more widespread synonym, 157 instances of French-derived forms were classified as CS in the sample analyzed, or an average slightly short of ten instances per participant, with individual figures ranging from 2 to 32.¹⁶

From a formal point of view, all speech in the eight sessions analyzed was mainly in MA, with limited constituents being switched to French. According to the Matrix Language Framework (Myers-Scotton 2002), it can then be stated that MA almost invariably constituted the Matrix Language and French the Embedded Language. Concerning the extension of the Embedded Constituents (ECs), they hardly went beyond noun-phrase level, with just one interviewee producing two entire clauses in French. This matches Bentahila & Davies’ (1995) data, which revealed that bilinguals with a post-Arabization education (like my participants) mainly engaged in a similar type of CS – with ECs mostly limited to noun phrases – more than 20 years before my data were collected. By contrast, bilinguals who received a Francophone (pre-Arabization) education engaged in a radically different type of CS in Bentahila and Davies’s data, with a greater use of clause- and sentence-wide ECs and with inter-sentential switches more frequent than intra-sentential switches. Below are some examples of CS from the Temara corpus, with different types of ECs involved, including

- morpheme-level ECs;
- noun-phrase-level ECs; and
- prepositional- and adverbial-phrase-level ECs.

In some cases, CS occurs below word level, with a French lexical morpheme being combined with MA inflectional morphemes, as in Examples (1) and (2).

¹⁵Further developments of the analysis presented here will also aim to check whether such regularities match the instances of CS identified in the Temara corpus.

¹⁶Occurrences of CS in speech directed at the researcher (e.g., during the answering task) were omitted, to avoid switches to French due to accommodation to a non-native speaker. However, cases in which the same term used in speech directed at the researcher was also employed with the other participant, or was adopted after the other participant had employed it for the same referent, were retained in the count.

- (1) MA (Temara corpus; GS-7)¹⁷
 hi:ja rˤa ka:-təbqa mfi:ksja
 3FSG ARG PRVB-stay;IPFV.3FSG fix;PTCP.F
 “It remains **fixed**” (< FR *fixer* ‘to fix’)
- (2) MA (Temara corpus; GG-11)
 wa:həd l-ma:la:bi:s lli ka:-tʔatiri bəzza:f d n-na:s
 INDF-clothes REL PRVB-attract;IPFV.3FSG many GEN DEF-people
 “Clothes that **attract** lots of people” (< FR *attirer* ‘to attract’)

When at least one lexeme is switched at noun-phrase level, CS can be limited to the noun alone (as in Example 3) or involve associated determiners (Example 4) and noun complements (Example 5).

- (3) MA (Temara corpus; DS-9)
 za:jdi:n li:-ha hu:ma f l-volym¹⁸
 increase;PTCP.ACT.PL to-3FSG 3PL in DEF-volume
 “They turned up its **volume**” (< FR *volume* ‘volume’)
- (4) MA (Temara corpus; IDG-10)
 fa:ʃ jəmfi d-dərri tə-jzi:w de turist
 CONJ go;IPFV.3MSG DEF-child.M PRVB-come;IPFV.3PL INDF.PL tourist
 “When the little kid leaves, **some tourists** come” (< FR *des touristes* ‘tourists’)
- (5) MA (Temara corpus; EMK-26)
 mu:hi:mm dək la sal d-atāt
 important DEM DEF room PREP-waiting
 “Anyway, it’s that type of **waiting room**” (< FR *salle d’attente* ‘waiting room’)

¹⁷Every interviewee is represented by a two- or three-letter code, which is always indicated next to the source. The number following the code indicates the total number of ECs calculated for the participant. Switched items on which the example is focused are in bold.

¹⁸Concerning the treatment of definite articles preposed to French ECs, it was observed that the /l-/ preceding masculine singular nouns starting with a consonant was often phonetically realized as [əl] and therefore sensibly different from French <le> ([lə]). For this reason, it is to be interpreted as the MA article /l-/ ([lə]); accordingly, it gets assimilated to a following sun (coronal) letter (cf. Example 12). Conversely, /la/ and /le/ are to be interpreted as the French feminine singular and plural articles respectively, and therefore as part of ECs. Boumans & Caubet (2000: 152) make similar observations in the context of Algerian Arabic-French CS. The issue of the confusion between the MA definite article and French <l’>, i.e., the form preceding singular nouns beginning with a vowel, will not be dealt with here as none of the examples happens to present such a case.

Several cases of switched prepositional (see Example 6) and adverbial phrases (Example 7) or conjunctions (Example 8) are also found, albeit more rarely.

- (6) MA (Temara corpus; DN-15)
 ā fas mʃa:-hum
PREP face with-3PL
 “Facing them” (< FR *en face* ‘opposite’)¹⁹
- (7) MA (Temara corpus; GS-7)
 otomatikmā hu:wa tə-jdxəl l-ha:di:k lli təht
automatically 3MSG PRVB-enter;3MSG to-DEM REL below
 “Automatically, he falls into that thing below” (< FR *automatiquement* ‘automatically’)
- (8) MA (Temara corpus; LT-32)
 a fak fwa ʃi wa:həd kə-jʒi
PREP each time INDF one PRVB-come;3MSG
 “Every time somebody comes ...” (< FR *à chaque fois* ‘every time’)

As regards the phonetic integration of French-derived forms, this varies from full integration into MA phonetics to full adaptation to the prescribed French pronunciation. Generally speaking, more proficient speakers tend towards the latter, although a certain degree of adaptation to local phonotactics is found in most cases. For instance, Table 3 shows different pronunciations of the vowels in the French *feu rouge* [fø ruʒ] ‘traffic light’, as produced by three participants. These are ranked by increasing (assumed) exposure to French in descending order. Based on this example, it would actually seem that greater exposure does lead to pronunciations closer to the standard. Conversely, the second column shows that the same factor did not necessarily entail greater engagement in CS during the test. In any case, analyses of a broader sample would be necessary to state to what extent proficiency is influential in the degree of integration of switched items.²⁰

While the number of ECs in the corpus is not quantitatively significant compared to other studies, certain sociopragmatic functions fulfilled through CS appeared to recur across individual uses. As I shall demonstrate in the next section, the

¹⁹In this case, the French adverbial locution is used in combination with the MA preposition /mʃa/ ‘with’ to form the MA prepositional locution /ā fas mʃa/ ‘opposite’.

²⁰Post’s (2015: 159–163) analysis of the correlation between CS practices and proficiency in French shows trends in the rate of different types of French constituents switched, but does not provide absolute numbers of switched constituents.

Table 3: Different phonetic integrations of *feu rouge* ‘traffic light’.

Interviewee	ECs	Description	Realization
DC	5	no schooling in French up to (and including) secondary level, unemployed	[fi:rʊʒ]
DN	15	no schooling in French up to (and including) secondary level, uses French as a working language, frequently goes to Rabat for work and leisure	[fø:rʊʒ]
LPI	5	all schooling in French, uses French as a working language, works in Rabat	[fø:ru:ʒ]

interest in pinpointing these functions lies in the new light they shed on MA speakers’ CS practices with respect to the data reported in previous literature.

6 Sociopragmatic functions of CS

6.1 The purpose of classifying functions

In order to show the various pragmatic functions fulfilled through the use of CS, an interactional analysis was carried out on each verbal exchange in which speakers made use of ECs. Following this, a tentative classification of the pragmatic functions was made based on those that recurred the most across the corpus. This endeavour freely follows the model provided and the analysis carried out by Gumperz with his list of conversational functions of CS, which he bases on parallel examples taken from three different pairs of alternated languages (Gumperz 1982: 75–84). However, it should be noted that the aim of the classification made in the present work is not to elaborate a comprehensive taxonomy, but rather to show how MA-French CS can also serve as an instrument to achieve greater inter-comprehension and/or convey information in a more effective way. This is already suggested by the context of these exchanges, which is the fulfilment of a verbal task that depends on the correct transmission of information from the speaker to the listener.

6.2 Information-conveying CS

Two types of communicative strategies have been identified in which CS is a means for achieving mutual understanding. Both of them can be assimilated to

Gumperz's "reiteration" strategy, which is when "a message in one code is repeated in the other code, either literally or in somewhat modified form. In some cases such repetitions may serve to clarify what is said, but often they simply amplify or emphasize a message" (Gumperz 1982: 78). It is argued that, in the exchanges analyzed here, clarification is what is aimed at.

The first strategy, or sociopragmatic function of CS, can be defined as one of 'clarification through translation'. This is obtained by juxtaposing a French-derived lexeme with its (Standard or Moroccan) Arabic translation, or vice versa. Even though the purpose is apparently that of making sure that the listener identifies the intended referent correctly, the speaker does not seem to wait for the interlocutor to request more clarification, but rather assumes that the latter may or may not understand if only the Arabic or the French form were employed. Examples (9) and (10) are taken from two different sessions.

- (9) MA (Temara corpus; GP-10)
 ha:da wa:həd ə... l-mazisiē sa:hi:r
 DEM INDF [hesitation] DEF-magician magician
 "This is a...magician, a magician" (< FR *magicien* 'magician')
- (10) MA (Temara corpus; DN-15)
 ka:-thəjjəd d-dʒa:ki:tʰa ka:-thəjjəd la vɛst
 PRVB-take.off;3FSG DEF-jacket PRVB-take.off;3FSG DET jacket
 "She takes off her jacket, she takes off her jacket" (< FR *la veste* 'the jacket')

As the pranks took place in Canada, it may be observed that in these two cases, which resemble many others in the corpus, the bilingual denomination is used for referents that are potentially associated with a foreign culture, namely a particular type of entertainer and an item of female clothing. Nonetheless, an Arabic lexeme does exist for each of these entities and is employed by the interviewees themselves. Therefore, while the foreignness of the signified may have an influence on the hesitation between the two languages, the choice of employing two signifiers, each one drawn from a different language, has pragmatic implications that go beyond language-culture association, as will be discussed below.

The second sociopragmatic function can be termed 'validation through translation'. In this order of cases, one participant employs French to translate or reformulate a string of speech that the other has just expressed in Arabic. Here, the foreign language appears to help the code-switcher make sure he understood his interlocutor's statement. Examples (11) and (12) are taken from two different sessions to illustrate this kind of use.

(11) MA (Temara corpus)

a. POP-1²¹

da:ba wa:həd smi:t-u wa:həd l-mətˁŷəm bħa:l l-mətˁŷəm
 now one name-3MSG INDF-restaurant like DEF-restaurant
 gəlti bħa:l
 say;PFV.2SG like
 “So it’s a... how’s it called... a restaurant, like a restaurant... you’d
 say... like...”

b. PT-7

ʔa:h rəstorã
 yes restaurant
 “Yeah, a **restaurant**!” (< FR *restaurant* ‘restaurant’)

(12) MA (Temara corpus)

a. GP-10

ha:da wa:həd ka:mi:ju dja:l l-həbs fi:-h zu:3
 DEM INDF lorry GEN DEF-prison in-3MSG two
 “This is a prison lorry, with two...”

b. LT-32

dja:l
 GEN
 “A what...?”

c. GP

dja:l l-həbs lli tə-jku:n fi:-h l-msəʒʒən hi:ja
 GEN DEF-prison REL PRVB-be.3MSG in-3MSG DEF-imprisoned 3FSG
 b-a:f tə-jhəwwlu:-h mən
 PREP-REL PRVB-transfer;IPFV.3PL-3MSG PREP
 “A prison [lorry], the one that has inmates inside, that is, to move
 them from...”

d. LT

ʔa:h dja:l t-trāspər dja:l le pɹizonje
 yes GEN DEF-transport GEN DEF.PL prisoner
 “Oh yeah! For the **transport** of **prisoners**!” (< FR *transport* ‘transport’
 and *les prisonniers* ‘the prisoners’)

- e. GP
hu:wa ha:da:k
3MSG DEM
“Exactly!”
- f. LT
vwala
INTRJ
“There you go!”

Interestingly, this use seems to imply that French can be more effective than Arabic in clarifying certain kinds of referents for some MA speakers in Morocco; this is particularly clear in Example (12), where LT does not seem to grasp GP’s tentative description of a prison lorry until he asks and obtains validation for his mixed French-MA reformulation of the description itself. The fact that the two examples both refer to concepts that were either introduced or renewed by the French colonizers during the protectorate may explain why their association with the European language appears to be strong for some participants.

As seen in the examples above, in which the main communicative goal was conveying information in a clear and effective manner, CS can fulfil sociopragmatic functions that go beyond the signalling of one’s linguistic skills – at least, one would argue, among speakers with low proficiency in, and/or lower exposure to, French. In parallel, what these uses of CS reveal is that individuals with different social, educational, and attitudinal profiles differ in their abilities to produce not only French, but also MA forms for specific items. The last excerpt reported is particularly eloquent in this respect. Further proof of this comes from the variety of lexemes used by the test participants for a single referent. Since the same hidden-camera videos were shown in different sessions, it is possible to see how participants diverged in the forms they used to denominate the same object. This is especially observed in lexical choices concerning objects crucial to the prank described. More often than not, the sets of words thus obtained contain both French and MA or SA items. Two examples are reported in Table 4 and Table 5.

The choice of whether to designate the same entity with an MA word (which could also be a relatively ancient borrowing, such as /bl̥a:ka/ in Table 4, which is presumably an old adaptation either from the Spanish *placa* or the French *plaque*,

²¹Letters within the example indicate different speech turns in a single exchange. Each turn is introduced by the interviewee’s code and his rate of ECs (the latter is only reported on the first turn).

2 The “unnecessary” use of French in Moroccan Arabic

Table 4: Different denominations for a noticeboard.

	<i>Noticeboard</i>	<i>Language variety</i>
LT-32	/banʒer/ (< bannière ‘banner’)	French
IDG-10	/wərqa/, /bl̥ˈa:ka/	MA, MA
FG-9	/pano/ (< panneau ‘board’)	French
GS-7	/bl̥ˈa:ka/	MA
OL-2	/bl̥ˈa:ka/	MA

Table 5: Different denominations for a doctor’s waiting room.

	<i>Waiting room</i>	<i>Language variety</i>
EMK-26	/sal datāt/ (< salle d’attente ‘waiting room’)	French
DN-15	/qa:ʔat ʔintid̥ˈa:r/	SA
DS-9	/sal datāt/	French
FG-9	/qa:ʔa d-ʔi:nti:d̥ˈa:r/	SA/MA
GM-5	/sal datāt/	French
EY-4	/s̥ˈa:la dʒa:l l-ʔi:nti:d̥ˈa:r/	SA/MA

both meaning “plaque”) or through CS could be due to the expectation of which form the interlocutor would understand most quickly, or simply to the chosen lexeme being the most familiar and therefore the most immediate one for the speaker. Whatever the reason may be, it is clear that different people are used to sourcing words from different varieties (French, MA, or SA) to denominate the same referent; additionally, as seen above, they are often aware that other individuals may not easily understand the lexeme with which they are most familiar. It is argued here that such divergence in the speakers’ lexical repertoires needs to be accounted for on the basis of the educational and social background of their language development.

The uses of CS illustrated so far appear to be motivated by a need to recur to French material for the sake of mutual understanding. For this reason, they cannot be motivated by the social status that French confers on an individual’s speech. That said, other instances of CS have been observed in the corpus in which the connotations of French must have played a role in the speaker’s linguistic choices. Nevertheless, what is advocated in the following paragraphs is the need to go beyond mere notions of ‘prestige’ and ‘social status’ when discussing the reasons that underlie the use of French for stylistic purposes.

6.3 Stylistic adjustment through CS

When CS is not used to clarify referential meaning, it serves the purpose of exploiting the indexical field²² associated with the French form in order to produce certain communicative effects. The sociopragmatic functions that are fulfilled in this group of cases will be grouped under the umbrella term ‘stylistic adjustment’ for the time being. Previous literature mainly restricted its view to these types of uses, attributing them to a generic indexation of prestige or young urban culture associated with the use of French, or French-MA CS. What is argued here is that a more in-depth look at the meaning of these code-mixing practices allows us to take into account how the indexicalities of the French forms employed originate from their juxtaposition with entities (people, objects, attitudes, settings, etc.) which are involved in the speakers’ familiarization with, exposure to, and/or use of such forms. A reading through the concept of indexicalities may also clarify what leads speakers to consider French forms as more prestigious, or what social values make them more appealing to young urban bilinguals, thus elucidating previous findings. The limited amount of data analyzed here does not make it possible to generalize as to the observations made in the remainder of this section; however, the few examples taken from the corpus provide a glimpse of the variety of indexical meanings that can underlie French-MA CS in the community studied and in the Moroccan context in general.

- (13) MA (Temara corpus; GP-10)

tə-jtkaffa mu:r^ˈa:-h
PRVB-hide.3MSG behind-3MSG

“He hides behind it” (< FR *cacher* ‘to hide’)²³

- (14) MA (Temara corpus; EMK-26)

ã fẽ d-kõt ka:-jʒi wa:həd-r^ˈ-r^ˈa:ʒəl
PREP end of-count PRVB-come.3MSG INDF-man

“In the end, a man comes ...” (< FR *en fin de compte* ‘ultimately’)

²²Peirce defines an index as a sign that “signifies its object solely by virtue of being really connected with it”, and not by virtue of laws (like linguistic signs) and visual resemblance (like icons; Peirce 1933: 361). I use the concepts of “indexicality” and “indexical field”, based on Peirce’s (1933) “index”, following their development by Silverstein (1976, 2003) and Eckert (2008).

²³While the original French verb only bears a transitive meaning, the fused MA-French verb used by GP is rendered intransitive by the MA infix /-t-/, which has a passivizing or reflexivizing grammatical function.

Neither of the switches occurring in Examples (13) and (14) stems from a problem of mutual understanding, as they both occur during the summary of one of the pranks that the speaker has viewed (i.e., neither GP nor EMK was seeking clarification from the other participant). A prestige-based interpretation would then claim that both consultants were *elevating* the style of speech. However, if the intended meaning of ‘elevating’ is ‘making appropriate to the communicative context’, then it seems equivalent to say that GP and EMK code-switched in order to *adjust* their style (hence the definition of this pragmatic function).

A reading from the perspective of indexicalities leads to the search for associations which the speaker makes with the use of the French forms involved in the switch. It was suggested that this search could be conducted by looking at the context of learning or habitual use of such forms, or exposure to them. This is admittedly complicated, as each single form might demand a dedicated enquiry. A possible starting point is the formal aspect of the switched constituent; while /ã fẽ dkõt/ (Example 14) is an adverbial locution pronounced according to standard rules, the switched item in /tə-jtkaffa/ (Example 13) is a single lexical morpheme that has been adapted to local morphology and phonotactics.²⁴ This suggests that EMK may have taken its locution from native or prescriptive (e.g., classroom) French speech, while GP may have acquired its mixed-form verb from interactions with Moroccan peers (or it may be an original form he created by following similar models of mixed-form verbs). However, the use of a switched form that is not fully adapted to MA phonetics does not necessarily stem from cultivated contexts of acquisition (see the switch in Example 15).

- (15) **direktəmā** hi:ja ka:-ddu:r[°] l l-kbi:r
 directly 3FSG PRVB-turn;3FSG to DEF-big
 “She **immediately** turns towards the grown-up man” (< FR *directement* ‘directly’)

Unlike EMK’s adverbial locution in Example (14), the French-derived adverb employed by DC in Example (15) is not located at the head of an utterance in standard French. However, its MA counterpart /ni:ja:n/ does admit this syntactic collocation. This suggests that DC’s use of /*direktəmā*/ derives from a direct translation of the MA adverb into French, and that, considering its collocation, this use may not have been picked up from proficient speakers of French — a possible indication that this, too, is a kind of use acquired from peers or youth speech.

Besides the context of acquisition or use, another important source of information is speech by the same interviewee outside the interview — which is unfortu-

²⁴See the gemination of /f/, which is non-geminated in the standard French verb [kafe].

nately only available in a few cases. The consultant whose speech is reported in Example (2) is one such case. During a previous field study in the same town, GG had also taken part in a group interview. In the course of both the interview and the test, he was often observed to adjust his own speech with respect to his usual, off-recorder way of speaking. This adjustment consisted of a more frequent engagement in both MA-French and MA-SA CS. By doing so, he was adopting a style which he obviously thought more appropriate for the research context. The overall view of GG's linguistic choices thus sheds more light on his use of /ka:tʔatiri/ in the utterance quoted in Example (2); this occurred in his description of the sexually appealing attire of an actress, whose way of dressing was functional to the prank he was summarizing. The use of an 'adjusted' French form at this point of the narration may be designed to diminish the potential vulgarity of the situation described. The same interviewee also resorted to French in another, similarly problematic context.

(16) MA (Temara corpus; GG-11)

mankē da:jri:n li:-h l-qa:d'i:b dja:l-u l-penis
dummy put;PTCP.ACT.PL to-3MSG DEF-penis GEN-3MSG DEF-penis
dja:l-u
GEN-3MSG

"A *dummy* to which they added a penis, a **penis**" (< FR *pénis* 'penis')

In Example (16), too, while reporting on a different prank, GG chose to code-switch to French (after using the SA word for the same referent) as he needed to refer to the male organ. The association between French and school, a context in which politeness is presumably given particular value, may underlie the preference for using this language to express sensitive content in the two examples taken from GG's test session; in (16), another language linked to schooling (SA) is even resorted to in addition to French. Therefore, the general tone adopted by GG throughout his contributions to the author's research helps frame his pragmatic uses of CS in Examples (2) and (16) as attempts to adjust to the communicative situation by building on the indexical association between French, its learning environment, and the social attitudes promoted within such an environment. Ultimately, it is this association that makes French a safe haven when one has to express contextually sensitive information.²⁵

²⁵Caubet (2002: 236) also identifies the use of a French word in CS with Algerian Arabic as sounding "neutral and more scientific or technical". CS used to express taboo words has also been found in MA-French CS in song lyrics (Bentahila & Davies 2002: 200–202) and even, occasionally, in MA-SA CS in the MA dubbing of soap-operas (Ziamari & Barontini 2013: 233–236).

7 Discussion and provisional conclusions

The uses of MA-French CS illustrated in the first part of the analysis add a new perspective on this practice, which is not only used to signal one’s linguistic skills, or as a group marker, but can also be, and often is, a means to improve the flow of information or to negotiate the meaning of what is being communicated. This aspect is not likely to be adequately appreciated if the focus of the analysis is limited to the structural features of the switched constituents. More importantly, it has been shown that cases in which this practice facilitates the conveyance of information reveal that it can also be a collaborative enterprise, whereby supposed gaps in the interlocutors’ linguistic repertoires are filled by referring to the same entity or concept in two different language varieties at the same time. This communicative strategy also has the effect of (partially) levelling inequalities in the speakers’ exposure to French, since those who are less familiar with the language have the chance to learn new words by hearing them associated with their translation in Arabic.

Nevertheless, the fact that some speakers are more at ease when a certain entity or concept is referred to in French, even though the MA form is more widespread in the community (i.e., when the French form is, apparently, ‘unnecessary’), also reveals that they have contrasting repertoires. It is easy to connect this linguistic inequality to the similarly unequal access to adequate French learning in Moroccan society, and different degrees of exposure to French inevitably entail different levels of proficiency in or familiarity with this language. In this sense, CS is also undoubtedly a social discriminant, inasmuch as it depends on the ability that the speaker has to engage in it, with obvious consequences for the presence and frequency of ECs in their speech. However, previous studies (especially Post 2015) have also pointed out how personal attitudes towards the linguistic varieties at stake may heavily affect the individual’s inclination to mix their native variety with foreign forms; in the Moroccan context, this is all the more relevant, since French is still seen today as a language through which values alien to the local society are being imported. This means that some speakers may choose not to make use of this distinctive linguistic skill even if they are able to. Another problematic point is the extent to which habitual MA-French code-switchers choose to accommodate to non-CS speakers. To clarify this, CS frequencies will have to be compared in the same interviewee’s speech in contexts other than the hidden-camera test (i.e., group interviews and spontaneous conversations), for those speakers for whom these data are available.

Concerning the cases in which CS serves the purpose of adjusting speech style, these suggest that practices which could all be included under the term ‘prestige’

can actually be based on different indexical meanings attributed to the same language. Even though additional data are needed to define more precisely what kind of associations are made by the speakers in reference to these forms, the mere existence of different contexts of learning implies that indexicalities are potentially quite divergent; while a French form learnt in the classroom may index politeness and cultural elitism, another French form learnt from a peer group may be associated with 'youth' or 'street' language, 'thug' registers and so on. Of course, hypotheses on contexts of use/exposition need more data to be supported and/or nuanced. Extending the sample of participants (including to women), making it more equally representative of speakers who have been schooled in Arabic or French, or asking the participants themselves for their feedback on the matter are all steps that may help in this endeavour.

The analysis of this small sample of interviewees has shed light on several elements that further research on CS (involving this and other pairs of varieties) should take into account, especially in contexts in which bilingualism is due to an ex-colonial, non-native language constituting a necessary tool for seeking employment and/or social upgrading. In these contexts, collaborative CS may be an option for speakers, and can be identified by analyzing what happens in interactions between speakers who usually code-switch with different frequencies, or have different proficiency levels in their non-native language. Besides, carrying out proficiency tests while at the same time collecting data on language use may help determine whether contact between proficient CS and non-proficient speakers contributes to the spread of linguistic knowledge. As for non-collaborative, stylistic CS, an analysis that rests on the concept of indexicalities makes it possible to move beyond the dichotomous interpretations of prestigious versus stigmatized speech, by revealing the multiple meanings that the embedding of one language into another may bear for bilinguals. All these efforts require extending the analysis to contexts of acquisition as well as use and exposure to the foreign forms, in order to get a fuller view of the sociolinguistic life of the speaker, and clarify why 'unnecessary' use of foreign forms under a purely linguistic perspective is actually necessary according to the individuals' social and pragmatic point of view.

Abbreviations

1, 2, 3	1 st , 2 nd , 3 rd person	INTRJ	interjection
ACT	active voice	IPFV	imperfective
ARG	argument-introducing particle	M	masculine
		MA	Moroccan Arabic
CONJ	conjunction	PFV	perfective
CS	code-switching	PL	plural
DEF	definite	PREP	preposition
DEM	demonstrative	PRVB	preverb
EC	Embedded Constituent	PTPC	participle
F	feminine	REL	relative particle
FR	French	SG	singular
GEN	genitive	SA	Standard Arabic
INDF	indefinite		

References

- Abbassi, Abdelaziz. 1977. *A sociolinguistic analysis of multilingualism in Morocco*. Austin: The University Texas at Austin. (Doctoral dissertation).
- Bentahila, Abdelâli & Eirlys E. Davies. 1983. The syntax of Arabic-French code-switching. *Lingua* 59(4). 301–330.
- Bentahila, Abdelâli & Eirlys E. Davies. 1995. Patterns of code-switching and patterns of language contact. *Lingua* 96(2-3). 75–93.
- Bentahila, Abdelâli & Eirlys E. Davies. 2002. Language mixing in rai music: Localisation or globalisation? *Language & Communication* 22(2). 187–207.
- Boumans, Louis & Dominique Caubet. 2000. Modelling intrasentential codeswitching: A comparative study of Algerian/French in Algeria and Moroccan/Dutch in the Netherlands. In Jonathan Owens (ed.), *Arabic as a minority language*, 113–180. Berlin: De Gruyter Mouton.
- Boutieri, Charis. 2012. In two speeds (*A deux vitesses*): Linguistic pluralism and educational anxiety in contemporary Morocco. *International Journal of Middle East Studies* 44(3). 443–464.
- Brunot, Louis. 1949. Emprunts dialectaux arabes à la langue française dans les cités marocaines depuis 1912. *Hespéris* Tome XXXVI. 347–430.
- Caubet, Dominique. 1998. Alternance de codes au Maghreb: Pourquoi le français est-il arabisé ? In *Plurilinguismes n°14, Alternance des langues et apprentissage en contextes plurilingues*, 121–142. Paris: CERPL, Université René Descartes.

- Caubet, Dominique. 2002. Jeux de langues: Humor and codeswitching in the Maghreb. In Aleya Rouchdy (ed.), *Language contact and language conflict in Arabic*, 233–255. London: Routledge.
- Caubet, Dominique. 2017. Darija and the construction of “Moroccanness”. In Reem Bassiouney (ed.), *Identity and dialect performance: A study of communities and dialects*, 99–124. London: Routledge.
- Chakrani, Brahim. 2010. *A sociolinguistic investigation of language attitudes among youth in Morocco*. Urbana: University of Illinois at Urbana-Champaign. (Doctoral dissertation).
- Chakrani, Brahim & Jason L. Huang. 2014. The work of ideology: Examining class, language use, and attitudes among Moroccan university students. *International Journal of Bilingual Education and Bilingualism* 17(1). 1–14.
- Davies, Eirlys E. & Abdelâli Bentahila. 2006. Code switching and the globalisation of popular music: The case of North African rai and rap. *Multilingua* 25(4). 367–392.
- Davies, Eirlys E. & Abdelâli Bentahila. 2008. Code switching as a poetic device: Examples from rai lyrics. *Language & Communication* 28(1). 1–20.
- de Ruiter, Jan Jaap. 2006. *Les jeunes Marocains et leurs langues*. Paris: L’Harmattan.
- de Ruiter, Jan Jaap, Karima Ziamari & Driss Meskine. 2014. *Le marché sociolinguistique contemporain du Maroc*. Paris: L’Harmattan.
- Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12(4). 453–476.
- Ennaji, Moha. 2005. *Multilingualism, cultural identity, and education in Morocco*. New York: Springer.
- Falchetta, Jacopo. 2019. *The social connotations of linguistic variation in a Moroccan urban context: The case of Temara*. Université d’Aix-Marseille (AMU). (Doctoral dissertation). <https://tel.archives-ouvertes.fr/tel-02571596> (15 February, 2022).
- Ferguson, Charles A. 1959. Diglossia. *Word* 15(2). 325–340.
- Ferguson, Charles A. 1991. Diglossia revisited. *Southwest Journal of Linguistics* 10(1). 214–234.
- Gumperz, John J. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.
- Hachimi, Atiqa, Jacopo Falchetta & Montserrat Benítez Fernández. 2022. Contextualizing the rise of vernacular Arabic in globalized North Africa. *International Journal of the Sociology of Language* 2022(278). 1–22.

- Khoumssi, Khawla. 2020. Attitudes of young Moroccans towards the use of code-switching in their Facebook interactions. *Journal of English Language Teaching and Linguistics* 5(3). 469–493.
- Labov, William. 1973. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Lahlou, Moncef. 1991. *A morpho-syntactic study of code-switching between Moroccan Arabic and French*. The University of Texas at Austin. (Doctoral dissertation).
- Manfredi, Stefano, Marie-Claude Simeone-Senelle & Mauro Tosco. 2015. Language contact, borrowing and codeswitching. In Amina Mettouchi, Martine Vanhove & Dominique Caubet (eds.), *Corpus-based studies of lesser-described languages: The CorpAfroAs corpus of spoken AfroAsiatic languages*, 283–308. Amsterdam: John Benjamins.
- Moatassime, Ahmed. 2006. *Langages du Maghreb face aux enjeux culturels euro méditerranéens*. Paris: L’Harmattan.
- Myers-Scotton, Carol. 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Peirce, Charles Sanders. 1933. On the algebra of logic: A contribution to the philosophy of notation. In *Collected papers of Charles Sanders Peirce*, 359–403. Cambridge: Harvard University Press.
- Pellegrini, Chloé. 2019. *L’enseignement des langues à l’école publique au Maroc: Construction des savoirs, identités et citoyenneté*. Université d’Aix-Marseille (AMU). (Doctoral dissertation).
- Post, Rebekah Elizabeth. 2015. *The impact of social factors on the use of Arabic-French code-switching in speech and IM in Morocco*. The University of Texas at Austin. (Doctoral dissertation).
- Redouane, Rabia. 1998. Arabisation in the Moroccan educational system: Problems and prospects. *Language Culture and Curriculum* 11(2). 195–203.
- RGPH. 2014. <http://rgphentableaux.hcp.ma/> (15 February, 2022).
- Sayahi, Lotfi. 2014. *Diglossia and language contact: Language variation and change in North Africa*. Cambridge: Cambridge University Press.
- Silverstein, Michael. 1976. Shifters, linguistic categories, and cultural description. *Meaning in Anthropology*. 11–55.
- Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23(3-4). 193–229.
- Tamer, Youssef. 2003. *Language and elementary education in Morocco: A sociolinguistic approach*. Rabat: Mohammed V University-Agdal. (Doctoral dissertation).

- Tamer, Youssef. 2006. What language situation do Moroccans favour? Arabization or bilingualism. In *L'arabe dialectal: Enquêtes, description, interprétations, Tunis, Actes du 6e colloque international de l'Association Internationale de Dialectologie Arabe*, 435–454. Tunis: Centre d'Etudes et de Recherches Economiques et sociales.
- Ziamari, Karima. 2007. Development and linguistic change in Moroccan Arabic-French codeswitching. In Catherine Miller, Enam Al-Wer, Dominique Caubet & Janet C. E. Watson (eds.), *Arabic in the city: Issues in dialect contact and language variation*, 275–290. London: Routledge.
- Ziamari, Karima. 2009. Moroccan Arabic-French codeswitching and information structure. In Jonathan Owens & Alaa Elgibali (eds.), *Information structure in spoken Arabic*, 243–259. London: Routledge.
- Ziamari, Karima. 2018. Determiner phrase: How specific is it in Moroccan Arabic-French codeswitching? In Stefano Manfredi & Mauro Tosco (eds.), *Arabic in contact*, 296–311. Amsterdam: John Benjamins.
- Ziamari, Karima & Alexandrine Barontini. 2013. Ana: Parlez-vous arabe marocain? Quand les séries réconcilient avec la darija. In Montserrat Benítez-Fernández, Catherine Miller, Jan Jaap de Ruiter & Youssef Tamer (eds.), *Évolution des pratiques et représentations langagières dans le Maroc du XXIe siècle*, vol. 1, 119–143. Paris: L'Harmattan.

Chapter 3

The discursive construction of code-switching in Yanito among the young population of Gibraltar

Marta Rodríguez García

University of Basel

Gibraltar is a small British enclave located in the south of the Iberian Peninsula. Despite its size, Gibraltar presents a historically, culturally, and linguistically rich landscape. Linguistically speaking, this speech community is characterised by a process of language shift in which Andalusian Spanish, the language of thousands of Gibraltarians before the 1970s, has slowly been replaced by the official language of the territory, British English. In this process, Yanito – a code-switching variety with lexicon from other Mediterranean languages – emerged as the vernacular language of the population. Despite the interest and complexity of this linguistic community, little is known about language variation or the use and evolution of this local dialect in younger speakers.

In this chapter, I study the structure and functionality of code-switching among young adults (16 to 35 years old) in Gibraltar in a sample of five focus groups collected between 2020 and 2021. I explore the use of code-switching in informal discourse from a more holistic perspective by combining structure and functionality. I first examine the fine line between structural categories and highlight categorical limitations between insertions and alternations. I then reflect on the importance of studying the functionality of code-switching from a sequential and interactional perspective. In doing so, this chapter offers a contribution to the literature on English and Spanish code-switching with an original methodology and from a European perspective.



1 Introduction

Perched on the southern tip of the Iberian Peninsula, Gibraltar is one of the two Pillars of Hercules that marked the western limits of navigation for the ancient Mediterranean world. Consisting of a ridge rising 421 metres above sea level, the territory spans 6.8 square kilometres and is visible from the neighbouring areas in Spain and Africa. With a current population of around 34,000, Gibraltar has always been a melting pot of different cultures, languages, and religions. It has been a British overseas territory since 1713, at which point British English became its official language – coexisting with Andalusian Spanish, which kept being used as a lingua franca until language shift settled. This Spanish variety is also the language of the neighbouring community and is still spoken by some Gibraltarians in familiar and more informal settings (Moyer 1993; Kellermann 2001). The coexistence of these two languages, which differ in prestige and status, has resulted in a vernacular and code-switching variety known as *Yanito*. Due to Gibraltar's location and history, its community is not only bilingual, but in fact polyglot. The local linguistic repertoire and code-switching varieties are enriched with terms from Genoese, Hebrew, Arabic, Maltese, and Portuguese as a result of the human and commercial hustle and bustle that has characterised the history of the territory.

The border community of Gibraltar has been extensively studied and analysed by historians and anthropologists (López de Ayala 1782; Gold 2005; Grocott & Stockey 2012). It has also captured the attention of linguists and sociolinguists (Lipski 1986; Moyer 1998; 2002; Kellermann 2001; Levey 2008; Loureiro-Porto & Suárez-Gómez 2017; among others), who, over the course of time, have aimed to describe the social, cultural, and linguistic situation of this community. In particular, this study seeks to fill a gap in the literature on bilingualism and code-switching as a conversational strategy in the bilingual discourse of young adults in Gibraltar. It does so by taking a conversationalist and interactional approach to exploring the use of structural and functional categories that pose a challenge to the account of code-switching. Thereby, the analysis presented here adds to the comprehensive and growing literature on English–Spanish code-switching from the broader European context, which is of interest for comparative inquiries into bilingualism on a wider basis. The general purpose of this study is to explore (1) the structure of code-switching and (2) the function of code-switching in a sample of 15 speakers, distributed across five focus groups of young adults in Gibraltar aged between 16 and 35, and to stake out the field for further research on this topic.

The chapter begins with a short description of the bilingual community of Gibraltar that reconstructs the social and historical processes that shape bilingualism and language shift (Section 2). The account of the methods that follows (Section 3) is unusually elaborate to allow for an informative description of the corpus generated by the five focus groups and a novel methodology, relying on a mystery game and the video-conferencing software Zoom. To achieve a more holistic picture of the complex phenomenon of code-switching in conversation, the analysis of the data combines two different angles: structural (Section 4) and functional (Section 5). Code-switching is first studied from a structural perspective, following Muysken's strategies of code-switching (Muysken 2000; 2013), before being examined as a conversational and interactional tool by exploring language negotiation, turn-taking, and shifts in settings and topic.

2 Language shift in Gibraltar

2.1 From Spanish as a lingua franca to an English-only policy:

Language shift in Gibraltar

Given what is happening in Gibraltar today, the way demographically Gibraltar is changing is concerning in some respects. For example, a lot of our young people are losing their second language, Spanish. We are losing bilingualism in favour of seeing a younger generation speak exclusively English. We want to control that if we can because bilingualism is an advantage. (Chief Minister Fabian Picardo, Uncorrected oral evidence: The UK-Spain agreement on Gibraltar 2021)

Despite major social and political changes, the linguistic situation in Gibraltar remained relatively stable until the 1970s (Mariscal Ríos 2014). After the British conquest in 1713, English was introduced as the sole official language. However, Spanish, or more precisely, Andalusian Spanish, was still spoken by the vast majority of the population in a wide range of contexts and was even considered a lingua franca of the territory (Moyer 1993; Mariscal Ríos 2014). After the Second World War, several factors contributed to major changes in the linguistic situation in Gibraltar and the establishment of an English-only policy. These included the evacuation of the civil population to other British territories and the return to Gibraltar of those families after the war, the introduction of the British schooling system, the closure of the border with Spain between 1969 and 1981, and the development of negative attitudes towards Spain and the Spanish language that promoted the use of English and the concept of English as 'the language of

opportunities'. From the 1970s onwards, British English has been increasingly established as a significant part of the linguistic landscape, along with Gibraltarian English (Levey 2015; Suárez-Gómez 2020), and a diminished Andalusian Spanish.

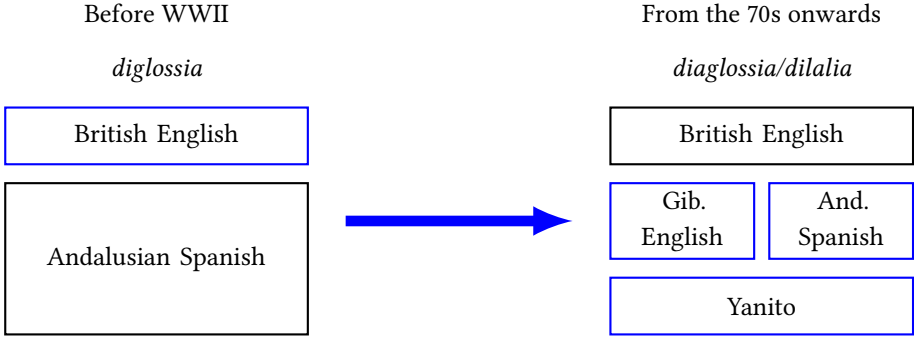


Figure 1: Language shift in Gibraltar from diglossia to diaglossia/dilalia (Rodríguez García & Goría 2021).

Figure 1 shows that the territory has undergone significant linguistic changes. From an extended diglossia (Fishman 1967; Auer 2005) with British English as the ‘high variety’ and Andalusian Spanish as the ‘low variety’, there has been a development to what could be better described as diaglossia (Auer 2005) or dilalia (Berruto 2005). This means that the structural and functional separation between the standard language and the local varieties, as well as the domains and specific pragmatic functions for each language, can no longer be clearly drawn. The bilingual situation is marked by intermediate variants between standard and dialect varieties and speakers shifting from a dialectal variant to a standard one, adapting to their situation and their audience (Rodríguez García & Goría 2021). Shifts can be encouraged by several factors and situations, such as the relationship between speakers and the formality of the speech events, the topic of the conversation, and the medium (Auer 2005; 2014). This is where Yanito, as the local code-switching variety, emerges.

Local writers and former educators such as Charles Durante believe there is a need to raise awareness of Gibraltar’s linguistic richness (Durante 2019) and value Yanito as a local form of communication that fosters social interaction, creates a sense of belonging, and reinforces the identity of the Gibraltarian, thereby allowing people to be identified both within and outside their community. However, as an oral and local variety, Yanito is exposed to constant variation across generations. This article examines the current variant of Yanito used by a new generation of young adults and highlights code-switching patterns observed in a small sample of young adult speakers.

2.2 Code-switching in Gibraltar

Code-switching is a widely used strategy to enhance linguistic and social identity in bilingual and multilingual contexts; it is also a gauge of sociolinguistic processes and changes in such communities. In this context, Gibraltar presents a complex sociolinguistic situation in which Yanito – which I define here as the local speech variety characterised mainly by English–Spanish code-switching – is widespread and enjoys certain covert prestige among speakers. However, the current language shift is affecting Yanito. The decline in popularity of Spanish among younger Gibraltarians, resulting from the introduction of English as a medium of instruction in school after the Second World War, and the dominance of English in private and public spheres are reshaping communication and the distribution of languages in conversation.

Code-switching eludes easy explanation, as recognised by experts like Myers-Scotton (2002), Gumperz (1977), and Auer (1998), who have all long tried to account for the multiple factors affecting and conditioning the choice of a code in conversations, discourses, and interactions. This general difficulty is exacerbated by the well-documented differences in various bilingual communities. Hence, for the purpose of this chapter, I combine Auer's (1998) and Gumperz's (1977) broad definitions of the term, which have served as suitable definitions for numerous studies on code-switching over the decades due to their flexibility and adaptability. Based on both authors, I understand code-switching as “the juxtaposition of passages of speech belonging to (at least) two different grammatical and semi-otic systems, within the same exchange” (Gumperz 1977: 1). This broad definition makes it possible to adopt a functional and holistic approach to the phenomenon of code-switching. This facilitates a nuanced understanding of the function and sequential organisation of code-switching in conversation and explains bilingual conversation in relation to language choices at a conversational and interactional level.

Existing research on the linguistic situation of Gibraltar points towards the generational loss of Spanish and to language shift (Moyer 1993; Kellermann 2001; Feijóo Rodríguez 2015) and focuses on the local code-switching variety and its form – accounting for both its use in familiar and informal situations and the complexity of its form and function. The twentieth century witnessed an increasing interest in the structures of bilingual speech in Gibraltar, and although much of this interest is confined to unpublished master's theses, there are several investigations and ongoing projects pointing to rich and variable code-switching strategies in Gibraltar. These strategies include the selection of a main language for interaction, the negotiation of languages between turns, and various examples of

intersentential and intrasentential code-switching behaviour (Moyer 1998). More recent studies on Yanito shed light on the reduction in the use and variety of bilingual structures. For instance, Goria (2017) reports the rise of a mixed code involving the use of fixed switching patterns in which the direction of the switch is constrained and the pragmatic meaning behind the use of certain bilingual markers is lost. Furthermore, Weston (2013) offers evidence of notable diachronic and synchronic differences at the individual level, demonstrating the need for further research at both individual and community level.

Although methods used to study Yanito have become more systematic, the corpora still lack sufficient data on younger generations of speakers. This data gap makes it difficult to provide an account of the current state of bilingual discourse in younger Gibraltarian generations. While the loss of Spanish in favour of English is undeniable at a community level, more has to be done to understand the state of Yanito. A preliminary study by Chevasco (2019) on language use and attitudes registers an unexpected increase in the use of and positive attitudes towards Yanito among young adults in informal and family settings, and reports Yanito as an important symbol of identity and as a strategy to reaffirm local identity in relaxed conversations (Chevasco 2019: 63).

An in-depth study of code-switching in younger generations which accounts for both structure and functionality promises a substantiated understanding of the linguistic situation of this territory and the future of code-switching in that setting. Based on the well-established knowledge that bilingual speech may be influenced by a great number of external factors (e.g., the communicative situation, the participants, the topic, etc.) and internal factors (e.g., grammatical constraints, the languages involved, the proficiency of speakers, etc.), I study the bilingual speech of young adults in Gibraltar as a communicative strategy with social meaning.

3 Methodology and research corpus

Previous literature has found that Yanito often appears in informal local conversations and that it can be triggered by various factors, such as the relationship between participants and the topic and purpose of the conversation (Moyer 1993; Kellermann 2001; Levey 2008; Weston 2013). Therefore, a flexible methodology which can account for the different situations that trigger code-switching is needed. A pilot study in which I conducted semi-directed interviews confirmed the value of focus groups, where participants discuss various topics with people who are close to them, such as friends or family members, following guidelines for the conversation that contain topics, pictures, statements, and a game.

The use of the (online) focus-group technique, as well as the guidelines with topics and activities for the conversation and the close relationship between participants, made it possible to observe language in a nearly natural manner while covering topics of interest for the study (Labov 1984). The guidelines comprise three parts. The first prompts participants to converse about various topics, from informal subjects (childhood memories, funny anecdotes, and holidays) to more formal topics (e.g., COVID-19 and political measures, education, job opportunities). In part two, participants are given visual and written prompts that encourage them to converse about the influence of different cultural, linguistic, and identity-related aspects. The pictures depict cultural festivities and items and images of the border, the late Queen Elizabeth II, or the Prime Minister of Gibraltar. The written statements are from Gibraltarians and refer to the local code-switching variety and the multiculturalism of the territory. Here, again, the formality of the topic changes according to the pictures and the situations the participants describe. Emotions such as nostalgia, happiness, or anger appear often and impact the participants' language. Finally, in part three, participants play a mystery game in which they have to solve a case by reading and talking about a series of bilingual cues. This part makes it possible to observe the language participants use while playing and how they react to Spanish input. Overall, the focus groups enabled me to gather data that allow a close analysis of code-switching and its functionality; namely, the well-known characterisations of situational, metaphorical, and conversational switching. After finishing this conversation, participants are asked to complete an online questionnaire that collects socio-demographic information and individual participants' views on language use and competence. Given the constraints on fieldwork during the COVID-19 pandemic, I opted for an online, corpus-based methodology with focus groups of three people each. As it was difficult to recruit younger participants (aged between 16 and 21), the online corpus was supplemented with data gathered in face-to-face interactions in May 2021 in three state secondary schools – Westside School, Bayside School, and Gibraltar College – following the same focus-group methodology.

The groups were established on the basis of age in order to determine how the language shift process is reflected in the language of young adults. The age groups were created attending to historical and social criteria. Concretely, the first age group (A) corresponds to school students with little experience abroad who were born in the 2000s (therefore, born and raised in a European and open-border Gibraltar); group (B) corresponds to young adults who were born during the launch of the European Economic Union and with possible experience abroad (educational or professional purposes); and group (C) represents those who were

born in a period of memories and resentment and a less porous border, but who also experienced a more balanced bilingualism within the community (those may also be the parents of a new generation). As it can be seen, the division of the groups according to age allows to explore language behaviour as a reflection of political, historical, social, and linguistic changes.

Table 1: Distribution of five focus groups (sample for analysis).

Age group A (16 to 21)	Age group B (22 to 28)	Age group C (29 to 35)	Mixed group
A1	B1 B2	C1	BA1

For this chapter, a first sample of five focus groups was studied (see Table 1): one focus group of younger speakers aged between 16 and 21 (A1); two focus groups of speakers aged between 22 and 28 (B1 and B2); one focus group of older speakers aged between 29 and 35 (C1); and one mixed group with speakers in the younger or middle age ranges (BA1). The sample was balanced, albeit small, totalling 15 participants. It included people with diverse cultural, ethnic, and religious backgrounds who had had diverse experiences abroad. In fact, 9 out of the 15 participants of this study confirmed having educational or professional experiences abroad. A gender balance was also achieved. The focus group discussions lasted approximately 30 to 45 minutes each. For the analysis, a total of 241 minutes (approximately 40,800 words) of interaction was transcribed in ELAN and subsequently analysed. In this first exploratory analysis, I used orthographic transcription to account for cases of code-switching, then classified the occurrences attending to the structural patterns. This was followed by a qualitative analysis of the instances with a focus on the sequence of the conversation and the structure of the discourse in order to explore the functions of code-switches. The results obtained from this provided a basis for determining which avenues of inquiry would be most promising for further research.

The examples provided for the analysis (Section 4 and 5) are extracted from the original transcriptions and followed by the English translation. For a better understanding of the examples, an appendix with transcription conventions is attached at the end of the chapter.

4 Yanito from a structural perspective

For the structural analysis, I focus on three patterns proposed by Muysken (2013): insertion, alternation, and backflagging. The first structural pattern, insertion, refers to a chunk of language B – in this case, Spanish – inserted into a sentence in language A – in this case, English (Muysken 2013: 712).

- (1) SOPHIA: ¹ *La querida* (.) why should why should *la querida* want to: maybe he's maybe he was gonna leave his wife *para la querida*
 SOPHIA: The lover (.) why should why should the lover want to: maybe he was gonna leave his wife for the lover

In Example (1), the lexical element *querida*, meaning ‘mistress’, appears together with the feminine definite article *la* inserted into an English clause. The expression *la querida* is taken from Spanish without affecting the structure of the sentence. In examples of insertion, we observe that single elements of language B are inserted within an otherwise language A clause. Sometimes, however, we find that a longer chunk of text of language B, which constitutes a clause, alternates with a chunk of text in language A, as in the following example from Sophia, a speaker from group C1.

- (2) SOPHIA: *Vamos a ver dónde está esto* | can you send me a screenshot?
 SOPHIA: Let's see where this is | can you send me a screenshot?

Example (2) is a clear case of alternation or “a succession of fragments in language A and B in a sentence” (Muysken 2013: 713). In this case, there is a first declarative statement in Spanish which is followed by a question in English. The appearance of a Spanish utterance at the left margin of a new context is very common among speakers from different focus groups. However, this phenomenon is not consistent among speakers, and alternations appear as a result of participant- and discourse-related switches, especially in language negotiation.

This leads to the next pattern: backflagging. According to Muysken (2013: 713), in this structural pattern, discourse markers from a heritage language are inserted in a sentence in an L2. This is the most common and consistent pattern among young adult speakers in Gibraltar. However, it is worth noting that the matrix language of the sentence, English, constitutes the L1 or language A for most of the speakers in my corpus. In the speech of young Gibraltarian adults,

¹In order to maintain the anonymity of the participants, a code was attributed to them. The names that appear in this chapter are fictitious, but they have been selected from names heard in Gibraltar during my research stays.

the insertion of Spanish markers into an otherwise English discourse is noteworthy (Goria 2017), even if they do not consider themselves intermediate or advanced speakers of Spanish, as stated in my socio-demographic and linguistic questionnaire and sometimes in the focus group itself. In this way, the markers could be some kind of ‘inherited’ material from previous generations or even a symbol of identity for younger generations. It is common among the five groups to find examples of the use of discourse markers such as *pero* and *bueno* or the question tag *¿no?* On many occasions, they appear as single insertions within an otherwise English clause. In Example (3), Sophia makes use of the discourse marker *bueno* to take her turn, change the topic of the conversation, and initiate the clause.

- (3) SOPHIA: *Bueno* should we think of like anecdotes or funny moments?
SOPHIA: Well should we think of like anecdotes or funny moments?

All three types of structural strategies described by Muysken are found in the corpus. In some cases, the three types appear in the discourse of the same speaker, as has been illustrated in the case of Sophia in Examples (1), (2) and (3), while in some groups, certain patterns are more consistent in one speaker than others. Table 2 outlines the phenomena found in the analysis of the five focus groups.

Table 2: Structural patterns found in the focus groups.

Code-switching strategies	A1	B1	B2	BA1	C1
Insertion	Sp > E	*Sp > E	Sp > E	Sp > E	*Sp > E
Alternation	(√)	√	(√)	(√)	√
Backflagging	√	√	√	√	√

As illustrated in Table 2, all patterns emerge in the five focus groups. However, insertion is the most common pattern among speakers, together with backflagging, which is also the most consistent one. In the case of insertion, I find both English and Spanish insertions in groups B1 and C1. The directionality of the switch is not always clear and this is why those groups are marked with an asterisk in the table. This is something worth studying, since previous studies on Gibraltar account for the insertion of elements from a dominant and prestigious language into a less prestigious or non-dominant language, but not in the opposite direction – something that is also very common in my corpus (Moyer 1998; Feijóo Rodríguez 2015). In the case of alternations, greater differences among speakers

in the same conversation can be observed, and while age seems to play a role in the use of different patterns of alternations and insertions, this is clearly not the only factor. In some cases, speakers who belong to the same age group show great variation in the use of alternation, such as in A1, B2, and the mixed group, BA1. In the process of classifying the examples from the corpus, the distinction between insertions and alternations occasionally becomes fuzzy. Speakers make use of longer insertions such as *y todo ese rollo* ‘and all that stuff’ which are difficult to classify according to structural parameters, as illustrated in Example (4). The Spanish sentence starts with the conjunction *y* ‘and’, followed by the indefinite pronoun *todo* ‘all’ and the demonstrative *ese* ‘that’, which precede the noun *rollo*, in this case, ‘stuff’. This cluster serves as a way of closing the previous topic but also serves to start the next alternation. In this case, Spanish is identifiable as language B in the quote, but the chunk results in a bigger switch to language B.

- (4) MOHAMED: The university (.) what have affected you guys? Tell us more about *bueno* you mentioned *que* they’ve contacted you for eh test when you go back home for Christmas *y todo ese rollo* so *es un rollo pero al mismo tiempo* | it’s maybe a mess *¿no?*

MOHAMED: The university (.) what have affected you guys? Tell us more about well you mentioned that they’ve contacted you for eh test when you go back home for Christmas and all that stuff so it’s (messy) stuff but at the same time | it’s maybe a mess isn’t it?

The functionality of the chunk *y todo ese rollo* in language B, Spanish, is also similar to structural markers or discourse organisers, since it serves as a recurrent mechanism for the speaker to close a topic. But later, the same noun *rollo* serves different purposes in the subsequent alternation. As I mentioned, the insertion of discourse markers is very common among young adult speakers. The examples observed again raise the question of whether they should be considered code-switching, as their form is similar to insertion. However, the use of more complex discourse markers and the combination of elements stress the need to pay special attention to the use of these pragmatic, discursive elements.

- (5) MOHAMED: Our identity rests solely on on on being part of [Britain *ahí está ahí está* on being part of Britain]
 AKRAM: [our history yeah the history of Gibraltar *sí señor*]
 MOHAMED: *escúchame el nuevo Bayside García y los colegas*
 AKRAM: social distancing social distancing

MOHAMED: *es que* none of us none of us were able to actually to enjoy the new Bayside *esa es la cosa*

MOHAMED: Our identity rests solely on on on being part of [Britain there you are there you are on being part of Britain]

AKRAM: [our history yeah the history of Gibraltar yes, sir]

MOHAMED: listen to me the new Bayside García and the friends

AKRAM: social distancing social distancing

MOHAMED: that is none of us none of us were able to actually to enjoy the new Bayside that's the thing

In this short extract from a conversation between two young males in Example (5), different markers enter into play: to express agreement, such as *ahí está* 'there you are' or *sí señor* 'yes, sir'; to get attention and initiate a new topic, such as *escúchame* 'listen to me'; and to initiate and close an intervention, such as *es que* 'is that' and *esa es la cosa* 'that's the thing'. As observed, their form and functionality are also of a different complexity, which makes it difficult to label them or group them into categories. Further research has to be done to study the functionality of those markers and to observe if their meaning and use comes from Spanish or English, or even if new meanings are arising from the contact of both languages. This question leads us to the same previous dichotomy of insertions versus alternations and even to code-switching versus code-mixing (Auer 1998). The complexity of the markers is higher among older speakers, as illustrated in an extract between two participants from group C1 in Example (6). Here, the combination of markers is very common among different participants (more than 100 instances of combinations of Spanish markers and around 70 combinations of Spanish and English markers), as in the example of *bueno espérate ¿te acuerdas?* 'well, wait, do you remember?'.²

- (6) LAURA: And it was dark on the way back and you were thinking if I run over a car (.) *no pasa nada* you have to run it over *¿te acuerdas?*

HELEN: yes (4.1) *bueno espérate ¿te acuerdas?* Was this the same time on the way back?

LAURA: ye::s that was pitch bla::ck!

LAURA: And it was dark on the way back and you were thinking if I run over a car (.) nothing happens you have to run it over do you remember?

HELEN: yes (4.1) well wait do you remember? Was this the same time on the way back?

LAURA: ye::s that was pitch bla::ck!

The function of the whole phrase *bueno espérate* is reminiscent of an utterance employed by Sophia – from the same focus group C1 – at another point in the conversation in which she makes use of the element *un momentito*, ‘a moment’ (Sophia: I’m just gonna quickly have a look at the questions *vale un momentito?* (.) I’m just gonna put [...]). In the case of *bueno espérate*, the chunk serves as strategy to give the speaker time to think about something and to do something in the case of *un momentito*, that is immediately followed by a stop. In particular, the use of discourse markers from the heritage language, Spanish, seems to show characteristics from both insertions and alternations, making it difficult to define the code-switching patterns. Sentential based classifications, such as Poplack’s (1980) inter- and intrasentential switching could help solving some of these difficulties, but it still seems clear that there is a strong need to further study the functionality of groups of structural patterns within discourse in order to account for the use of both languages in Gibraltar.

By analysing the structure of code-switching, we can better visualise the factors, strategies, and outcomes in code-switching. In addition, it also helps to justify language shift within a community by studying the matrix language and the directionality of the switch; English is by far the preferred language and the matrix language in most cases, while switches to Spanish appear often, whereas previous studies on first and second generations have shown completely the opposite distribution of languages (Moyer 1993; Kellermann 2001; Weston 2013). Something I also observed in the examples is the peripheral location of the switches to Spanish, which tend to appear to the left (in most cases) or to the right of the turn (ongoing analysis, see Goría 2021). However, it seems that structure is not enough to understand how speakers code-switch and why they do it. Therefore, a detailed study of the functionality of those patterns in conversations and interactions is needed in order to better understand the motivation of the speakers, which I will turn to in the next section.

5 Yanito from a discursive and conversational perspective

5.1 Language negotiation

When initiating a conversation or when there is a change in a setting or in terms of the participants, speakers adapt their speech to the new environment. In this section, I present some examples that account for the alignment of participants in interaction and the distribution of languages by turns. In some cases, the three participants of the focus group change the code from the beginning of the conversation and show a rich repertoire of bilingual structural strategies, as in the case of C1 in Example (7).

- (7) SOPHIA: Girls
HELEN: [okay]
SOPHIA: [...]
HELEN: okay *bueno* [let's start]
SOPHIA: [I don't wanna touch] my phone and I've got the questions *mira en el laptop*
LAURA: *y yo y yo* my [laptop is (...)]
HELEN: *bueno* [let's (.) *¿qué dice?*]
LAURA: we met in school!
HELEN: *bueno espérate* let's go through the first section ((reading the guideline)) "your friendship (.) how did everything start?"
LAURA: I don't know
SOPHIA: Westside *¿no?* I think it started
LAURA: we were all in different schools *¿no?*
HELEN: *sí yo* [*fui a:*]
SOPHIA: yeah Westside
HELEN: *yo fui a* Saint Catherine's first school *y después* Sacred Heart
- SOPHIA: Girls
HELEN: [okay]
SOPHIA: [...]
HELEN: okay well [let's start]
SOPHIA: [I don't wanna touch] my phone and I've got the questions look on the laptop
LAURA: and I and I my [laptop is (...)]
HELEN: well [let's (.) what do you say?]
LAURA: we met in school
HELEN: well wait let's go through the first section ((reading the guideline)) "your friendship (.) how did everything start?"
LAURA: I don't know
SOPHIA: Westside wasn't it? I think it started
LAURA: we were all in different schools weren't we?
HELEN: yes I [I went to]
SOPHIA: yeah Westside
HELEN: I went to Saint Catherine's first school and then Sacred Heart

In Example (7), the speakers start the conversation by establishing English as their matrix language, but Helen quickly starts inserting Spanish elements, such as the marker *bueno* 'well' to get the attention of the other speakers, to which

Sophia reacts with a bigger switch: an alternation with an English insertion (*mira en el laptop* 'look on the laptop') at the end of her intervention, accepting the switch as a strategy for the conversation. Laura also initiates her turn in Spanish (*y yo y yo* 'and me and me') as a response of agreement to the use of both languages in interaction. In this example of language negotiation, we observe that even though English is used as the language of communication, speakers agree on using Spanish for participant-related purposes, such as agreeing on ideas – *y yo y yo* 'and me and me' and *sí* 'yes' – or reacting to a previous interaction – *¿qué dice?* 'what do you say' and *espérate* 'wait'. Spanish is also used for discursive purposes by framing the sequentiality of the events, for example with *y después* 'and then'. Speakers agree on the language of interaction because they all know that their proficiency and linguistic strategies are similar, which also makes speakers display a great number of structural strategies of code-switching within the conversation. This is more an exception than the norm, since in most of the focus groups, speakers take longer to negotiate the language of communication.

In the B1 group, two speakers with a high level of proficiency in both languages (Daniel and David) switch from English to Spanish on multiple occasions, while Alex uses English most of the time, as can be seen in Example (8).

- (8) ALEX: We used to be in different schools and in different football teams and then we played together (.) for the selection (.) for the gfa
 DANIEL: ((addressing David)) *¿te acuerdas que jugabas con él o no? ¿y era bueno de chico?*
 DAVID: *malísimo* (.) *bueno* ((addressing Alex)) recently *te di la foto*::
 ALEX: yeah recently you gave me (.) actually
 DANIEL: *¿y eso qué era gfa?*
 DAVID: *Entonces (año) nueve*
 DANIEL: (under the) eleven rather than
 DAVID: and eh:: eleven *sería ¿no?* or on the ninth ((addressing Alex))
 ALEX: or the ninth ninth (.) we were younger
 DAVID: yeah
 ALEX: ((addressing Daniel)) *¿y tú?*
 DANIEL: me what?
 ALEX: knowing David
 DANIEL: David? (0.7) *¿de dónde nos conocemos?* ((addressing David)) *ahí en la escuela ¿no?*

ALEX: We used to be in different schools and in different football teams and then we played together (.) for the selection (.) for the gfa

DANIEL: ((addressing David)) do you remember that you played with him or not? And was he good as a child?

DAVID: very bad (.) well ((addressing Alex)) recently I gave you the picture

ALEX: yeah recently you gave me (.) actually

DANIEL: and what was that gfa?

DAVID: then (year) nine

DANIEL: (under the) eleven rather than

DAVID: and eh:: eleven would be wouldn't it? Or the ninth ((addressing Alex))

ALEX: or the ninth ninth (.) we were younger

DAVID: yeah

ALEX: ((addressing Daniel)) and you?

DANIEL: me what?

ALEX: knowing David

DANIEL: David? (0.7) from where do we know each other? ((addressing David)) there in the school wasn't it?

Speakers switch and accommodate in many parts of this conversation when addressing each other. In Example (8), Alex accommodates Daniel by using Spanish when asking him how he met David with the question *¿y tú?*, 'and you?'; Daniel responds in English, accommodating Alex's language preference (me what?) and switches again to Spanish when he looks for confirmation from David as to where they got to know each other (*¿de dónde nos conocemos?*). In the case of B1, the group seems to be very aware of their linguistic preferences, even though they are probably not aware of their multiple accommodations during the conversation and the increase in the use of Spanish in Alex's interventions. They even mention this explicitly when discussing language shift and the linguistic situation of Gibraltar (see Example 9).

(9) ALEX: No we mix a lot *sin darnos cuenta, la verdad*

DANIEL: *pero lo bueno es que los padres hablan español y los niños [no hablan ¿cómo puede ser?]*

DAVID: [*y los niños no (no hablan) mi her[mano]*]

ALEX: [*mi hermana*] (.) *mi hermana igual*

DANIEL: ((addressing David)) *bueno tú hablas mucho más español que tu hermano*

DAVID: *mi hermano no lo habla pero todos los amigos de mi hermano son ingleses de Soto (.) esos son tan English*

DANIEL: *pero lo sabe*

ALEX: yeah I know

DANIEL: *pero ¿los padres tuyos?*

ALEX: it's more English yeah

DANIEL: ah they don't speak Spanish

DAVID: they don't speak *así*

ALEX: *mi hermana no habla nada eh*

DANIEL: *claro le cuesta (.) tú ahora hablas mucho más que antes (.)* probably because of us

ALEX: probably

ALEX: Not we mix a lot without realising, that's the truth

DANIEL: but well the thing is that the parents speak Spanish and the children [don't speak how can it be?]

DAVID: [and the children don't (don't speak) my brother]

ALEX: [my sister] (.) my sister is the same

DANIEL: ((addressing David)) well you speak much more Spanish than your brother

DAVID: my brother doesn't speak it but all the friends of my brother are English from Soto (.) those are so English

DANIEL: but he can/knows

ALEX: yeah I know

DANIEL: but your parents?

ALEX: it's more English yeah

DANIEL: ah they don't speak Spanish

DAVID: they don't speak like that

ALEX: my sister doesn't speak at all

DANIEL: of course it's hard for her (.) you speak now much more than before (.) probably because of us

ALEX: probably

In Example (9), the speakers discuss the loss of Spanish in younger generations, starting with the following statement: 'Parents speak Spanish and their children don't' but they also reflect on their own situation and their linguistic preferences. They discuss the fact that their siblings are also speaking less Spanish because of their English friends and that Alex speaks Spanish even though his parents speak more English at home. Here, Daniel concludes that the fact that Alex has improved his level and use of Spanish may be *because of us*. This could refer to their current friendship, but probably also to the fact that they teach in the same school, which makes them spend a lot of time together.

Even though younger generations make more use of English, as the three speakers from the B1 group state, I still observed differences between participants in group A1. While most of them prefer to use English almost exclusively, others make use of both languages in discourse and even force English speakers to accommodate. In Example (10), the three speakers aged 17 and 18 discuss their holidays in Spain. In this part of the conversation, it seems that Dan wants Mike to switch to Spanish since he keeps asking him in this language ('where do you go', 'where in Spain?', 'with your family?', 'your friends as well or not?') and only stops asking for more information when Mike finally replies in Spanish: *sí hermano* 'yes, brother' and *no*.

- (10) MIKE: I didn't go on holiday last year 'cause Covid same as Lyan (.) and (.)
I'll be going on holidays this year in July middle of July
LYAN: ¿qué más? ¿adónde [vas?]
DAN: [dónde vas]
MIKE: Spain
LYAN: pero ¿a dónde en España?
DAN: Estepona::
MIKE: yes Estepona
LYAN: with a pool (.) bar?
MIKE: a villa
DAN: ah vale con tu familia ¿no?
MIKE: sí hermano
DAN: ¿tus amigos también o no?
MIKE: no
DAN: (0.7) ah vale (.) perfecto
LYAN: vale está bien ahora vamos hablar de (...)

MIKE: I didn't go on holiday last year 'cause Covid same as Lyan (.) and (.)
I'll be going on holidays this year in July middle of July
LYAN: what else? Where are you [going?]
DAN: [where are you going]
MIKE: Spain
LYAN: but where in Spain?
DAN: Estepona::
MIKE: yes Estepona
LYAN: with a pool (.) bar?
MIKE: a villa
DAN: ah well with your family right?

MIKE: yes brother

DAN: and your friends as well or no?

MIKE: no

DAN: (0.7) ah okay (.) perfect

LYAN: okay it's fine now let's talk about (...)

Generally speaking, switching between languages is accepted by speakers despite their linguistic preferences. Younger speakers start their conversations in English; however, as I showed in Examples (8) and (10), there are still some cases of accommodation when discussing different topics or changing the formality of the discourse. In this example, Mike finally accommodates Dan by responding with *sí hermano*. In the case of group BA1 in Example (5), in which the speakers have a Moroccan background, there is a part of the conversation in which participants start talking about their families and background, but when one of them switches to Moroccan Arabic, he immediately stops to say *sorry*. Speakers automatically accept this switching between languages by responding in Arabic and switching between languages.

In general, I observe that participants negotiate and re-negotiate language during the interaction and co-construct some kind of communicative style for the purpose of the conversation that is not directly related or correlated to pre-established sociolinguistic variables from each participant. It seems that language negotiation and accommodation is a common strategy among all speakers. Speakers shift from one code to the other, adapting to the participants and form of the conversation. Despite the individual and group differences, when completing their questionnaires, all speakers from the examples confirmed the use of English and Spanish, as well as Yanito in their homes (except for Helen in group C1, who just mentioned both languages) and when out and about or with friends (except Mike in group A1, who only mentioned English). Speakers also consider themselves good speakers of Yanito, assessing their speaking ability in Yanito as 4 or 5 out of 5 (except Mike, with 3 out of 5). As expected, most of the speakers evaluating their speaking abilities in Yanito as 'native' (5/5) were the ones displaying a more mixed speech. It is interesting though, that younger speakers from group A who also considered themselves native users of Yanito, still showed a clear preference for English in discourse (less than 250 Spanish words throughout the whole conversation).

5.2 Topic management

As Auer (2005) affirms, activities are not strictly linked to a certain language. In fact, some speakers may discuss a topic in language A while others will discuss

the same topic in language B. In my corpus, most topics are not tied to a specific language; however, it is true that some topics and activities – due to specific reasons that I will highlight in the examples – trigger the use of code-switching and lead speakers to renegotiate their language of interaction.

In Section 2, I discussed the concepts of diglossia and language shift in the territory, which are central to understanding why certain topics in Gibraltar may lead to a renegotiation of languages in conversation. In fact, previous literature on Gibraltar highlights the use of Spanish in informal conversations and topics (Moyer 1993) such as food, holidays, and free time. On the other hand, English continues to be used in formal settings and for educational, professional, political, and economic purposes. As I previously described, Gibraltarian young adults constitute a central group in the process of language shift. In this subsection, I examine the use of code-switching in some topics that were prompted by the guidelines used in the focus group conversations. I first refer to three topics: food, holidays, and free time. I then present examples of the use of Spanish in unexpected topics such as education. I examine this from a discursive and conversational perspective to get an understanding of the reasons behind the switch.

One of the most striking topics in terms of its rich vocabulary is food. Gibraltar's gastronomy is very varied and rich, and even though the transformation to a more British diet is observable on restaurants' menus, most families still follow some kind of Mediterranean diet, in which a lot of typical dishes keep their Spanish, Portuguese, or Maltese names. There are also many restaurants with Spanish names, such as *El Pulpero* mentioned in Example (11). This explains why conversations about food and gastronomy are commonly enriched by Spanish insertions and names; cultural terms that enrich local vocabulary. Furthermore, the proximity to Spain and the differences in prices also encourage some Gibraltarians to cross the border to visit the cheaper supermarket *Mercadona* or to enjoy a meal every now and then in a restaurant. In Example (11), we observe how speakers from group C1 describe their experience and memories of restaurants and markets in COVID-19 times. It is very interesting to observe how Laura automatically initiates her interaction in Spanish when she first describes her experience: *Todo está cerrado, yo no soy mucho de salir a cenar ni a comer* 'Everything is closed; I'm not much for going out for dinner'. However, she automatically alternates languages and continues her interaction in English. Despite this, Helen's short response, *ya* – an affirmation marker of response that literally means 'already' but is used as 'I know' – allows this reaffirmation of code-switching as a good strategy for conversation.

- (11) LAURA: *Todo está cerrado* (.) *yo no soy mucho de salir a cenar ni a comer* anyway so I don't miss anything because I like to cook *pero* every now and then I ordered takeaway
 HELEN: *ya*
 SOPHIA: right? I miss El Pulpero
 LAURA: I knew that (0.5) I don't really go out much
 HELEN: what I miss to be honest (.) I miss *como vamos a ver* like if I were in Gib *obviamente* I miss the chicken *pin-* the tandoori chicken *pinchito con las patatas* thing
 SOPHIA: I miss (...) *qué bueno*
 LAURA: *no bueno* but you could have that (.) restaurants are delivering
 HELEN: *ah bueno vale*
 LAURA: *oh me está entrando hambre*
 SOPHIA: *yo* home and cooking a lot more I'm doing lasagne *y le meto le meto* really nice eh lentils
- LAURA: Everything is closed (.) I'm not much for going out for dinner or lunch anyway so I don't miss anything because I like to cook but every now and then I ordered takeaway
 HELEN: yes
 SOPHIA: right? I miss El Pulpero
 LAURA: I knew that (0.5) I don't really go out much
 HELEN: what I miss to be honest (.) I miss like let's see like if I were in Gib obviously I miss the chicken the tandoori chicken chicken skewer with fries thing
 SOPHIA: I miss (...) how tasty!
 LAURA: no well but you could have that (.) restaurants are delivering
 HELEN: ah well okay
 LAURA: oh I'm getting hungry
 SOPHIA: I home and cooking a lot more I'm doing lasagne and I put I put inside really nice eh lentils

In Example (11), I also observed the use of a more complex insertion, *pinchito con las patatas*, in which the typical chicken dish *pinchito* 'chicken skewer' is accompanied with fries or *las patatas* but is also premodified by the noun phrase *tandoori chicken*. When discussing food, speakers show very rich ways of combining words and languages, adding English and Spanish adjectives and nouns as premodifiers and postmodifiers that are not easy to classify and require further analysis. It is also observable that at the beginning of the conversation, Spanish

insertions do not trigger the use of Spanish; instead, switching to Spanish appears as insertions of noun phrases or certain fixed expressions such as *qué bueno* or contextualisation elements such as *como vamos a ver* 'like, let's see'. However, by the end of this conversation, Helen and Laura show a preference for the use of Spanish in short interventions, and Sophia responds with the first pronoun in Spanish, *yo*, to take her turn and alternates between English and Spanish in her intervention. The way Sophia takes her turn with the first Spanish pronoun and continues the sentence in English is very interesting and not a common strategy for turn-taking, as I will show later in this section.

Two other topics of interest for the analysis of code-switching are holidays and free time. The proximity of Spain to the British enclave makes it possible for people to cross for leisure and for a holiday. Many participants mention their holidays in the nearby provinces of Malaga or Cadiz, referring to places and experiences. The memories of trips to Spain are sometimes expressed in Spanish, as can be observed in the following conversation from Example (12), in which speakers talk about their last holidays in Spain after the COVID-19 pandemic made them cancel their trip to Croatia: *No fuimos a Croacia por dos semanas, ¿no? Íbamos a ir* 'We didn't go to Croatia for two weeks, did we? We were going to go'.

- (12) ALEX: We were gonna go to Croatia but then we couldn't
DAVID: *no fuimos a Croacia por dos semanas, ¿no? Íbamos a ir*
DANIEL: *con un coche*
ALEX: *después de Ibiza*
DAVID: *a ustedes a ustedes Ibiza (0.5) a ustedes más es verdad (.)* like it affected you more
DANIEL: *de Ibiza*
ALEX: yeah
DANIEL: for sure
ALEX: *porque a mí no*
DAVID: *sí (.) yo*
DANIEL: *los holidays dices tú*
ALEX: *bueno ya tenías el fútbol por eso*
DANIEL: *ah porque tenía cosas ya*
ALEX: *bueno you weren't able to go anyway*
DAVID: *por eso*
DANIEL: *pero tú no tenías ningún holiday planned? (.) para este verano?*
DAVID: *yo quería ir a Ibiza*

- ALEX: We were gonna go to Croatia but then we couldn't
 DAVID: we didn't go to Croatia for two weeks, did we? We were going to go
 DANIEL: with a car
 ALEX: after Ibiza
 DAVID: to you to you Ibiza (0.5) to you more actually (.) like it affected you more
 DANIEL: of Ibiza
 ALEX: yeah
 DANIEL: for sure
 ALEX: because me it didn't
 DAVID: yes (.) I
 DANIEL: the holidays you say
 ALEX: well you got football already that's why
 DANIEL: ah because you already had things
 ALEX: well you weren't able to go anyway
 DAVID: that's why
 DANIEL: but you didn't have any holiday planned? (.) for this summer?
 DAVID: I wanted to go to Ibiza

In this extract from the conversation between Alex, Daniel, and David, Alex – who is dominant in English and uses this language in most of his turns – chooses to accommodate and renegotiate his language use in his intervention on a holiday in Ibiza: *Después de Ibiza* 'After Ibiza'. The same happens in his intervention about football when he addresses Daniel to explain that he had football and that is why he did not join them: *ya tenías el fútbol por eso*. However, it is not only the topic but also the sequence and negotiation of languages that are striking in this extract of conversation. It is interesting to observe that participants alternate between languages sequentially within turns and also insert the noun *holiday* from language A, with the masculine and plural pronoun *los*, but then follow this same noun by the participle *planned*, meaning they insert the premodifier in Spanish and the postmodifier in English.

Speakers from the youngest group, A1, also renegotiate language use when referring to football, a leisure activity that Gibraltarans sometimes pursue in Spain. The prestige and greater number of Spanish football clubs sometimes see young Gibraltarans cross the border and join a team in Spain. However, I did not find the reference to the term *football* itself in Spanish, as in B1, but rather an increase in alternations between both languages and Spanish insertions with interactive and discursive purpose (see Example 13).

- (13) DAN: We used to play in the same football club when we were small (0.5)
y ahora somos amigos (1.6) *ehm ¿qué más?*

LYAN: that's how we really got to know each other *¿no?*

DAN: yeah

MIKE: yeah yeah

DAN: football (0.5) starts with football (.) our parents knew each other (.)
and *ya::* (.) we started going out with each other more (0.5) so in school

LYAN: *y aquí estamos en* the same class

DAN: yeah at school we're all in the class as well

DAN: We used to play in the same football club when we were small (0.5)
and now we are friends (1.6) ehm what else?

LYAN: that's how we really got to know each other isn't it?

DAN: yeah

MIKE: yeah yeah

DAN: football (0.5) starts with football (.) our parents knew each other (.)
and then(.) we started going out with each other more (0.5) so in school

LYAN: and here we are in the same class

DAN: yeah at school we're all in the class as well

When discussing how they got to know each other, playing football in the same football club, speakers from group A1 still use English as the main language of conversation. However, switches into Spanish appear as peripheral elements for discursive and interactive purposes, such as closing an idea or showing a result in the case of *y ahora somos amigos* 'and now we are friends', asking participants for more information like *¿qué más?* 'what else?', or initiating a turn and concluding with *y aquí estamos* 'and here we are'. It seems that here, the function of the switch goes beyond the topic and can be better understood as a contextualisation clue. It is also interesting to observe how Dan omits the subject pronoun *it* before the verb *starts* in his third intervention.

To sum up, code-switching and Spanish elements are more present in some parts of the conversation in which speakers are discussing familiar and informal topics such as holidays, funny anecdotes, or personal experiences. More formal topics such as the British royal family, politics, or health (e.g., the COVID-19 pandemic) are commonly discussed in English, although code-switching is also observed when describing physical appearance in the pictures; and when discussing common politics and problems with Spain, or people's behaviour during the COVID-19 pandemic. However, even though some topics may trigger the use of insertions, this is not so clear for the speakers in my corpus. It seems that

alternation of languages and their sequentiality is much better understood by analysing the functions of the switch and studying interactions with a sequential, interactive, and conversational approach. This idea becomes evident when observing the use of code-switching in more formal topics such as education, job opportunities, or politics. Despite the use of English lexical elements to refer to the school system, such as *class*, *classroom*, or *middle school*, speakers in group BA1 and B1 constantly use markers and expressions in Spanish and alternate languages when talking about their experiences in school, making it difficult to define the matrix language in some of their interventions. As such, even though I expected to find that English was the matrix language of all interventions in younger participants, for a small number, this is clearly not the case. Further research needs to be done to account for individual and group variation. A similar pattern emerges in Example (14) below.

- (14) AKRAM: *¿Qué te digo man?* We've been friends really since primary school (0.6) *o antes, ¿no?* the nursery more or less

MOHAMED: *hombre yeah* (.) *guardería* (.) *ahí está* (0.7) St Marie's first school eh?

AKRAM: *ahí estamos mismo class de todo*

MOHAMED: *de to' la vida como dirían en yanito ¿no? de to' la vida bro*

AKRAM: *ahí estamos* (1.3) *después fuimos a Bishop*

MOHAMED: *ahí está*

AKRAM: secondary school

MOHAMED: middle school yeah

AKRAM: What can I tell you man? We've been friends really since primary school (0.6) or before haven't we? The nursery more or less

MOHAMED: of course yeah (.) nursery (.) there you go (0.7) St. Marie's first school eh?

AKRAM: there we are same class everything

MOHAMED: since forever as they would say in Yanito right? Since forever bro

AKRAM: there we are (1.3) then we went to Bishop

MOHAMED: there you are

AKRAM: secondary school

MOHAMED: middle school yeah

Rather than the topic, it is the level of formality required which seems to play a role in the selection of languages here. Interestingly, speakers constantly switch

to Spanish in a peripheral position for interactive and discursive purposes by using clusters of discourse markers or fixed expressions such as *¿qué te digo?* ‘what can I say?’, *ahí estamos* and *ahí está* ‘there we/you are’. This is also observable, for instance, in the extended use of discourse markers such as *bueno* ‘well’, a common strategy among speakers in Gibraltar as a way of making a conversation more informal, relaxed, and friendly. The use of *bueno* as a polite and friendly strategy to shift the topic of the conversation appears again in (15).

- (15) EMMA: *Bueno* guys (.) ((reading the guidelines)) “your relationship and friendship” (.) I did not have a:: choice (.) in meeting you (.) it was just forced

OLIVIA: (could be worse)

ISABELA: it was just forced

OLIVIA: wait can you hear me *bien?* (2.3) *¿se escucha?*

EMMA: yes

ISABELA: yeah

OLIVIA: okay *vale* cool

EMMA: *sí se escucha bien*

EMMA: Well guys (.) ((reading the guidelines)) “your relationship and friendship” (.) I did not have a:: choice (.) in meeting you (.) it was just forced

OLIVIA: (could be worse)

ISABELA: it was just forced

OLIVIA: wait can you hear me well? (2.3) can you hear me?

EMMA: yes

ISABELA: yeah

OLIVIA: okay okay cool

EMMA: yes we can hear you well

Speakers make use of the discourse marker *bueno* to change the topic of the conversation, but it also serves as a strategy to take their turn and add ideas. The direct questions like *¿se escucha?* ‘can you hear me?’ are good examples of the interactive and discursive purposes of the switch. Rather than the topic itself, it seems interesting to further study language choice and code-switching related to emotions and feelings (Pavlenko 2007; Dewaele 2010). The examples presented above refer to a variety of emotional themes for the participants; early memories, memories of school, hobbies, and hard times such as the COVID-19 pandemic. Despite the different contents of the conversation, examples show a common link: the emotional charge. Several studies (see Lantto 2014; Acuña Ferreira 2017)

advocate the need to consider the affecting function of code-switching and point to the use of L1 and L2 to express various emotions.

5.3 Turn-taking strategies

A focus-group methodology is crucial for the study of interactions and turn-taking strategies. To finish this section on the functionality and the discursive value of code-switching, I focus on how speakers initiate or finish their interventions. As I mentioned in the previous section, the discourse marker *bueno* is a recurrent strategy used by speakers to intervene in the conversation with the aim of finishing a topic or an activity and redirect the discussion; however, it also serves as a common marker for turn-taking and turn-yielding management.

Furthermore, different strategies are deployed by speakers to intervene and organise their discourse, as well as to request, compel, or encourage the other participants to contribute to the discourse. Spanish discourse markers and question tags often appear and serve this purpose. In particular, I exemplify here the use of the well-known request marker *¿no?* and reflect on the use of *bueno* as a turn-taking device with similar functions. Following this, I briefly mention the case of *mira* and the appearance of more complex and combined markers for turn-management purposes.

First, I would like to highlight the widespread use of the Spanish question tag *¿no?* as a turn-giving device and a request for agreement. This marker has already been described by Moyer (2000: 493) as a constant element expressing a “yes-no request” and a “request of agreement”. Goria (2017: 445) also finds multiple examples of *¿no?* as a pragmatic marker used to request agreement, but also stresses the need to study the question tag more extensively and points to the construction of this marker as a turn-giving device and a transition between argumentations and ideas. In my corpus, the use of *¿no?* is recurrent across all groups in both functions, but it always has an interactive component, making other participants contribute to the discourse, as in the following two examples from group B2.

- (16) OLIVIA: I think going on holidays together *¿no?* it's just [such good memories]

ISABELA: [I remember yeah]

OLIVIA: I think going on holidays together right? It's just [such good memories]

ISABELA: [I remember yeah]

- (17) OLIVIA: No but yes just maybe being over there (.) *¿no?* like together
ISABELA: good time (.) [yeah definitely]
EMMA: [it helped]

OLIVIA: No but yes maybe being over there (.) right? Like together
ISABELA: good time (.) [yeah definitely]
EMMA: [it helped]

The use and functionality of the marker *¿no?* as a turn-management strategy in Gibraltarian speech has already been studied by the above-mentioned authors. However, the use of *bueno* as a discourse or pragmatic marker with a similar functionality in interactions has not been addressed yet in the bilingual context of Gibraltar. As I mentioned before in Section 4, the use of this marker is recurrent across all groups. Speakers often use *bueno* to change the topic, but also to initiate their turns, changing the dynamic of the conversation by becoming the sender of the information or giving the conversation a new direction (18).

- (18) LYAN: I have to play with you one day
DAN: don't ((laugh))
LYAN: ((laugh))
DAN: *bueno*: I think we've spoken enough yeah?

LYAN: I have to play with you one day
DAN: don't ((laugh))
LYAN: ((laugh))
DAN: well I think we've spoken enough yeah?

In Example (18), Dan decides to take a turn in order to finish the conversation. Here, the marker serves as a strategy to initiate the turn and conclude the conversation or finish the task. A similar strategy is used by Emma in Example (19) below. In this case, the participants want to move on, but they seem tired. Emma uses the marker *bueno* to take a turn after a moment of silence and to move forward with the conversation. In this same intervention, the marker is used again to reaffirm that she is still going to hold her turn and introduce her own opinion, *My summer was a bit corta rollo* 'boring, restricted', despite having asked a question that was in the guidelines.

- (19) ISABELA: So let's move on (0.5)
EMMA: *bueno* the holidays ((reading the guidelines)) "how was your summer?" (.) *bueno* my summer was a bit *corta rollo* let's be honest

ISABELA: So let's move on (0.5)

EMMA: well the holidays ((reading the guidelines)) "how was your summer?" (.) well my summer was a bit boring/restricted let's be honest

Another marker that appears in group C1 is the attention-getter *mira* 'look', but its appearance is much less common in the other groups. This marker has previously been studied by Gorla (2017: 447) as an attention-getter and a strategy to hold a turn (see Example 20).

(20) SOPHIA: No we weren't (.) I think maybe Helen ¿*no*? were you in my class?

HELEN: yeah I was with e::h Mr Doel ¿*no*?

LAURA: I wasn't

SOPHIA: yeah ((laugh)) Doel ((laugh)) *mira* I'm just gonna quickly have a look on the questions (...)

SOPHIA: No we weren't (.) I think maybe Helen were you? Were you in my class?

HELEN: yeah I was with e::h Mr Doel right?

LAURA: I wasn't

SOPHIA: yeah ((laugh)) Doel ((laugh)) look I'm just gonna quickly have a look on the questions (...)

In group C1 and in group BA1, this function seems to be also fulfilled by more complex attention-getting devices and expressions such as *escucha* or *escúchame* 'listen' or 'listen to me', always used in second person singular regardless of the number of participants. The use of 'look', on the other hand, seems to be replaced. Further studies on the use of pragmatic markers and response tags need to be conducted to get a better understanding of the functionality and structure of code-switching in interactions.

6 Conclusion

The data presented in this chapter show a variety of structural and functional switches to both Spanish and English and reveal the complexity of Yanito used by younger generations. I observed that different code-switching structures are present in the conversations analysed, but not always in a homogeneous manner throughout the conversation. The context and the audience are crucial in the selection of strategies, but the formality required by the topic and the implicit emotions and feelings in the discussion may also change the entire dynamic of

the conversation. Even though we may be facing a generation of speakers who are not that proficient in Spanish, speakers show a repertoire of code-switching structures and functions to interact with their friends and family members.

From a structural perspective, I analysed three patterns proposed by Muysken (2000; 2013) and found that nominal insertions and backflagging – or, in this case, the use of Spanish discourse markers – are the most consistent and systematic strategies. Despite variability between age groups and individual variation within the focus groups (that could be based on language preferences and exposure to different languages), speakers from all focus groups employ Spanish elements in their discourse, especially discourse and interactive markers such as *escucha*, *oye*, *mira*, or *bueno* and vocatives like *tío* or *hermano*. In particular, I remarked on the use of fixed and vernacular expressions, such as *ahí está*, *esa es la cosa* and *todo ese rollo*, and a preference for the location of the switch in a peripheral position. The complexity of the expressions employed by speakers makes it difficult to classify them in fixed categories. Furthermore, the broad use and functionality of these expressions – together with changes in the direction of the switch – at different stages of the conversation suggest that they can be defined within various structural and functional categories. This points to the need to review definitions by considering the functionality of code-switching structures in the speech of Gibraltarian young adults.

Regarding function, I explored code-switching as a tool for topic-management, language negotiation, and turn-taking, and stated that traditional models of ‘predicting’ code-switching are also not consistent in Yanito. Regarding topic-management, my results coincide with those of Auer (1995), which account for the complexity of code-switching and the difficulty in defining specific situations or moments in which code-switching is going to appear, and advocate the need to study structure and functionality of code-switching from a more interactional and conversational perspective (Weston 2013; Gorla 2017; 2021). However, a low level of formality in the conversation and the emotional charge of the participants seems to play a role in the increase of Spanish insertions and switches. With respect to language negotiation, speakers negotiate and renegotiate their language use throughout the entire conversation and show a high degree of accommodation for other participants’ linguistic preferences. Here, code-switching serves as a tool to allow speakers to accommodate another speaker for a particular reason, as for instance, agreeing on something. Finally, code-switching also seems to play an important role in turn-taking management, allowing speakers to react to something (*sí* or *ahí está*), change the topic (*bueno*), or catch the attention of speakers (*mira* or *escucha*). The use of Spanish elements and the combination of markers for turn-taking purposes also supports the appearance of Spanish structures and switches as peripheral elements in conversations.

As an overall conclusion from what I have presented in this chapter, macro and sociolinguistic aspects of the speech community offer clues as to how a linguistic community operates. Furthermore, in order to analyse the meaning behind code-switching, it is important to combine a functional and structural analysis using interactional and conversational approaches. Further work has to be done to study the application of structural and functional categories of code-switching and to determine the current and future status of Yanito in Gibraltar. For this purpose, this study shows that analysing the speech of young adult Gibraltarians provides useful clues and avenues for future research.

Transcription conventions

- [text] Square brackets indicate an overlap
- (0.0) Numbers in round brackets indicate the amount of time elapsed in silence
- (.) A dot in parentheses indicates a tiny gap between utterances
- :: Colons indicate prolongation of the prior sound
- “text” Quotation marks indicate a quotation or reading of instructions
- (text) Parentheses indicate the transcriber’s difficulty in hearing/understanding what was said
- (...) Parentheses with ellipsis indicate the transcriber’s inability to hear/understand what was said
- ((text)) Double parentheses contain the transcriber’s descriptions or multimodal observations

Acknowledgements

Special acknowledgment goes to my first supervisor, Dr Sandra Schlumpf-Thurnherr, for her remarkable academic and professional advice and support throughout the COVID-19 pandemic. I would also like to express my gratitude to the *Seminar für Iberoromanistik* and the *Hermann Paul School of Linguistics* at the University of Basel for financing this project and my stays in Gibraltar. My gratitude also goes to the educational and cultural institutions in Gibraltar that supported me in the search for participants, and especially to the participants for making my research possible during difficult and uncertain times. The transcriptions of the focus groups were realised with the valuable assistance of Laura Renna and Ina Sarah Abel (University of Basel). Finally, I would also like to thank Johannes Ritter (University of Basel) for his invaluable support in the preparation and revision of this chapter.

References

- Acuña Ferreira, A. Virginia. 2017. Code-switching and emotions display in Spanish/Galician bilingual conversation. *Text & Talk* 37(1). 47–69.
- Auer, Peter. 1995. The pragmatics of code-switching: A sequential approach. In Lesley Milroy & Pieter Muysken (eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, 115–135. Cambridge: Cambridge University Press.
- Auer, Peter. 1998. *Code-switching in conversation: Language, interaction and identity*. London: Routledge.
- Auer, Peter. 2005. A postscript: Code-switching and social identity. *Journal of Pragmatics* 37(3). 403–410.
- Auer, Peter. 2014. Language mixing and language fusion: When bilingual talk becomes monolingual. In Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder & Achim Rabus (eds.), *Congruence in contact-induced language change: Language families, typological resemblance, and perceived similarity*, 294–334. Berlin: De Gruyter.
- Berruto, Gaetano. 2005. Dialect/standard convergence, mixing, and models of language contact: The case of Italy. In Frans Hinskens, Paul Kerswill & Peter Auer (eds.), *Dialect change: Convergence and divergence in European languages*, 81–95. Cambridge: Cambridge University Press.
- Chevasco, David. 2019. *Contemporary bilingualism: Llanito and language policy in Gibraltar: A study*. Cádiz: Editorial UCA.
- Dewaele, Jean-Marc. 2010. *Emotions in multiple languages*. London: Palgrave Macmillan.
- Durante, Charles. 2019. Charles Durante: Gibraltarian language and identity (I). *ReachExtra*. <https://reachextra.com/charles-durante-gibraltarian-language-and-identity-i/> (21 February, 2022).
- Feijóo Rodríguez, Sonia. 2015. “Somos más British que los British”: *Gibraltar: Code-Switching im Dienst der Identitätskonstruktion*. Freiburg: Albert-Ludwigs-Universität, Universitätsbibliothek. (Doctoral dissertation).
- Fishman, Joshua A. 1967. Bilingualism with and without diglossia; Diglossia with and without bilingualism. *Journal of Social Issues* 23(2). 29–38.
- Gold, Peter. 2005. *Gibraltar: British or Spanish?* London: Routledge.
- Goria, Eugenio. 2017. Functional markers in llanito code-switching. In Chiara Fedriani & Andrea Sansó (eds.), *Pragmatic markers, discourse markers and modal particles: New perspectives*, 439–457. Amsterdam: John Benjamins Publishing Company.

- Goria, Eugenio. 2021. Complex items and units in extra-sentential code switching: Spanish and English in Gibraltar. *Journal of Language Contact* 13(3). 540–572.
- Grocott, Chris & Gareth Stockey. 2012. *Gibraltar: A modern history*. Cardiff: University of Wales Press.
- Gumperz, John J. 1977. The sociolinguistic significance of conversational code-switching. *RELC Journal* 8(2). 1–34.
- Kellermann, Anja. 2001. *A new New English: Language, politics, and identity in Gibraltar*. Heidelberg. (Doctoral dissertation).
- Labov, William. 1984. Field methods of the project on linguistic change and variation. In John Baugh & Joel Sherzer (eds.), *Language in use: Readings in sociolinguistics*, 28–66. Englewood Cliffs, NJ: Prentice Hall.
- Lantto, Hanna. 2014. Code-switching, swearing and slang: The colloquial register of Basque in Greater Bilbao. *International Journal of Bilingualism* 18(6). 633–648.
- Levey, David. 2008. *Language change and variation in Gibraltar*. Amsterdam: John Benjamins Publishing Company.
- Levey, David. 2015. Gibraltar English. In Daniel Schreier, Edgar W. Schneider, Jeffrey P. Williams & Peter Trudgill (eds.), *Further studies in the lesser-known varieties of English*, 51–69. Cambridge: Cambridge University Press.
- Lipski, John M. 1986. Sobre el bilingüismo anglo-hispánico en Gibraltar. *Neu-philologische Mitteilungen* 87(3). 414–427.
- López de Ayala, Ignacio. 1782. *Historia de Gibraltar*. En Madrid: por don Antonio de Sancha.
- Loureiro-Porto, Lucía & Cristina Suárez-Gómez. 2017. Language contact in Gibraltar English: A pilot study with ICE-GBR. *Alicante Journal of English Studies / Revista Alicantina de Estudios Ingleses* (30). 93–119.
- Mariscal Ríos, Alicia. 2014. Consecuencias de la evacuación de la población civil de Gibraltar durante la Segunda Guerra Mundial en la situación sociolingüística y educativa actual. *Lengua y migración/Language and Migration* 6(1). 67–79.
- Moyer, Melissa G. 1993. *Analysis of code-switching in Gibraltar*. Universitat Autònoma de Barcelona. (Doctoral dissertation). <http://www.tdx.cat/handle/10803/4918> (15 January, 2022).
- Moyer, Melissa G. 1998. Bilingual conversation strategies in Gibraltar. In Peter Auer (ed.), *Code-switching in conversation: Language, interaction and identity*, 215–234. London: Routledge.
- Moyer, Melissa G. 2000. Negotiating agreement and disagreement in Spanish-English bilingual conversations with *no*. *International Journal of Bilingualism* 4(4). 485–504.

- Moyer, Melissa G. 2002. Pieter Muysken, Bilingual speech: A typology of code-mixing. Cambridge: Cambridge University Press, 2000. Pp. xvi, 306. Hb \$ 59.95. *Language in Society* 31(4). 621–624.
- Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.
- Muysken, Pieter. 2013. Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16(4). 709–730.
- Myers-Scotton, Carol. 2002. Review of Bilingual speech, a typology of code-mixing. *Language* 78(2). 330–333.
- Pavlenko, Aneta. 2007. *Emotions and multilingualism*. Cambridge: Cambridge University Press.
- Poplack, Shana. 1980. Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a typology of code-switching. *Linguistics: An Interdisciplinary Journal of the Language Sciences* 18(7-8). 581–618.
- Rodríguez García, Marta & Eugenio Goría. 2021. *The role of pragmatic markers in an emerging mixed code: The case of Gibraltar*. Panel contribution, 17th International Pragmatics Conference. Winterthur (Switzerland).
- Suárez-Gómez, Cristina. 2020. English in Gibraltar: Applying the EIF model to English in non-postcolonial overseas territories. In Sarah Buschfeld & Alexander Kautzsch (eds.), *Modelling World Englishes*, 347–370. Edinburgh: Edinburgh University Press.
- Weston, Daniel A. 2013. Code-switching variation in Gibraltar. *International Journal of Bilingualism* 17(1). 3–22.

Part II

Language contact and group identity

Chapter 4

Language change and stance in a remote Mennonite community in Canada

Miriam Neuhausen

University of Freiburg

This study investigates Canadian Raising patterns in the Pennsylvania German English speech of nine Old Order Mennonites in southern Ontario with a focus on linguistic context and stance-taking. Although this community is traditionally isolated and resistant to change, some speakers are now increasingly in contact with the local English-speaking community and have begun to participate in the ongoing shift towards Canadian Raising, a process that is largely complete in the wider Canadian English speech community. The extent to which linguistic resources are available to the speakers and used for social work, such as stances, hinges on the individual degree of contact with the English community. The speakers deal differently with the new vowel; /aʊ/-raising, imbued with the social meaning of Canadianness, might be adopted or avoided, while /aɪ/-raising, below the speakers' social awareness, indicates contact with the linguistic resources. The context of this study is perfectly suited for understanding local and translocal dynamics in a dialect contact situation, particularly in relation to stylistic practices in Third Wave sociolinguistics.

1 Introduction

The Old Order Mennonites (OOM) comprise a religious community of Swiss-German origin (Burridge 1998: 72) who reject modern technology, including the internet, mobile phones, and cars (Epp 2012: 39–41). They have resisted not only social change but also linguistic change to an extent; unlike the majority of immigrant groups in North America, they have maintained Pennsylvania German (PG) as L1 for nearly the past 400 years (Burridge 2002: 203). Similarly, their L2



English has also been comparably resistant to the surrounding language changes and still displays old features of Canadian English, such as the lack of Canadian Raising. However, due to increasing exposure to the English-speaking community, some OOMs have begun to participate in the ongoing change towards Canadian Raising, a process that is already complete in the local mainstream English community.

Research on language and identity has shown that identity plays a major role in linguistic variation (Eckert 2000, Johnstone et al. 2006, Nycz 2018). Playing with (context-dependent) social meaning, speakers can use linguistic variants to create locally meaningful identities (Podesva 2007). In choosing between non-raised and raised vowels, OOMs can create Pennsylvania German identities that linguistically distance them from local mainstream speakers or show linguistic integration with mainstream society.

While linguistic research on OOMs in Canada is mostly restricted to PG (Richter 1969, Burridge 1992), the majority of studies conducted on Pennsylvania German English (PGE) is based on communities in the US (Huffines 1984, 1986, Kopp 1997, 1999) – and only anecdotally in some cases (Springer 1980, Shields 1987). Furthermore, the existing literature on both PG and PGE is fairly dated, with the only exception being Anderson's (2011) PhD thesis discussing dialect contact and salience in PGE.

Drawing on six sociolinguistic interviews with nine OOMs, I investigate the linguistic context of Canadian Raising, before exploring stance – both quantitatively and qualitatively – as a possible factor accounting for some of the observed linguistic variation. It should be noted that the nine speakers do not have the same amount of contact with English; as the language change of Canadian Raising is still in its incipient stages in the OOM community, a high degree of inter- and intra-speaker variation can be expected.

2 The Old Order Mennonites

In previous research, members of religiously conservative communities have been shown to produce sociophonetically different speech from their secular neighbours. For instance, a Mormon community in Alberta displayed less Canadian Raising (Meechan 1999) and /æ/-raising (Rosen & Skriver 2015) than their immediate secular neighbours. And even within a Mormon community there may be differences: Baker-Smemoe & Bowie (2015) report significant linguistic differences based on how active Mormons in Utah were; inactive members fronted pre-nasal /ʌ/ significantly less than active members.

Unlike the vast majority of immigrants, the OOMs have maintained their L1 PG since the early eighteenth century (Draper 2010: 216), when they left Europe for freedom of religion, which was promised in Pennsylvania (Frantz 2017: 131–132). Despite its Swiss German roots, PG is based on Palatinate German and shaped by its long-term close contact with English. The Anabaptists likely originally spoke Upper German and Swiss German (Raith 1996: 317). When they fled Zurich and were staying in the Palatinate, they shifted to Palatinate German within one generation (Gratz & Geiser 1973 ctd. in Raith 1996: 317). The shift to Palatinate German thus took place before the migration to Pennsylvania and provided the basis for what would later become known as PG (Raith 1996: 316). Today, members of the community are usually bilingual and acquire the two languages successively – PG at home as L1 and English in school as L2 (BurrIDGE 1998: 85–86).

Distinct views prevail in the community concerning PG and PGE. Speakers frequently comment on linguistic variation in PG, for example the increase of English words, without evaluating it. By contrast, strong prescriptivism prevails concerning English (BurrIDGE 1998: 85). For example, during fieldwork, I observed that the speakers were aware of the fact that their English deviated from their secular neighbours' and even corrected each other's usage of English, such as one speaker's production of *legacy* with [dʒ] instead of [g].

Even though the community attempts to sustain its isolation and restrict interactions with the secular world, numerous OOMs work in English-speaking domains, for example, local corner shops, quilt stores, and Canada's biggest farmers' market. With more than 200 vendors, the market is the largest year-round farmers' market in Canada (see also their website, <https://stjacobsmarket.com/about-us/>, for more information). Situated in Waterloo County, it advertises itself as being "home to the largest population of Old Order Mennonites in Canada" and therefore attracting local OOM farmers who travel to the market by horse and buggy. Every day, numerous tourists from Toronto and beyond visit the market and buy traditional Mennonite-made goods, such as quilts and cheese curds. As a consequence, while some Old Order Mennonites are in regular contact with English-speaking locals and tourists, others work in PG-speaking domains, for example, farms and domestic work, and have barely contact with English.

PGE may be referred to as a religion-based ethnolinguistic repertoire (Rosen & Skriver 2015: 110; Benor 2011: 142), since it is not only the cultural heritage that keeps the community together, but also – and in particular – religion, serving as a "source of ethnic regional differentiation" (Frantz 2017: 147) from mainstream society. The notion of linguistic repertoire (as opposed to variety) implies that speakers of a group do not behave uniformly but choose linguistic variants both consciously and unconsciously to perform identities (Benor 2011).

3 Previous research

3.1 Canadian Raising

Canadian Raising, first noted by Joos (1942), describes the raised onsets [ʌɪ] and [ʌʊ] of the vowels /aɪ/ and /aʊ/ prior to voiceless consonants and /t/-flaps (Dailey-O’Cain 1997, Rosenfelder 2007, Fruehwald 2008). Raising is, however, not always restricted to this environment and may also occur prior to voiced consonants (such as /r/) and nasals or word-finally. Examples of this are provided in case studies in Canada, as in Victoria, BC (Rosenfelder 2007), but also in the US, as in Ann Arbor (Dailey-O’Cain 1997) and Philadelphia (Fruehwald 2008).

The earliest attestations of Canadian Raising in Ontario can be traced back to the 1880s (Thomas 1991: 148, Chambers 2006: 107), 100 years after the first OOM settlers arrived there (Epp 2012: 17). Due to increasing contact with the English-speaking community, some OOMs now find themselves in the early stages of the language change and have begun to adopt Canadian Raising.

Canadian Raising is commonly linked to general Canadian English speech (Chambers 1989, Niedzielski 1999), despite its attestations in the US (Labov 1972a, Roberts 2016) and in the UK (Moore & Carter 2018). In particular the two lexical items *out* and *about* have become stereotypes to both American and Canadian speakers in the Labovian sense (Labov 1971: 200). These two words are often produced with (phonetically inaccurate) hyper-raised nuclei: “oot and aboot” (Nycz 2013: 50). Investigating Canadian speakers who had moved to the New York City region, Nycz (2013: 56–57) found that half of the 15 speakers raised in all raising contexts, while the other half raised only the vowels in *out* and *about*. By contrast, regarding /aɪ/-raising, there seems to be much less social awareness (Chambers 1989: 76).

Concerning the speech of the OOMs, it can be expected that the two lexical items *out* and *about* pattern differently from other lexical items in general and from /aɪ/-raising in particular. Despite the differing degrees of contact with the mainstream Canadian community, all nine OOMs maintain close ties with settlements in the US through regular visits, letter correspondence and Old Order journals like the *Brotherhood Journal*. Therefore, despite their restricted geographic mobility, it can be assumed that these speakers are aware of the two shibboleths *out* and *about*.

3.2 Stance and identity

When language variants become linked to social meaning, speakers can use them to create locally meaningful identities. Linguistic variants have multiple potential

social meanings that are constantly negotiated within the context of language, interaction, and sociocultural values (Du Bois 2007: 139). Once a linguistic variant has gained social meaning and is no longer an *indicator* but a *marker* in Labov's (1971: 192–193) sense, speakers can use it to do identity work. That way, a variant can become enregistered as indexing group membership (Eckert 2012: 94). For example, in Pittsburgh, a set of linguistic features indicating socioeconomic class was first linked to place and then associated with a “Pittsburghese” dialect (Johnstone et al. 2006). Similarly, a white American boy at a California high school called Brand One draws on features of African American Vernacular English because they index (black) masculinity (Buchholtz 1999).

The OOMs have different linguistic means to index otherness in PGE: Firstly, features that are archaic in Canadian English can be used to index separation from the mainstream community and secondly, L1 transfer from PG may index membership in a local (PG-speaking) Mennonite community. Importantly, however, Podesva (2007: 496) notes that the association evoked by a given linguistic variant remains “open to interpretation, on the part of both the linguist and [the interlocutor]”. In other words, it is not just a matter of the speaker interpreting a variable to use it for social work; their interlocutors (and the researchers) also need to be able to read and interpret it as such.

Analysing stylistic practices, such as stance, from a Third Wave perspective (cf. Eckert 2012) is essential to understanding how speakers use language to create social identities. Stance-taking occurs when speakers evaluate and position the object they are talking about and align themselves with regard to the object and listener (Du Bois 2007). Through the accumulation of these stance-taking acts, speakers create identities (Rauniomaa 2003 and Du Bois 2002 ctd. in Buchholtz & Hall 2005: 596) that are constantly negotiated and renegotiated.

There is no one way of conceptualising stance. Some linguists examine affective and epistemic dimensions (Gadanidis et al. 2021), while others include alignment (Barnes 2018) or explore investment, interlocutor positioning, and voicing (Bohmann & Ahlers 2021). Including alignment, affective stance, and topic, Nycz (2018) investigated Canadian Raising in the speech of mobile Canadian speakers residing in Washington, DC. She detected more raising, associated with Canada, when ambivalence or emotional distance from the US was expressed and less raising when closeness or positive affect was shown. Following Nycz's (2018) definition of stance, I evaluate its usefulness for the analysis of Canadian Raising in the speech of nine OOMs.

4 Methodology

4.1 Procedure

I conducted sociolinguistic interviews with OOMs based in Waterloo County, Ontario, between 2018 and 2019. The fieldwork took place within the framework of Sali Tagliamonte's Ontario Dialects Project (2003, 2007, 2010; cf. <http://ontariodialects.chass.utoronto.ca>), which documents linguistic variation in English across Ontario.

As a cultural outsider, I entered the community as a “friend of a friend” (Milroy 1980: 53). Residing outside the community, I spent five months with the community, familiarising myself with the culture and the languages. Even though we found common ground thanks to my German-European background, I represented a cultural outsider at all times.

Interviewing members of the OOM community who openly reject modern technology proved quite difficult at times, particularly as many OOMs felt extremely uncomfortable around the recording device. It was impossible to eliminate the “observer's paradox” (Labov 1972b: 113), but the range of informal topics resulted in comparably casual and emotionally coloured speech. Being the interviewer in all speech situations, I should note that I generally do not raise the diphthongs /aɪ/ and /aʊ/, which may or may not have affected the participants during the interview situation.

I used a Roland R-09 recording device and its integrated stereo microphone for the interviews. Recorded data were digitised at the sampling rate of 44.1 kHz and submitted for acoustic analysis. For the acoustic analysis, I measured the height of the vowel onset, i.e. the F_1 value at 20 per cent duration. Future work should include tracking the entire vowel trajectory, as done by Moore & Carter (2018). Fronting was not included in the analysis as an initial exploration of the data did not indicate any changes in F_2 .

Preparing the data for quantitative analysis, I took the following steps: Where necessary, I removed noise in Audacity (Ash et al. 2015) before normalising the sound.¹ I segmented and transcribed the interviews in ELAN (Max Planck Institute for Psycholinguistics 2019), before force-aligning and extracting the vowels in FAVE (Rosenfelder et al. 2014). I manually checked approximately 15 per cent of the relevant contexts for accuracy in the force-aligned PRAAT script. Where necessary, I modified the alignment and re-ran the vowel extraction. After the removal of unstressed vowels in R (R core team 2020), I removed tokens shorter

¹Normalising sound in Audacity does not change the sound quality but amplifies sound without cutting off sound waves.

than 80ms ($\sim 8.65\%$, $n = 67$ for /aʊ/ and $\sim 11.3\%$, $n = 301$ for /aɪ/; see also Nycz 2013: 55). As a next step, I removed outliers beyond three standard deviations for each vowel and formant individually ($\sim 1.13\%$, $n = 8$ for /aʊ/, and $\sim 1.81\%$, $n = 44$ for /aɪ/). After this, I was left with a data set of 700 /aʊ/ tokens and 2,378 /aɪ/ tokens for the quantitative analysis. I then normalised the vowels based on the vowel-extrinsic and speaker-intrinsic Lobanov method in R.²

For the quantitative analysis, I manually coded for /t/-flapping and for stance; /t/-flapping was coded both acoustically and auditorily and stance was defined by topic and alignment (cf. subsection 4.3). I then matched the codes with the measured vowels using Python (Python software foundation 2019). In R, I categorised following context into raising contexts (voiceless sounds and /t/-flaps), non-raising contexts (voiced sounds and pauses), and all cases of *out* and *about* (cf. Chambers 1973, Moreton & Thomas 2007, Sadlier-Brown 2012, Nycz 2013: 115). Preceding segments were grouped into the different manners of articulation and word-initial contexts. Syllable type was manually coded per word as heterosyllabic or tautosyllabic. For the qualitative analysis, raising was determined auditorily. For reasons of simplicity, the auditory analysis is binary (as opposed to the quantitative approach).

4.2 Speakers

This data set comprises nine speakers in six interview situations, totalling 8.25 hours of speech. Each of these interviews covers 1.25 to 1.75 hours of speech, of which approximately three quarters represent PGE and one quarter PG. As the analysis is restricted to snippets of these speakers' speech, patterns emerging from the acoustic analysis should be considered suggestive rather than representative of general patterns in the OOM community.

All nine speakers grew up in the Township of Woolwich or in Wellington County and have spent their entire lives in the area. The speakers constitute a homogeneous group in terms of the following variables:

- Religious affiliation: baptised Old Order Mennonites
- Ethnicity: white European Canadians
- L1: PG acquired at home
- L2: PGE acquired in school

²For more information on the recommended order of operations in sociophonetic analysis for the purpose of comparability, please consider Stanley (2022).

- Education: eight years of either public or parochial school³
- Non-mobility: transportation modes restricted to horse and buggy rides and the occasional ride in a car driven by a non-member

Language competence was not measured, as all speakers are fluent in both languages and can switch effortlessly between the languages. Two speakers are in their 20s (Chloe and Leah), four speakers are in their 40s and 50s (Ada, Elisa, Naomi, and Rachel), and the remaining three speakers in their 60s (Isaac, Katie, and Phoebe).⁴ The current data set covers eight women and one man. It was more difficult for me to recruit men as interview partners because they tend to have less contact with cultural outsiders – their lives are often restricted to PG-speaking domains. Another reason is gender segregation in the community; a young woman would easier interview other women than men.

All speakers in this data set are somewhat open to cultural outsiders as they agreed to being recorded. In a community that openly objects to modern technology, this already presupposes some openness towards and trust in cultural outsiders. I did not record speakers that were completely shielded from outside communities. Based on my impressionistic observation, these speakers may not raise at all; however, it does not seem likely that linguists will gain access to these speakers in the near future.

Despite their homogeneous sociodemographic background, the OOMs – as any other community – are a highly heterogeneous group of speakers. While the language change of Canadian Raising is already complete in the wider Canadian population, the OOMs have only begun to participate in the process. Thus, the nine OOMs display great inter-speaker variation which can be observed in Figure 1. The x-axis visualises the degree of raising, where lower numbers indicate more raising. Ada and Naomi feature the highest degree of raising, while Leah, Isaac, and Phoebe display the least. Both Ada and Naomi are independent, unmarried women and work in the English-speaking domain, while the three speakers with the smallest degree of raising – Isaac, Leah, and Phoebe – work in the PG-speaking space.

The y-axis visualises the speakers' respective standard deviation of raising and therefore shows how much speakers vary in their raising behaviour. Higher numbers indicate greater variation. The graph illustrates that Naomi, Ada, and Chloe,

³The parochial school system was established in the 1960s (Epp 2012: 41) and reinforced social and linguistic isolation. Prior to that, children from the OOM community went to public school taught by non-members and attended by 'English' children, i.e. children growing up in mainstream Canadian society. The types of school are not statistically significant in this analysis, but future research could explore the role of school in linguistic variation further.

⁴All names were anonymised.

plotted at the top, are among the speakers who shift the most between raising and non-raising. Strikingly, all three women are surrounded by English speakers in their everyday lives. By contrast, Isaac – the only male speaker in this data set – appears to follow a consistent pattern, not raising much in the first place and hardly varying this behaviour; he may thus avoid the innovative variant or not have full access to it (yet). The remaining speakers, Katie, Rachel, Elisa, Phoebe, and Leah, behave similarly in terms of how much they vary between their raised and non-raised variants.

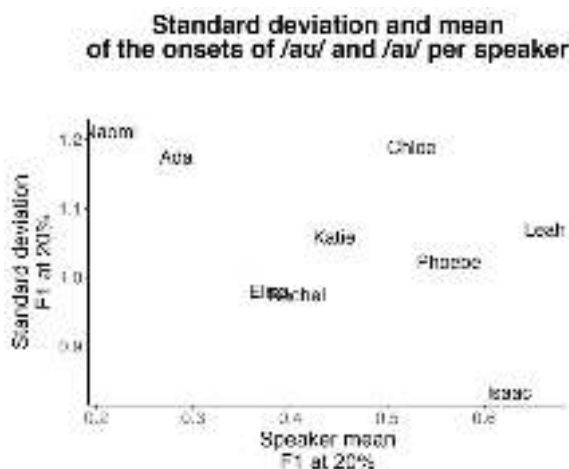


Figure 1: Inter-speaker variation of the onsets of /aɪ/ and /aʊ/ is displayed on the x-axis and intra-speaker variation on the y-axis. Higher numbers on the x-axis represent a lower degree of raising, i.e. speakers on the right raise less than speakers on the left. Higher numbers on the y-axis indicate more intra-speaker variation, i.e. speakers on the top vary greatly in their raising behaviour, while speakers on the bottom do not vary much.

4.3 Stance

The coding process was text- and content-based; any influence of phonetic context was avoided. Following Nycz (2018), I coded for topic and affective stance/alignment. The different codes are explained and examples of each are provided in the following.

Topic was divided into OOM-specific topics, mainstream topics, and neutral topics.⁵ The first include narratives and descriptions of community-specific norms

⁵If a stretch of speech represented multiple topics, I chose the one that was most relevant for the speaker's overall statement.

and customs, like the concept of a maid in (1). Maids are young girls who help other young families with child care and domestic work before they are allowed to work for wages at the age of 16. Mainstream topics cover subjects related to the (English-speaking) secular mainstream community, as in (2). Neutral was coded for when neither of the former two applied, as in (3).

- (1) My sister went out and helped another family, uhm, if they maybe have a whole bunch of small kids, that's what we usually do our first couple of years, 'til we're old enough to work away, 'til sixteen or whatever. (Rachel, OOM).⁶
- (2) It is interesting with the people we meet from all over the world, like you. They are for the history that we didn't get. (Katie, mainstream)
- (3) The teacher didn't like when we whispered, like we always, like sometimes was when she'd catch us but then she would get us to write out: "Let's not whisper! Let's not whisper!" (Naomi, neutral)

Speakers may express closeness to or distance from a given topic, which is captured by the alignment coding. Following Nycz (2018), I coded for both alignment and affect, which resulted in the following levels: aligned, non-aligned, positive, negative or neutral. The aligned stance expresses both solidarity and closeness, for example, expressions of belonging or fitting in, as in (4). Non-aligned is used for feelings of not belonging to or not agreeing with a community, as in (5). Here, Chloe distances herself from another community, another branch of Mennonites, by explaining that they do not share the same variety of German. Positive and negative affect covers the evaluation of a described object. An example of positive affect is provided in (6) and of negative affect in (7). Neutral alignment, illustrated in (8), represents stretches of speech where no explicit affective or alignment stance was used.

- (4) Although it's still a very pleasant life in our thinking, is a great opportunity to raise a family in a setting where they've got something to do. (Isaac, aligned)
- (5) Or there's Mexican Mennonites, [...] their German is different from ours. (Chloe, non-aligned)

⁶To increase legibility, I inserted commas and full stops for pauses. Words indicated in single quotation marks in square brackets represent a translation; an ellipsis inserted in square brackets indicates missing parts of speech. Extracts may be shortened and restricted to relevant context.

- (6) They'd know that's the icing on the cake, so that used to be fun to gather eggs and then feed them! (Phoebe, positive)
- (7) And grammar was my worst subject. I never understood why we had to analyse sentences. I did not see any sense in that, I didn't understand it. [...] I quit analysing sentences when I quit school. (Isaac, negative)
- (8) A few times that we didn't go to school that I remember of, like where it snowed so bad that– or it was extremely cold or something that they cancelled school. (Leah, neutral)

The data set covers 475 aligned codes, 463 non-aligned codes, 225 positive codes, 35 negative codes, and 1,880 neutral codes. Due to the comparably low number of negative alignment levels, they were excluded from the statistical analysis. After the removal of outliers and the negative alignment codes, I was left with 1,085 vowels coded for OOM topics, 481 for mainstream topics, and 1,477 vowels for neutral topics. Translated into vowels, this left me with 694 tokens of /aʊ/ and 2,349 tokens of /aɪ/.

For the analysis of topic and alignment, contact with English needs to be assessed for every speaker individually. As the OOM community finds itself in a long-term contact situation with ongoing linguistic changes, it is vital to understand to what extent individual speakers have contact with English and are open to cultural outsiders. Scholars have addressed the issue in different ways. Hazen (2000: 150–151), for example, draws a line between local- and expanded-identity speakers. Local-identity speakers are oriented towards the core of the community, while expanded-identity speakers are oriented towards other communities; this orientation shows both socially and linguistically. In the case of the OOMs, all speakers have somewhat extended identities, as they allowed me to record them and are thus necessarily open towards people beyond the community lines. Baker-Smemoe & Bowie (2015) distinguish between active and non-active members in a Mormon community, which shows linguistically. Active members, as opposed to inactive members, self-identified as participating at least weekly or nearly weekly in organised religious activities. This distinction is not useful for the present community, as all OOMs are active members. It becomes clear that the issue of which measures to use strongly hinges on local context. In lesser-researched communities, such relevant community-specific social factors may only emerge during ethnographic fieldwork (cf. Stanford 2009, Neuhausen 2023) and can only then be used to accurately define what contact with another language (or community) means in a given community.

I developed a community-specific scale attempting to quantify the speakers' individual degree of contact with English and openness to non-members. The following categories serve to help measure it:

- using English at least sometimes at the work place,
- working part-time/full-time in an English domain,
- having contact with English family members/neighbours/friends outside of work,
- seeking contact with cultural outsiders, and
- being an unmarried woman (who often leave the PG-speaking home farms to live in apartments in the English neighbourhood and/or who may run their own businesses with English customers).

Based on these categories, I calculate the individual degrees of contact with English as follows: The starting point for each speaker is score 0 and represents almost no contact with English; the higher the score, the more contact with English a speaker is predicted to have. For every category that pertains to a given speaker, a score of +1 is added. This measure is then transformed into categorical variables ranging from "(almost) no" contact to "regular" contact with English. According to this measure, Elisa and Leah have the least contact with English; Rachel, Isaac, Katie, and Chloe are grouped as speakers with some contact with English; Phoebe and Ada represent speakers with moderate contact; and Naomi is the speaker with the highest score, with regular contact with English (see also Figure 2). It is noteworthy that here, Naomi's status aligns with her vowels being the most strongly raised and most strongly varied, as can be seen in Figure 1. Despite the measure pointing to some speakers having "(almost) no" contact with English, it is important to keep in mind that all speakers have some contact with English as they all acquired English as L2 and live in a community that uses English as lingua franca with outside communities.

5 Statistical analysis

The statistical analysis includes both linguistic context and stance. For the linear mixed-effects model, the F_1 value was treated as an independent variable, with linguistic context and stance as dependent variables. Speaker and word were incorporated as random effects and vowel was fitted as random slope for speaker.

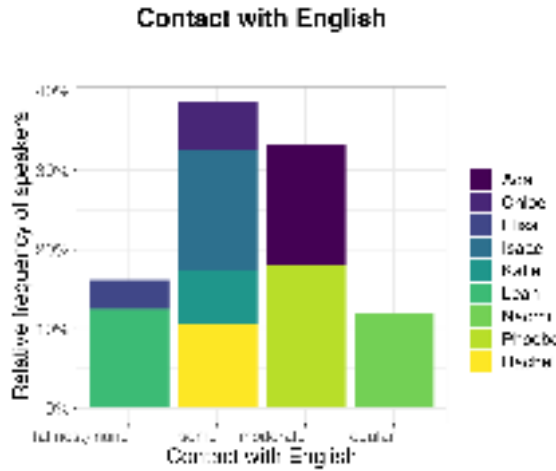


Figure 2: The categorised contact with English is visualised on the x-axis. Different colours indicate different speakers; the different width of each bar represents the relative frequency of vowels analysed per speaker.

In order to investigate the linguistic context of Canadian Raising in the speech of the OOMs, the following language-internal variables were included in the model: vowel, duration, following voicing context, preceding manner of articulation, and syllable type. Phonetic stress and age were not significant and not included in the model.

Concerning language-external factors, stance can be expected to strongly hinge on the degree of contact with English as the OOMs use English to varying degrees. Thus, I examined the interaction of contact with English and alignment and topic.

All categorical variables were sumcoded. The reference levels of the sumcoded variables were set to word-initial contexts for preceding manner of articulation and non-raising contexts for following voicing contexts, i.e. word-final and voiced contexts. Neutral codes were used as reference levels for alignment and topic and the least contact with English was used as reference level for contact with English.

Multicollinearity and heteroscedasticity were both tested for, using the performance package (Lüdtke et al. 2021). They are not a concern; for the same model without the interaction, the highest corrected VIF scores are at 1.95 for alignment and 1.94 for topic (see also Levshina 2015: 160). In the following, the results of the model are briefly discussed and visualised in form of predictions. The predictions based on the statistical model were calculated using the ggeffects (Lüdtke 2018)

and the lme4 packages (Bates et al. 2015). The p-values were calculated using the broom.mixed library (Bolker & Robinson 2022). The calculated R^2 value for the model (from the performance package) predicts that 30.3% of variance can be accounted for by the model.

Table 1: The mixed-effects model treats raising as an independent variable; linguistic context and stance – i.e. the interaction of contact, alignment, and topic – are treated as fixed effects. Speaker and word were incorporated as mixed effects, with vowel as random slope for speaker. Lower estimates displayed in Table 1 indicate more raising. Only significant interactions of stance were included in Table 1.

Mixed model predicting Canadian Raising in the speech of Old Order Mennonites							
<i>N</i> = 3,043 (<i>n</i> =694 for /aʊ/, <i>n</i> =2,349 for /aɪ/)							
term	estimate	standard error	t-value	p-value	95% confidence interval		signif.
					lower bound	upper bound	
(intercept)	0.1387	0.2135	0.6495	0.5162	-0.2804	0.5577	
Vowel: /aʊ/	0.2082	0.0403	5.1623	0.0000	0.1265	0.2898	***
Duration [in s]	3.5748	0.2612	13.6859	0.0000	3.0626	4.0870	***
<out> </about>	-0.3742	0.1373	-2.7260	0.0098	-0.6524	-0.0960	**
Raising context (voiceless & /t/-flaps)	0.1359	0.0735	1.8497	0.0707	-0.0119	0.2836	
Prec. plosive	-0.0633	0.0515	-1.2290	0.2197	-0.1645	0.0379	
Prec. fricative	0.2029	0.0760	2.6717	0.0080	0.0534	0.3525	**
Prec. vowel	0.0781	0.0773	1.0112	0.3120	-0.0734	0.2296	
Prec. nasal	-0.2537	0.0599	-4.2359	0.0000	-0.3714	-0.1360	***
Prec. rhotic	0.0634	0.0732	0.8654	0.3874	-0.0806	0.2073	
Prec. sibilant	-0.2353	0.0727	-3.2383	0.0013	-0.3779	-0.0972	**
Prec. lateral	-0.3819	0.0889	-4.2949	0.0000	-0.5567	-0.2071	***
Prec. approximant	0.0399	0.1246	0.3200	0.7494	-0.2064	0.2862	
Prec. affricate	0.0487	0.2362	0.2064	0.8366	-0.4161	0.5136	
Tautosyllabic syllable	-0.0968	0.0344	-2.8183	0.0055	-0.1647	-0.0290	**
[Moderate contact - aligned stance]	0.7997	0.3104	2.5760	0.0100	0.1910	1.4083	**
[Aligned stance - mainstream topic]	0.9271	0.4017	2.3079	0.0211	0.1394	1.7148	*
Moderate contact - aligned stance - mainstream topic	-1.4124	0.4059	-3.4796	0.0005	-2.2083	-0.6165	***
[Regular contact - aligned stance - mainstream topic]	2.1791	0.8324	2.6177	0.0089	0.5468	3.8113	**
Random effects structure: Random intercepts for participant and word; by-participant random slopes for vowel							

In the following, the model will be interpreted based on the predicted values of F_1 . The predicted values of F_1 are adjusted for the following linguistic context: /aʊ/ vowels; at a duration of 0.16 seconds; in the items *out* and *about*; with preceding plosives; in tautosyllabic contexts. The adjusted values for the following social factors are: speakers with moderate contact to English; producing non-aligned stances; and mainstream topics.

Regarding the vowel, less raising is predicted for /aʊ/ (predicted value for F_1 = 0.7) than for /aɪ/ (predicted value for F_1 = -0.4, p = 8.17e-6). This finding is interesting as /aɪ/ is arguably below the speakers' conscious level, while /aʊ/ may indicate Canadianness. While the spread of /aɪ/-raising indicates that Canadian

Raising has entered the community, the lack of /aʊ/-raising may suggest that speakers avoid the variant, which is socially imbued with Canadian identity outside their community (see also Figure 3). In other words, /aɪ/ may be more systematic in both raising and non-raising contexts, while for /aʊ/-raising, speakers may feel the need to mark the distinction between the raised and non-raised variant of /aʊ/ more strongly as it is linked to the social meaning of Canadianness. It should be noted, however, that more than three thirds of the data set cover tokens of /aɪ/ (77.19%, $n = 2,349$).

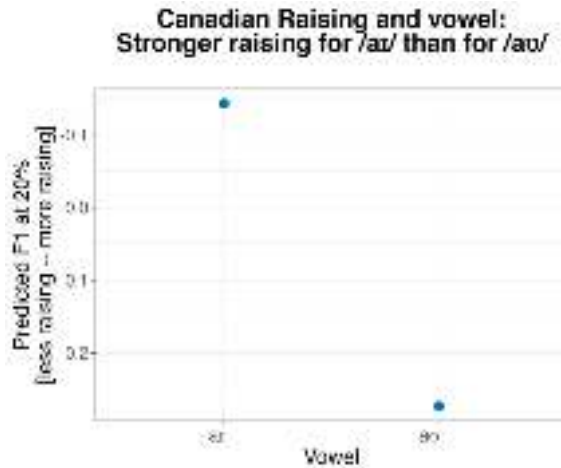


Figure 3: The two vowels /aɪ/ and /aʊ/ are indicated on the x-axis and the predicted amount of raising is plotted on the y-axis, where lower numbers indicate more raising. /aɪ/, plotted at the top, is predicted to feature more raising than /aʊ/, plotted at the bottom. This may suggest that speakers avoid the variant that carries the social meaning of Canadian identity.

The predicted probability score of raising with a vowel duration of 0.5 seconds is 1.49 (see Figure 4). It increases to 3.27 for a vowel duration of 1 second. In other words, the longer the duration, the less raising is predicted in these contexts ($p = 2.21e-41$). In American English, vowels tend to be longer when preceding voiced consonants (Kendall & Fridland 2021: 119). This aligns with the present findings, where longer vowels, assumably prior to voiced contexts and pauses, are predicted to feature less raising (see also Figure 4).

Following voicing context was grouped into raising and non-raising contexts. A third level was included describing all vowels in *out* and *about*. Nycz (2013) demonstrates that these words show different raising patterns in the speech of mobile Canadian speakers in the New York City region. The same pertains to the present data set: the words *out* and *about* are significantly more raised (0.27,

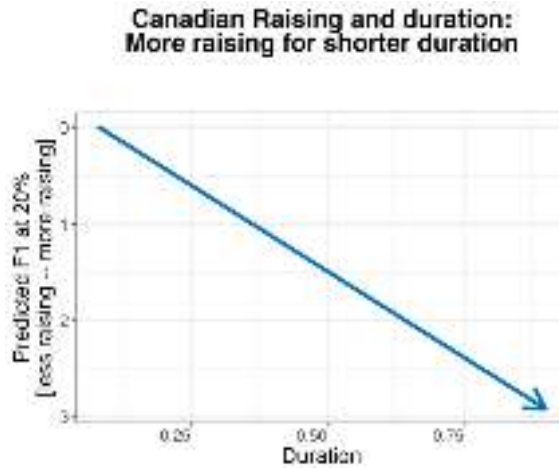


Figure 4: Duration (in seconds) is displayed on the x-axis, while predicted raising is indicated on the y-axis. Longer vowels are associated with less raising.

$p = 9.77e-3$) and predicted to feature more raising than vowels in raising contexts ($0.78, p = 7.07e-2$) and in non-raising contexts (0.89 ; see also Figure 5). This finding aligns with Nycz' study who found that some mobile Canadian speakers constantly raised while others only raised the vowels in *out* and *about*.⁷ Remember that /aʊ/ generally co-occurs with less raising than /aɪ/; yet, the two cultural shibboleths of *out* and *about* significantly co-occur with raising, which potentially indicates that the two words carry social meaning. This may also suggest that *out* and *about* accelerate the language change as shibboleths towards /aʊ/-raising or are selectively used for raising, while the remaining tokens of /aʊ/ lag behind.

In terms of preceding manner of articulation, laterals ($-0.5, p = 2.19e-5$), nasals ($0.08, p = 2.71e-5$), sibilants ($0.10, p = 1.25e-3$), and plosives ($0.27, p = 2.20e-1$) are predicted to favour raising in that order. Examples of preceding laterals are *line* and *realised*, of preceding nasals *mind* and *amount*, of preceding sibilants *outside* and *shout*, and of preceding plosives *kind* and *about*. These are followed by approximants ($0.38, p = 7.49e-1$, *while* and *otherwise*), affricates ($0.39, p = 8.37e-1$, *child* and *lunch hour*), rhotics ($0.40, p = 3.87e-1$, *right* and *around*), other vowels ($0.42, p = 3.12e-1$, *so I* and *go out*), and fricatives ($0.54, p = 8.01e-3$, *find* and *with-out*). The smallest degree of raising is predicted in word-initial vowels (*I* and *out*, see also Figure 6).

⁷Unlike in her data where *out* and *about* occurred in more than half of all occurrences, *out* and *about* only account for 31.56% of the /aʊ/ data set.

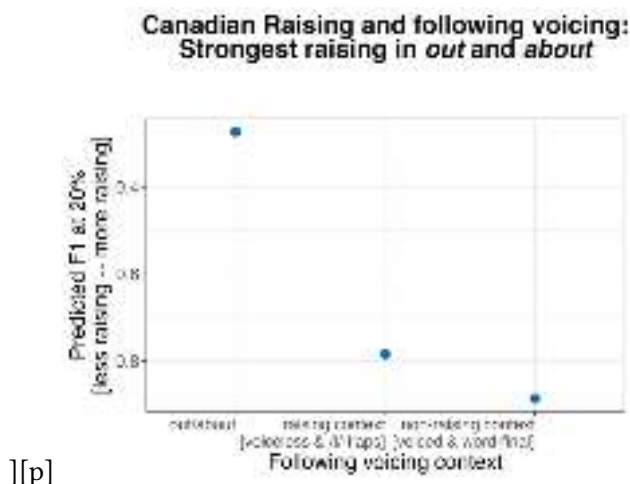


Figure 5: Following voicing context is displayed on the x-axis. By far the most raising is predicted for the words *out* and *about*. Vowels prior to voiceless contexts (including /t/-flaps), are linked to more raising than in non-raising contexts (following voiced sounds and pauses).

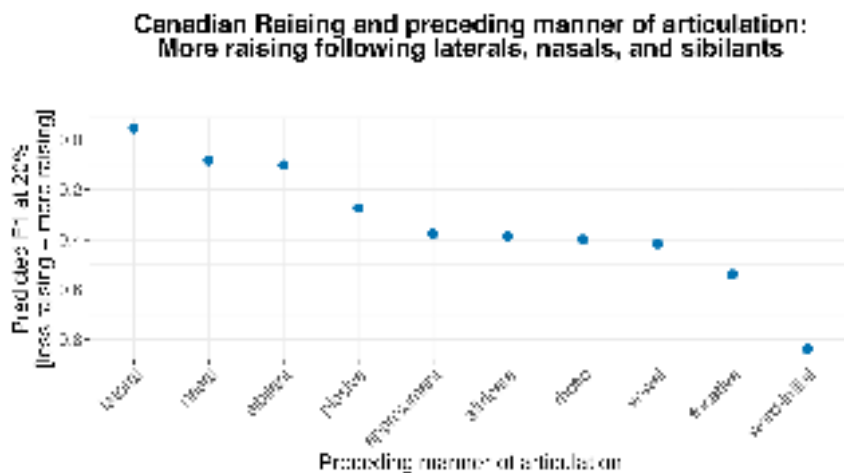


Figure 6: The x-axis represents the preceding manner of articulation. From left to right: Vowels following laterals, nasals, and sibilants are predicted to be the most raised, followed by plosives, approximants, affricates, rhotics, and other vowels in that order. Vowels following fricatives and word boundaries are predicted to be raised the least.

According to Chambers (1989: 79), tautosyllabic contexts favour Canadian Raising over heterosyllabic contexts. Investigating the distribution of raising across the two syllable types is worthwhile, as the community finds itself in the initial stages of the language change towards Canadian Raising and it cannot be assumed that all speakers in the present data set have access to the linguistic constraints. Supporting Chamber's statement for the OOM community, in the present data set, vowels in tautosyllabic contexts (0.7, $p = 5.47e-3$) are predicted to occur more often as raised tokens than in the latter (0.47; see also Figure 7). This finding may suggest that the OOMs indeed have access to (at least some of) the linguistic resources of Canadian Raising.

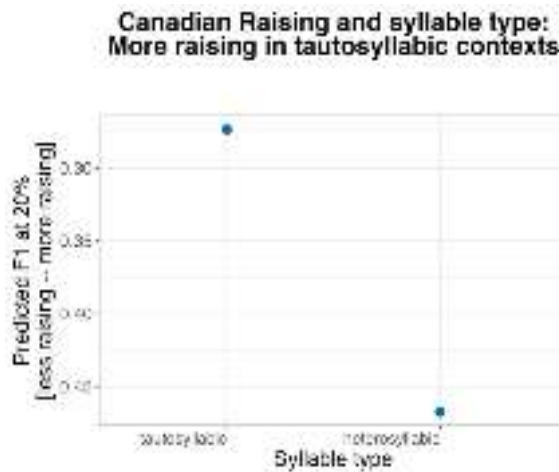


Figure 7: Heterosyllabic and tautosyllabic contexts are indicated on the x-axis. The graph supports previous literature on Canadian Raising, predicting raising to favor tautosyllabic contexts (Chambers 1989: 79).

Stance, defined as the interaction of contact with English, topic, and alignment, represents the only social factor in the model. For the analysis, I only consider the significant interaction with the highest amount of interaction levels with three levels. As one of the two significant interactions (regular contact–aligned stance–mainstream topic) only features four tokens, I only focus on the other significant three-level interaction: moderate contact–aligned stance–mainstream topic ($n = 66$). For speakers with moderate contact with English, the aligned stance is predicted to occur with more raising when covering mainstream topics (0.03, $p = 5.10e-4$, as opposed to 0.78, 0.55, and 1.26 for speakers with almost no contact, some contact, and regular contact in that order). This is also visualised in Figure 8. It is an interesting finding, as it suggests that rhetorical closeness may

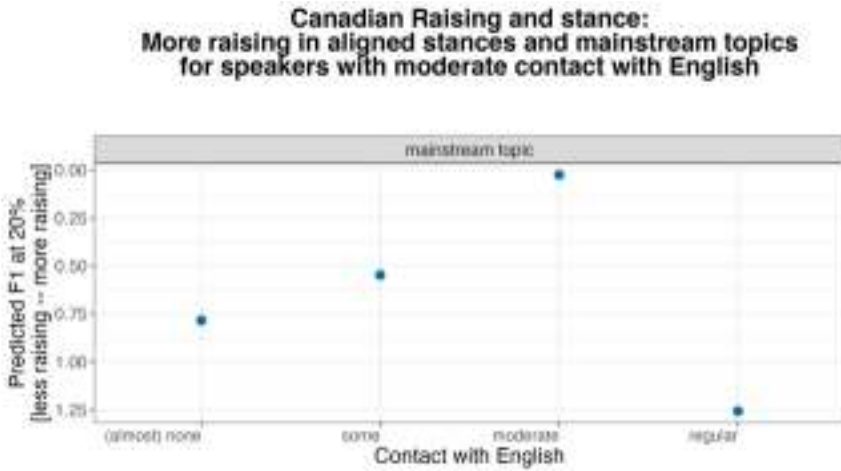


Figure 8: The combination of moderate contact, aligned stance, and mainstream topic (n=66) is predicted to feature raised vowels.

entail linguistic closeness for speakers with moderate contact with English. In other words, speakers who rhetorically align with mainstream topics may also converge linguistically. In the following section, the significant interaction of aligned stance, mainstream topic, and moderate contact with English is assessed qualitatively in the speech of Phoebe.

6 Phoebe

Phoebe, in her sixties, is active in both PG and English domains, but spends significantly more time in the former, where she plays an important social role. In terms of raising degree and intra-speaker variation, her behaviour clusters with the other speakers (with the exception of Naomi and Ada, cf. Figure 1).

Having grown up with English-speaking family members, she likely has full access to the linguistic constraints of Canadian Raising and appears confident when speaking English. Phoebe expresses strong sentiments, showing solidarity with and distancing herself from concepts linked to the English-speaking community. In (9), she emphasises that from a spiritual perspective, there is no difference between people, no matter what community they are affiliated with. In (10), she distances herself from church services in the English-speaking mainstream community, where sermons are (necessarily) monolingual and restricted to English only. In her view, having only one language to pray in, the mainstream

community misses out on some of the depth and meaningfulness that the German language can add in the bilingual context of the Mennonite church. Here, she distances herself from mainstream society and implicitly states that the Mennonites profit from bilingual sermons, as speakers can use the resources of more than one language to express their needs (and prayers).

- (9) So yeah, we're all just people anyhow, regardless what colour hats we wear or dresses or cars or vehicles we drive or – [...] we're all just people.
- (10) There's something missing, that added depth isn't there because they're confined to their use of their English vocabulary.

In the following, two stretches of speech produced by Phoebe will be analysed in more depth. Both address mainstream topics and represent aligned stances. Raised vowels are indicated in italics, non-raised vowels are underlined>. In utterance (11), Phoebe describes how her younger brother Elo acquired English by spending time with his cousin Jacob. Growing up in an English-only environment outside the OOM community, Jacob brought English into the games with his PG-speaking cousin. Phoebe expresses her solidarity with her English cousins and simultaneously distances herself from them by describing them as “English”: “we have good memories of our *English* cousins”. Solidarity is further expressed by her emphasis of how close the two boys were: Firstly, sharing the same birthday and year, they were called twins by their families; secondly, they spent a lot of time together (“they often came”); and thirdly, they were still children. Phoebe makes the point that children “get along well” even without speaking the same language. In the following stretch of speech, she expresses her adoration of how Elo started to learn English without effort—“without thinking”—thanks to his socialising with his English cousin. Here, Phoebe expresses admiration of how easily children adapt to new situations and how they profit from interacting with other (English) children.

Linguistically, in terms of /aʊ/, she does not raise in non-raising contexts, such as *our* and *how*. In raising contexts, she only raises the vowel in the shibboleth *out*, but not in *without*, which represents a raising context. By contrast, for /aɪ/, all non-raising contexts feature non-raised variants, i.e. *my* and *time*, and the only raising context features a raised vowel, i.e. *like*. This utterance may support what the quantitative findings suggest: /aʊ/ is not raised across the board but in *out* and *about*, while /aɪ/ may be consistently raised in raising contexts. Regarding /aɪ/, Phoebe conforms to the traditional raising contexts, potentially indicating that she has full access to the linguistic constraints.

- (11) We have good memories of our English cousins and that definitely, my youngest brother Elo had a twin with an English cousin and you know how children are; they don't need to know the same language to get along well, they just know by their actions and by the tone of voice. Uhm, they were both, he was four—they were both four, of course, 'cause they were twins—and after they left, they often came, but after they left this one time when he was four, my mother was making a cake—or was it, I'm not sure who was making the cake—but, uhm, we told him to bring their wooden spoon, like to stir it, and he got it out of the drawer and he said: “This?” That was his first English word that we know that he said without thinking, you know, it was just there because he had been interacting so much with Jacob, his cousin.

In (12), she describes how her mother always emphasised that there was no difference between people, whether they were Mennonites or not. Phoebe strongly agrees with her mother and adds that it may take longer for some people to figure this out, but they will do so eventually. Rhetorically, she aligns with members of mainstream society here as she does not make a difference between them and herself.

In terms of /aʊ/-raising, the first two tokens of /aʊ/, i.e. *out*, are not raised despite being in raising contexts. The third occurrence of *out* is audibly just a little bit raised. Concerning /aɪ/-raising, non-raising contexts feature non-raised vowels, i.e. *my*, *line*, and *find*, and raising contexts feature raised vowels, i.e. *life*. The only exception is represented by *why*, a raised vowel in a non-raising context.

- (12) And my mother, she's the one who always constantly was saying: “We're all just people.” And you know what? The bottom line is: we are all just people! Some people find it out sooner, some people find it out later in *life* but eventually there will come a time when you will find out and discover first hand: we are all just people! *Mir sin all jusht leit* [‘we are all just people’], no kidding! And she was good with that and she also said: “What happens to others can happen to us. *Why* do we think we're better than they are?” Mh. Amen. *Sell is wahr*. [‘That is the truth.’]

It is interesting that in this stretch of speech, in two out of three times, Phoebe does not raise the vowel in *out*. Following the quantitative findings, one would expect to see a strong tendency for raised *out* and *about*. This may suggest that Phoebe deviates from a more general raising pattern for *out* and *about* by avoiding to raise in these contexts. As she is one of three speakers who have moderate or regular contact with English, this raises the question whether the other

two speakers with a similar degree of contact behave similarly and whether increased contact with English can lead to the avoidance of raised *out* and *about*. If speakers are confident enough speaking English, they may not feel the urge to speak ‘proper’ English anymore, i.e. linguistically conform to their monolingual English neighbours, and use the raising of these vowels for identity work. Doing so, Phoebe may distance herself linguistically in these contexts. Moreover, while she conforms to the traditional raising contexts for /aɪ/ in (11), she over-raises /aɪ/ in one instance in (12). This may be a first indicator that she has access to the linguistic constraints of /aɪ/-raising but may manipulate both vowels /aʊ/ and /aɪ/ to make a statement, particularly as the over-raised vowel occurs in a rhetorical question summarising her statement. It should be noted that the qualitative analysis remains very superficial at this point and only provides a snapshot of the linguistic complexity that is going on in the data.

7 Discussion

According to the quantitative analysis, linguistic context mostly correlates with the expected raising behaviour. This applies to what is known about vowel duration and syllable type. Yet, the qualitative analysis also reveals that following voicing contexts do not consistently feature the raising pattern documented in previous studies. Maybe this cannot be expected in a contact situation with ongoing linguistic changes and speakers with varying degrees of contact with English.

Across the board, the vowels in *out* and *about* behave significantly differently from voicing contexts in the statistical model. It seems likely that the shibboleths *out* and *about* are salient to all speakers subject to the analysis. The speakers maintain close ties with their American settlements and all of them are in contact with English to varying degrees (even the most isolated ones still live in a community that is surrounded by English). Potentially, *out* and *about* accelerate the language change towards /aʊ/-raising as shibboleths. However, as a snapshot of Phoebe’s speech illustrates, not everybody raises *out* and *about* across the board. This is particularly interesting because Phoebe has moderate contact with English and it is safe to assume that she is aware of the social meaning imbued on these two words.

The quantitative analysis demonstrates that /aɪ/ is raised more than /aʊ/. The phenomenon of Canadian Raising provides a particularly interesting linguistic variable for the analysis of variation in a contact situation as the present speakers are likely aware of the social meaning attached to raised /aʌ/ (particularly in *out* and *about*), which can be used or avoided for identity work.

Being able to use innovative linguistic forms for identity work hinges on sufficient contact with the new linguistic variant. Therefore, I attempted to develop a measure for the individual speaker contact with English. This is a highly community-specific matter; in a different community, the same measure will be based on different local social factors. Ethnolinguistic fieldwork is vital for researchers to understand what categories may be socially relevant in a given community and can be used for that purpose. More research is required to develop solid measures that can potentially even be transferred to other multilingual/diaspora communities.

It may not come as a surprise that the interaction of contact and stance is only significant for speakers who are in regular contact with English. Only speakers with at least moderate contact with English may use the new vowel for identity work, while speakers with fewer contact points may still be in the process of acquiring the linguistic constraints. A possible explanation as to why the interaction is not significant for the only speaker with regular contact with English may be that she is already linguistically integrated in mainstream society and conforms to the traditional raising pattern in all alignments and topics.

The variable age was not significant in the quantitative analysis; this may be the case because nine speakers are not enough to provide a wide age range. In contexts with ongoing change, age should also be included in the measure of contact to assess ongoing change. As younger speakers are increasingly in touch with non-members, such as English customers, their way of speaking may (consciously or not) be converging towards the English spoken outside the community. Other (older) speakers, such as Phoebe, may avoid raised /aʊ/ for two reasons: either countering the ongoing movement towards raising the vowel in *out* and *about* or marking themselves as members of a different community.

8 Conclusion

The OOMs represent a community in a long-term contact situation with varying degrees of contact with the language changes happening around them. This paper describes an attempt to capture contact with English without relying on speaker surveys.

All in all, the combination of both a quantitative approach and a qualitative analysis yields fascinating insights and raises more questions. According to the quantitative analysis, the analysed OOMs adopt Canadian Raising overall in the expected linguistic contexts but not to the same extent as previously documented. The OOMs raise mostly in shorter vowels and tautosyllabic syllables, but they

raise the two shibboleths *out* and *about* significantly more than vowels in other raising contexts. The qualitative analysis, however, shows that linguistic context alone cannot fully explain individual raising patterns.

The greater amount of overall variation in /aʊ/ may be explained through openness to social (and linguistic) change. The status of *out* and *about* as indexes of Canadianness is well-established across community boundaries, which makes these two items easily available to OOM speakers for a wide range of stance-taking acts. By variously avoiding or adopting it in the unfolding interviews, they can use the feature to position themselves with regard to topics and objects talked about. Depending on the speaker, the cumulative result of such stance-taking is to emphasise a traditional Mennonite identity or to represent selective identification with the Canadian mainstream. For the two speakers who are in moderate contact with English, the interaction of raising with topic and alignment is significant. The analysed snippets of Phoebe's raising behaviour suggest that she does not simply use the lexical shibboleths idiosyncratically but rather avoids them and overraises in some /aɪ/ contexts.

Despite remaining suggestive, the findings of this paper shed light on the importance of qualitative research and understanding individual biographies. Due to the scope of this paper, the qualitative approach has remained superficial but already indicates interesting reasonings. Speakers seem to deal with the ongoing linguistic change individually. This has become apparent in Phoebe's under-raising of *out* and *about*, the vowels in which, according to the quantitative analysis, are predicted to be raised. As the participation in ongoing language change very much hinges on contact with English, the speakers' individual social contexts must be outlined and understood in order to paint the full picture and develop appropriate measures for the quantification of contact with English.

Stance may not tell us everything, especially not in the initial stages of a language change. What it does, however, is provide us with a fine-grained picture that enables us to study how speakers use a newly available linguistic resource to position and align themselves with regard to topics and listeners. Such stance-taking is first and foremost part of their individual efforts to express identity through language, but in the long run it will lead to the emergence of more consolidated ethnolinguistic repertoires for the community.

Abbreviations

OOM Old Order Mennonites

PGE Pennsylvania German English

PG Pennsylvania German

Acknowledgements

I would like to thank the Mennonite community and the Elmira locals for their heart-warming hospitality and their willingness to share their time and life stories with me as well as letting me become part of their world. The field trip was made possible by a DAAD scholarship and the generous support of Christian Mair and Sali Tagliamonte. Last but not least, I would like to thank the two anonymous reviewers for their productive feedback.

References

- Anderson, Vicki Michael. 2011. *Bidialectalism in intense language variety contact: An “unexpected” development in the death of Pennsylvania Dutchified English*. Indiana University. (Doctoral dissertation).
- Ash, Richard, James Crook, Roger Dannenberg, Benjamin Drung, Vaughan Johnson, Leland Lucius & Martyn Shaw. 2015. *Audacity*. <https://www.audacityteam.org>.
- Baker-Smemoe, Wendy & David Bowie. 2015. Linguistic behavior and religious activity. *Language & Communication* 42. 116–124.
- Barnes, Sonia. 2018. Copula variation in Asturian Spanish and the multidimensionality of stancetaking in interaction. *Journal of Sociolinguistics* 22(1). 29–54.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.
- Benor, Sarah Bunin. 2011. *Mensch, bentsh, and balagan*: Variation in the American Jewish linguistic repertoire. *Language & Communication* 31(2). 141–154.
- Bohmann, Axel & Wiebke Ahlers. 2021. Stance in narration: Finding structure in complex sociolinguistic variation. *Journal of Sociolinguistics*. 1–19.
- Bolker, Ben & David Robinson. 2022. *broom.mixed: Tidying methods for mixed models*. R package version 0.2.9.4. <https://CRAN.R-project.org/package=broom.mixed>.
- Buchholtz, Mary. 1999. You da man: Narrating the racial other in the production of white masculinity. *Journal of Sociolinguistics* 3(4). 443–460.
- Buchholtz, Mary & Kira Hall. 2005. Identity and interaction: A sociocultural linguistic approach. *Discourse Studies* 7(4–5). 585–614.
- Burridge, Kate. 1992. Creating grammar: Examples from Pennsylvania German, Ontario. In Kate Burridge & Werner Enninger (eds.), *Diachronic studies on the languages of the Anabaptists*, 199–241. Bochum: Brockmeyer.

- Burridge, Kate. 1998. Throw the baby from the window a cookie: English and Pennsylvania German in contact. In Anna Siewierska & Jae Jung Song (eds.), *Case, typology and grammar*, 71–93. Amsterdam & Philadelphia: John Benjamins.
- Burridge, Kate. 2002. Steel tyres or rubber tyres – maintenance or loss: Pennsylvania German in the ‘horse and buggy’ communities of Ontario. In David Bradley & Maya Bradley (eds.), *Language endangerment and language maintenance: An active approach*, 203–229. London: Routledge.
- Chambers, Jack K. 1973. Canadian Raising. *Canadian Journal of Linguistics* 18. 113–135.
- Chambers, Jack K. 1989. Canadian Raising: Blocking, fronting, etc. *American Speech* 64(1). 75–88.
- Chambers, Jack K. 2006. Canadian Raising retrospect and prospect. *The Canadian Journal of Linguistics* 51(2/3). 105–118.
- Dailey-O’Cain, Jennifer. 1997. Canadian Raising in a midwestern U.S. city. *Language Variation and Change* 9(1). 107–120.
- Draper, Barb. 2010. *The Mennonites of St. Jacobs and Elmira: Understanding the variety*. Kitchener: Pandora.
- Du Bois, John W. 2007. The stance triangle. In Robert Englebretson (ed.), *Stance-taking in discourse. Subjectivity, evaluation, interaction*, 139–182. Amsterdam & Philadelphia: John Benjamins.
- Eckert, Penelope. 2000. *Linguistic variation as social practice: The linguistic construction of identity in Belten High*. Malden, MA: Blackwell.
- Eckert, Penelope. 2012. Three Waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41. 87–100.
- Epp, Marlene. 2012. *Mennonites in Ontario*. 3rd edn. Waterloo, ON: The Mennonite historical society of Ontario.
- Frantz, John B. 2017. Religion. In Simon J. Bronner & Joshua R. Brown (eds.), *Pennsylvania Germans. An interpretative encyclopedia*, 131–147. Baltimore: John Hopkins University.
- Fruehwald, Josef. 2008. The spread of Raising: Opacity, lexicalization, and diffusion. *Selected Papers from NWAV36* 14(2). 83–92.
- Gadanidis, Timothy, Angelika Kiss, Lex Konnelly, Katharina Pabst, Lisa Schlegl, Pocholo Umbal & Sali A. Tagliamonte. 2021. Integrating qualitative and quantitative analyses of stance: A case study of English *that*/zero variation. *Language in Society* 52(1). 27–50.
- Hazen, Kirk. 2000. *Identity and ethnicity in the rural south: A sociolinguistic view through past and present be*. Durham, NC: American Dialect Society.

- Huffines, Marion Lois. 1984. The English of the Pennsylvania Germans: A reflection of ethnic affiliation. *The German Quarterly* 57(2). 173–182.
- Huffines, Marion Lois. 1986. Intonation in language contact: Pennsylvania German English. In Werner Enninger (ed.), *Studies on the languages and the verbal behavior of the Pennsylvania Germans I*, 25–36. Stuttgart: Franz Steiner.
- Johnstone, Barbara, Jennifer Andrus & Andrew E. Danielson. 2006. Mobility, indexicality, and the enregisterment of ‘Pittsburghese’. *Journal of English Linguistics* 34(2). 77–104.
- Joos, Martin. 1942. A phonological dilemma in Canadian English. *Language* 18(2). 141–144.
- Kendall, Tyler & Valerie Fridland. 2021. *Sociophonetics*. Cambridge: Cambridge University Press.
- Kopp, Achim. 1997. Die Phonologie des Englischen der Pennsylvaniadeutschen als Indikator für Spracherhalt und Sprachverlagerung. *Zeitschrift für Dialektologie und Linguistik* 64(1). 1–36.
- Kopp, Achim. 1999. *The phonology of Pennsylvania German English as evidence of language maintenance and shift*. Selinsgrove: Susquehanna University.
- Labov, William. 1971. The study of language in its social context. In Joshua A. Fishman (ed.), *Basic concepts, theories and problems: Alternative approaches*, 152–216. Berlin & Boston: De Gruyter Mouton.
- Labov, William. 1972a. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania.
- Labov, William. 1972b. Some principles of linguistic methodology. *Language in Society* 1(1). 97–120.
- Levshina, Natalia. 2015. *How to do linguistics with R. Data exploration and statistical analysis*. Amsterdam & Philadelphia: John Benjamins.
- Lüdecke, Daniel. 2018. ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software* 3(26). 772. <https://cran.r-project.org/web/packages/ggeffects/index.html>.
- Lüdecke, Daniel, Mattan Ben-Shachar, Indrajeet Patil, Philip Waggoner & Dominique Makowski. 2021. performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software* 6(60). 3139.
- Max Planck Institute for Psycholinguistics. 2019. *ELAN The Language Archive*. Nijmegen. <https://tla.mpi.nl/tools/tla-tools/elan/>.
- Meechan, Marjory. 1999. *The Mormon drawl: Religious ethnicity and phonological variation in southern Alberta*. Ottawa: University of Ottawa. (Doctoral dissertation).
- Milroy, Lesley. 1980. *Language and social networks*. Oxford: Blackwell.

- Moore, Emma & Paul Carter. 2018. Natural phonetic tendencies and social meaning: Exploring the allophonic raising split of PRICE and MOUTH on the Isles of Scilly. *Language Variation and Change* 30. 337–360.
- Moreton, Elliott & Erik R. Thomas. 2007. Origins of Canadian Raising in voiceless-coda effects: A case study in phonologization. In Jennifer Cole, José Ignacio Hualde & Aditi Lahiri (eds.), *Laboratory Phonology 9*, 37–64. Berlin & Boston: De Gruyter Mouton.
- Neuhausen, Miriam. 2023. Understanding, collecting, and presenting data in New Englishes research: Insights from ethnographic fieldwork in an Old Order Mennonite community. In Guyanne Wilson & Michael Westphal (eds.), *New Englishes, New Methods*, 222–242. Amsterdam & Philadelphia: John Benjamins.
- Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18(1). 62–85.
- Nycz, Jennifer. 2013. Changing words or changing rules? Second dialect acquisition and phonological representation. *Journal of Pragmatics* 52. 49–62.
- Nycz, Jennifer. 2018. Stylistic variation among mobile speakers: Using old and new regional variables to construct complex place identity. *Language Variation and Change* 30. 175–202.
- Podesva, Robert J. 2007. Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics* 11(4). 478–504.
- Python software foundation. 2019. *Python 2.7.10*. <https://www.python.org>.
- R core team. 2020. Vienna: The R foundation. <http://www.r-project.org/index.html>.
- Raith, Joachim. 1996. Religiöse Migration. In Hans Goebel, Peter H. Nelde, Starý Zdeněk & Wolfgang Wölck (eds.), *Kontaktlinguistik: Ein internationales Handbuch zeitgenössischer Forschung*, vol. 1, 311–320. Berlin: Walter de Gruyter.
- Richter, Manfred. 1969. *The phonemic system of the Pennsylvania German dialect in Waterloo County, Ontario*. Toronto: University of Toronto. (Doctoral dissertation).
- Roberts, Julie. 2016. Internal boundaries and individual differences: /aʊ/ raising in Vermont. *American Speech* 91(1). 34–61.
- Rosen, Nicole & Crystal Skriver. 2015. Vowel patterning of Mormons in southern Alberta, Canada. *Language & Communication* 42. 104–115.
- Rosenfelder, Ingrid. 2007. Canadian Raising in Victoria, BC: An acoustic analysis. *AAA: Arbeiten aus Anglistik und Amerikanistik* 32(2). 257–284.
- Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini, Scott Seyfarth, Kyle Gorman, Hilary Prichard & Jiahong Yuan. 2014. *FAVE (Forced Alignment and Vowel Extraction)*. <https://github.com/JoFrhwld/FAVE>.

- Sadlier-Brown, Emily. 2012. Homogeneity and autonomy of Canadian Raising. *World Englishes* 31(4). 534–548.
- Shields, Kenneth. 1987. Germanisms in the English of eastern Pennsylvania. *Journal of English Linguistics* 20(2). 163–180.
- Springer, Otto. 1980. The study of the English of the Pennsylvania Germans. In Joshua A. Fishman & J. L. Dillard (eds.), *Perspectives on American English*, 195–203. The Hague: Mouton.
- Stanford, James N. 2009. Clan as a sociolinguistic variable: Three approaches to Sui clans. In James N. Stanford & Dennis R. Preston (eds.), *Variation in indigenous minority languages*, 463–484. Amsterdam & Philadelphia: John Benjamins.
- Stanley, Joseph A. 2022. Interpreting the order of operations in a sociophonetic analysis. *Linguistics Vanguard* 8(1). 279–289.
- Tagliamonte, Sali A. 2003. *Linguistic changes in Canada entering the 21st century*. Research grant #410-2003-0005. Social sciences & humanities research council of Canada (SSHRC).
- Tagliamonte, Sali A. 2007. *Directions of change in Canadian English*. Research grant #410-070-048. Social sciences & humanities research council of Canada (SSHRC).
- Tagliamonte, Sali A. 2010. *Transmission and diffusion of linguistic change in Canada*. Research grant #410-101-129. Social sciences & humanities research council of Canada (SSHRC).
- Thomas, Erik R. 1991. The origin of Canadian Raising in Ontario. *Canadian Journal of Linguistics* 36(2). 147–170.

Chapter 5

Why a language dies: The case of Bəṭaḥrēt in Oman

Fabio Gasparini

Free University of Berlin

This chapter gives an account of my experiences conducting fieldwork on Bəṭaḥrēt (commonly referred to as Baṭḥari), a Semitic–Afroasiatic language that is critically endangered and spoken by fewer than 10 elders in the eastern part of the governorate of Dhofar, Oman. Using data and observations collected during fieldwork in the area between 2016 and 2017, I will address community and speakers’ attitudes in order to understand why the Baṭaḥira tribe have switched almost completely to the dominant Arab–Bedouin identity, which will inevitably lead to the disappearance of traditional Baṭaḥira heritage in just a few decades. While the influence of colonialism and the growth of Arab nationalism since the start of the twentieth century played a crucial role in shaping the contemporary cultural landscape elsewhere in the Middle East (Khalidi et al. 1991; Miller 2003), Dhofar, its people, and its cultures remained disconnected and almost unknown to outsiders until the 1970s, when the country was unified and underwent a process of rapid Arabization. Prior to this, its inhabitants lived in seminomadic tribal groups. The Baṭaḥira have undergone a process of identity reshaping since then. The shift toward the local Bedouin Arabic culture inevitably also involved a process of language shift; in fact, Arabic has now replaced Bəṭaḥrēt in every social domain within the local community.

1 Introduction

Bəṭaḥrēt is one of the six Semitic–Afroasiatic Modern South Arabian languages,¹ the other five being Mehri, Hobyōt, Ḥarsūsi, Soqotri, and Jibbāli (also known

¹Henceforth MSALs.



as Shehri, Šherēt, or Gəblēt). Modern South Arabian is an endangered group of unwritten minority languages currently spoken by around 200,000 people in Eastern Yemen and Soqatra, Western Oman, and the southernmost part of Saudi Arabia. Bəṭaḥrēt is listed by the *UNESCO Atlas of the world languages in danger* (Moseley 2010) as “critically endangered,” and is the most endangered among the MSALs. It is spoken to varying degrees of fluency by fewer than 10 elders belonging to the Baṭāḥira tribe and thus destined to disappear in a few decades, at best. Current literature uses the name *Baṭhari* for the language, which is the Arabized form (with the *nisba* ‘relation’ suffix *-i*). However, native speakers refer to it as *Bəṭaḥrēt*; therefore, this label will be used throughout this chapter.

Minorities can only be defined within the wider social context in which they are found. As Miller (2003: 2) says, “the concept of minority implies the notion of inequality either in terms of demographic [...] or sociopolitical weight.” Theoretically, this description aptly describes the case of the Baṭāḥira until recent times; according to local narratives concerning tribal life before the 1970s (the time of the unification of Oman), this small tribe used to live in geographic and political isolation, with few – albeit constant – relationships with the neighboring Janaba, Ḥarāsī, and Mahra tribes (Morris 2017). However, today, the Baṭāḥira tend not to identify themselves as a separate minority within the multifaceted cultural landscape of contemporary Oman. Younger generations look to a unifying national Omani identity and the traditional Bedouin cultural heritage of neighboring Janaba as a means of self-representation. This allows them to feel like an integrated part of the Arab world on a wider scale, while at the same time maintaining a well-marked regional identity as a means of preserving an ethno-anthropological peculiarity (albeit partially divergent from that of their ancestors).

This case study aims to portray my experiences with the Baṭāḥira tribe while conducting fieldwork. In doing so, I will highlight the processes that led the Baṭāḥira to abandon their tribal language and traditional culture. This chapter is based on certain considerations initially presented as an introduction to my unpublished PhD thesis (Gasparini 2018) and on a series of field notes.

2 The Baṭāḥira

This section aims to give a brief and basic description of the setting in which Bəṭaḥrēt speakers previously lived and live today. A comprehensive, rich ethnographic portrait can be found in Morris (Forth.). However, a broad outline at least is needed to understand why the language is on the verge of extinction and why it remained understudied for such a long time – and will also afford a better understanding of the fieldwork conditions.

2.1 The history of the tribe

Until recently, very little was known in general about the small Baṭāḥira tribe. Mentions of the tribe in literature are almost non-existent, most likely because it occupies an area that is far from hospitable, hard to reach, entirely desert and barren – apart from a few springs. This area must have appeared considerably less welcoming than the lush, reassuring coasts of Salalah, which meant that Western visitors and scholars preferred to focus their interests far from the area inhabited by the Baṭāḥira.

The first documented report to mention the tribe was written by the famous British explorer Bertram Thomas (1929: 100). His words about his encounter with a member of the tribe are rather unflattering and are a source of annoyance to some Baṭāḥira today. A list of words is presented in a later paper (Thomas 1937). In addition to a brief mention by Dostal (1960), some minor information can be found in Janzen (1967). The only person who has been able to conduct fieldwork with the Baṭāḥira was Miranda J. Morris, who worked in the area between the 1970s and 1980s and returned many years later, starting from 2013 and until 2016. She collected many recordings that had been left unpublished until recently, and much of the material collected at the time – together with more recent recordings (Morris 2016) – is publicly available. Furthermore, a printed edition of a selection of texts is nearing publication (Morris Forth.) and a descriptive grammar is in the making (Morris & Gasparini Forth.). Morris is also the author of two papers concerning Bəṭaḥrēt; one (Morris 1983) deals with Bəṭaḥrēt songs and poems, while the other (Morris 2017) reports on thoughts and problems concerning the study of endangered languages (particularly addressing Hobyōt and Bəṭaḥrēt), with various considerations and samples from the languages in question. Further fieldwork and linguistic analysis on the language was recently conducted by Gasparini (2017; 2018; 2021a; 2021b) and Bettega & Gasparini (2020).

Nowadays, the Baṭāḥira are settled in the coastal area of the far east of Dhofar, near the border with the al-Wuṣṭā governorate. Members of the tribe are scattered from the villages of Likbi to Ṣowḵarə, with most of them living in Ashwaymiyah (Figure 1).

A clan belonging to the tribe, the Bayt Kdaš, moved to the area around Salalah in the past; they still live there, and having been assimilated into the Ṣherēt-speaking communities, they no longer speak Bəṭaḥrēt. Reports from some of the members of the tribe (although it should be noted that the Baṭāḥira living in Salalah are particularly insistent on this point) say that their tribal territory once reached the two Wādi Ġadōn, which can be found approximately 30 km to the west of Tamrīt and 10 km east of Ṣowḵarə, respectively, stretching south



Figure 1: Map of the area inhabited by the Baṭāḥira (Base layer: ESRI World Imagery. Data edited by E. Croce, F. Gasparini, and M. J. Morris, 2022).

near the mountains that divide Salalah from the desert inland. Migration and invasions by the Janaba from the northeast and from the Mahra coming from Yemen reportedly pushed the Baṭāḥira toward the area which they inhabit today. alTabuki (1982: 54–55) places this event up to around 300 years ago, but the absence of historical records prevents further investigation. Because of these invasions, the tribe lost control of their former land, suffered a dramatic decrease in numbers, and were subjugated to the conquerors. A recurrent folk etymology connects the tribe's name to the term *baṭḥ* 'dust, sand,' which would allude to them previously being as numerous as grains of sand. The dramatic reduction in their numbers is explained through two disjointed myths. The first refers to the slaughtering of *nākat Ṣāliḥ* 'the she-camel of the Prophet Ṣāliḥ,' for which God sent a plague of insects that ate all the Baṭāḥira's belongings as punishment, leaving this once prosperous people to starve to death. The second myth is connected to the memory of fierce fighting against the so-called *Burtuḡaliyīn* 'the Portuguese.' Narratives related to these events constitute common lore among the younger members of the tribe and are imbued with a strong sense of pride.

During my stay in Ashwaymiyah, I was shown a cave along the beach of Warx, an ancient settlement at the mouth of a *wādī* to the east of Ashwaymiyah. The cave, which can only be reached by boat, is where the foreign invaders would be imprisoned; a large number of graves grouped together in the same area is ascribed to a great massacre of women and children perpetrated by the invaders – according to traditional folklore; this caused the drastic decrease in numbers of the tribe.

Any statement concerning the position of the Baṭāḥira within the context of past and current tribal power relationships should best be avoided. According to Thomas' (1929) informants, the Baṭāḥira were considered to belong to the lowest social scale. However, intertribal marriage between Baṭāḥira women and men from the Janaba and the Ḥarsūsi-speaking tribes in the north was relatively frequent (less so for Baṭāḥira men); the Baṭāḥira had ownership over some fishing sites and frankincense trees; in times of famine, it was common for the bedouins living inland to share food with the Baṭāḥira fishermen and vice versa. In all likelihood, the situation was not as simplistic as Thomas portrayed it; furthermore, these issues are still relevant to the Baṭāḥira and constitute a salient part of their identity, and must therefore be approached carefully by nonmembers of the community.

While the interactions between the Portuguese and Oman are well documented in the northern part of the country, the same cannot be said for Dhofar, where the records are scarce and far from exhaustive. What is known is mostly due to sailors, travelers, and adventurers; Portuguese sailors, members of Vasco de Gama's navy, did indeed stop in Hallāniyāt (formerly known as Kuria Muria), an island not far from the shores of Ashwaymiyah, for several months between 1502 and 1503. The raising of one of the sunken ships from that expedition, together with written records from the period, confirms this historical event (Mearns et al. 2016). However, there is no tangible evidence of any conflict between the Baṭāḥira and the Portuguese navy, which apparently had good relations with the islanders, while no contact with the inland population is documented. Some truth may yet be found as to the origin of these narratives, but given the lack of historical records, these reports can only be treated as oral folk tradition.

2.2 Environment and lifestyle

The area traditionally inhabited by the Baṭāḥira is characterized by severely dry weather. The climate is not affected by the *xarīf* season (that of tropical monsoons, between June and August), which makes the plain of Salalah prosperous and fertile. The desert and desolate landscapes of the area inhabited by the

Baṭāḥira were not particularly hospitable, and a paucity of natural springs, vegetation, pastures, and wild animals to hunt made traditional life very hard, according to recollections of the eldest members of the tribe (who are also the last remaining speakers of Bəṭaḥrēt).

The Baṭāḥira's diet was composed almost exclusively of fish, shellfish, and turtles (abundant off the coast of Oman), camel and goat milk (from animals they bred), and, occasionally, rice and dates, depending on the time of the year and trade. Narratives and personal stories constantly emphasize the absolute lack of sufficient means of survival, with chronic starvation and diseases cyclically decimating the numbers of the tribe. The elders understandably carry with pride the fact of them being able to survive nonetheless in such harsh conditions. Water was fetched by women from various coastal springs, often located miles away from their areas of settlement. Living a seminomadic way of life, the Baṭāḥira would be either cave-dwellers or build small stone houses, which are still recognizable, especially in Mingíy, another abandoned settlement west of Šarbithat. Daily activities were carried out almost exclusively during daylight hours, as leaving the camp during the night was extremely dangerous and done only in case of emergency.

These harsh conditions, made worse by chronic starvation and disease, eventually came to an end after the unification of Oman in the 1970s. The Baṭāḥira completely gave up their nomadic lifestyle and now live a quiet, relatively healthy life in modern houses with all the standard comforts of the contemporary world. The whole tribe quickly switched to Arabic (and other MSALs), and at the time of writing, there are fewer than 10 relatively fluent Bəṭaḥrēt native speakers remaining.

The reasons that led to this swift process of language shift will be specifically addressed in the following sections.

3 Working with the Baṭāḥira: The difficulties of fieldwork

The following section provides a general description of my experiences in the field with the Baṭāḥira.² While my primary goal was to gather material for what would later become my PhD dissertation (Gasparini 2018), I also tried to analyze

²Unfortunately, further fieldwork after 2018 was initially prevented by a lack of additional funding and, after the start of my postdoctoral studies at Freie Universität in April 2020, by the global COVID-19 pandemic. This prevented me from traveling to Oman until recently (January 2022), when I finally managed to get back in the field. Data from the current fieldwork were not included due to time constraints.

what I witnessed from a sociolinguistic point of view, by paying attention to attitudes and opinions toward the use of language within the community.

3.1 Fieldwork description

In order to find a way to connect with the Baṭāḥira community, I got in touch with Dr. Miranda J. Morris from the University of St. Andrews, Scotland, who very kindly agreed to help me by sharing her local connections and part of her Bəṭaḥrēt recordings. An initial meeting with her and her main field collaborator, Khalifa Hamoud alBaṭḥari, a member of the community himself, took place in November 2015 in St. Andrews, Scotland, where the two were at that time working on Morris' corpus of ethnographic recordings. During my stay, I was able to discuss the possibility of a period of fieldwork in Oman with Khalifa alBaṭḥari, to which he eventually agreed.

Social scientists approach field research in different ways, according to their epistemological assumptions and goals. Often, the main goal of fieldwork in linguistics is to conduct research in the service of academia and the linguist's individual career, whereas the potentially exploitative effects of being researched on are tentatively minimized but do not necessarily become crucial to the researcher. Within this framework, fieldwork is conducted in such a way that the linguist collects linguistic data for their own purposes, with little consideration for the speakers' needs and desires in respect to their language and culture. This kind of approach is referred to as *ethical* (Cameron et al. 1992: 15). *Advocacy* research (Cameron et al. 1992: 15) brings the researcher to actively use their knowledge to promote the researched's needs. Finally, the *empowering* model (Cameron et al. 1992: 22) requires the researcher to work in strict collaboration with speech community members, with the aim of building a truly reciprocating relationship between the researcher and the community. The latter approach surely is the best way to care for language data, especially when dealing with endangered and minority communities.

The initial stages of my fieldwork were conducted over the course of two stays between October and November 2016 and March and April 2017. During my first stay, I settled in Šəlīm, a small town – mainly inhabited by local workers from South Asia and some Mahra – on the plateau surrounding the plain of Ashwaymiyah. I paid daily visits to the elder Bəṭaḥrēt speakers together with Khalifa alBaṭḥari, who would arrange the meetings for me; otherwise, the elders would not have agreed to work with me, as happened to other researchers who previously tried to work on Bəṭaḥrēt. Due to temporal and financial constraints,

I was unable to spend enough time in the community to completely earn people's trust and better understand community needs. I could not perform immersion fieldwork, understood as the observation of how language is actively used by the community during a consistent period of time (Aikhenvald 2007), since Bəṭəḥrēt is effectively a moribund language known by only a handful of elders. Furthermore, at the start of my research, my interest in the community was primarily of a purely linguistic nature, which led me to adopt an *ethical* approach to my fieldwork. It was not long before I understood how inadequate my linguistic analysis would have been without a better understanding of the cultural setting. I found myself working on a language none of the members of the community had any interest in working on with me (apart from Khalifa, who is not a proficient speaker); the amount of work the last remaining speakers did with Morris right before my arrival was more than enough for them to pay her back for everything she had done for the tribe in the past. This is how I came to terms with the importance of developing a reciprocal relationship with the community. The production of a descriptive grammar did not meet primary community needs, but reminiscing and illustrating personal histories and traditional heritage (and being shown respect for their knowledge) was the main priority to many of the speakers. It became necessary to mediate between my needs as a researcher and the needs of the speech community. Therefore, I soon abandoned mostly unsuccessful direct elicitation from the contact language (Arabic) and tried to focus on monologic, monolingual, unscripted speech, discussing relevant aspects of traditional culture and spending my time with Khalifa alBaṭḥari translating the recordings into Arabic when not conducting interviews. This way I learnt how honour and respect were the most important values for the elders of the community, and without building solid trust first I would not have achieved much in my fieldwork.

My second round of fieldwork saw some considerable improvements. I was able to secure an apartment in Ashwaymiyah, a small yet lively center today famous for its rich fish trade and where most of the tribe is settled. This made the whole project much easier due to daily contact with the community, and I also had the chance to meet elders outside the controlled context of interviews.

During my second stint, I was able to get in touch with the younger members of the tribe, some of whom were the grandsons and nephews of the elder Bəṭəḥrēt speakers. I would spend my evenings after fieldwork sitting outside coffee shops and talking to the curious young men who would approach me, gathering their impressions of my work and their cultural heritage through informal discussions.

I would now like to turn the focus from the observer to the observed, that is to say the interviewees I worked with.

3.2 The interviewees

Two main groups of interviewees were considered, namely the speakers of Bəṭaḥrēt and the younger members of the tribe (up to 30 years old), all of them living in Ashwaymiyah.

The first group, which was also my target group for my primary linguistic inquiry, included the elderly men and women of the tribe whose mother tongue is Bəṭaḥrēt and who were adjudged by Khalifa and the community itself to be reliable speakers. Of the few who were left, I was able to work with six men and three women. Their exact ages were not clear, but they were certainly born years before Sultan Qaboos acceded to the throne, by which time they were young adults, meaning they would now be between 60 and 80 years old. All of them are illiterate, unlike most of their descendants, who underwent free schooling after the unification of Oman.

Nowadays, all the interviewees are bilingual with Arabic, which has become their daily means of communication, and all of them but one know at least one other MSAL (either Mehri or Ḥarsūsi) as a consequence of frequent intertribal marriage.

Whenever I was left to spend time by myself in Ashwaymiyah, I would have the chance to socialize with the younger male members of the Bəṭaḥira tribe living there (educated children and young adults, some of whom attended or were attending university in Muscat). They are for the most part the (great-)grandsons/nephews of the elder Bəṭaḥrēt speakers and were genuinely curious about my interest in the old language of the tribe, which was completely obscure to them. Of the dozen young men I would frequently meet during those evenings, two became my main interlocutors. This second group would later come to be a precious – and unexpected – source of knowledge. Local gender restrictions did not allow me to interact with members of the opposite sex who were the same age as me. However, as their male peers and the elders themselves reported, not even the young women of the tribe have any competence in Bəṭaḥrēt, despite the fact that they spend more time in the household in close contact with female elders. Šəlim did not offer any such chances of socialization, since it is mostly inhabited by South Asian workers, and the small community of Omani residents – with whom I had no contact – belong to the Mahra tribe.

Day by day, I was able to collect the impressions of these young men through informal discussions during or after dinner. I did not record these sessions since they would happen by chance, mostly out in the open air and surrounded by other people passing by. However, I did make extensive use of handwritten notes.

4 Linguistic repertoires and community attitudes toward Baṭāḥrēt

Nowadays, the Baṭāḥira tribe shares the area with other local tribes: Janaba families can be found east of Ashwaymiyah; Šherēt-speaking families populate most of the villages along the coast toward Salalah; and the Mahra inhabit the inner land. There is also the overwhelming presence of South Asian laborers working either in nearby oil fields and local small shops or as servants in private houses. Furthermore, intertribal marriage and subsequent relocation have always been common practice. Ashwaymiyah has also become an important fishing harbor due to its particularly rich waters. For this reason, the area is well known to traders, fishermen, and even tourists interested in fishing. All this shows that the local community has become an integrated part of the wider Omani society.

My report concerning the speech community of Ashwaymiyah, where speech community is understood as “a whole of people, of undetermined size, sharing access to a whole of language varieties and united by some sort of socio-political grouping”³ (Berruto 2002: 60), does cover two important segments of the local population; however, I could not get past very formal levels of interaction with the generation between these two. This is much to my regret, as a closer look at their communicative habits would have been critical since they are the generation where the change in the community’s linguistic behavior happens. However, there is no direct evidence that their linguistic repertoire is divergent from that of their fellow tribesmen as reported here.

Arabic is undoubtedly the primary language of the community. Morris (1983; 2017) refers to the variety spoken by the Baṭāḥira by the label *Janaba Arabic*, which is the dialectal variety of Arabic spoken by the neighboring Janaba tribe, with whom the Baṭāḥira have always been in close contact. It is important to note that Modern Standard Arabic carries an ‘overt’ prestige only on certain formal occasions, for example when talking to a foreigner such as myself who would not likely be taken for a native speaker of Arabic. In fact, the younger, educated members of the tribe would talk to me in Modern Standard Arabic, presumably seeing our interactions as pertaining to a very formal register. The elders, however, do not have access to this variety since they were not formally educated – and they would occasionally be mocked for this.

Few of those young men show any interest in their linguistic heritage. Some can understand a few words and basic expressions, but none of them has any real

³Own translation; the original reads: “un insieme di persone, di estensione indeterminata, che condividano l’accesso a un insieme di varietà di lingua e che siano unite da una qualche forma di aggregazione socio-politica.”

competence in the language. The most common reason for their lack of interest is that Bəṭaḥrēt would be completely useless in their daily lives, since no one outside their hometown would understand them – not even the neighboring Mahra, with whose language there is only scarce intelligibility. The young Bəṭaḥira view Arabic as a powerful tool to enrich themselves and move to bigger cities (mainly Salalah, Muscat, or the Emirates). In fact, evident prestige was accorded to those who could master Modern Standard Arabic at a higher level. A good example of this was the case of M.; 24 years old, with a degree in engineering and a high level of Modern Standard Arabic, he would be automatically elected as the most entitled one to guide a group conversation with me, and would be addressed as a medium between me and the rest of the group. Most of the other youngsters spoke only Arabic (in its vernacular and standard forms) and the local Pidgin Gulf Arabic variety, while one individual was fluent in Ḥarsūsi, as his mother came from the Ḥarāsīs tribe.

The time I spent with the elders was for the most part dedicated to learning more about Bəṭaḥrēt. Usually, our recording sessions would revolve around a series of ethnographic topics. When questioned about the reasons which led them not to speak Bəṭaḥrēt with their children anymore, all the elders agreed in stating the uselessness of Bəṭaḥrēt in the rapidly changing post-revolutionary Oman. The last remaining speakers do not seem to have any will or see any need to teach the language; they feel too old to start, lack the energy to see it through and are scarcely interested in speaking it anyway. It is a common opinion – not only in the area, but in the Arab-speaking world in general (see Kaye 2001) – that ‘proper’ languages are taught at school, while daily, home speech has to be considered vernacular. In fact, all of the remaining speakers consider Bəṭaḥrēt to be a virtually dead language – which is true from a sociolinguistic point of view, with the use of the language being maintained in no social domain whatsoever apart from arranged interviews. There is only one occasion which sees the speakers use Bəṭaḥrēt; during social gatherings, some traditional Bəṭaḥrēt songs are occasionally sung, mainly by the men of the tribe, and recordings are shared using smartphones. The overall meaning of these songs is not always preserved, though, and sometimes not even the elders remember the meaning of the words.

However, it must be recognized that, after Morris’ work, a rediscovered sense of pride in the adventurous narratives of past daily struggles to survive in extreme conditions is palpable. The young men I met feel at least great respect for their grandfathers and their material culture, even though they do not have the will to maintain it. Their social networks are usually wider than those of their grandparents, and frequently include individuals and tribes from other villages. Furthermore, they socialize also with the great numbers of immigrant workers

from South Asia (who often speak Gulf Pidgin Arabic) and foreigners in general, mostly (but not only) for the sake of daily life and work-related reasons. This situation is common throughout the whole Gulf area (Holes 2011: 137) and has radically changed the sociolinguistic landscape of the region.

The suppressed Omani revolution and the subsequent “renaissance” overturned the traditional way of life of the Baṭāḥira. In the new Omani society that emerged, the role of Arabic was that of a unifying and empowering tool that the tribe was very keen to fully adopt for the sake of their involvement in the globalized world. Therefore, there was no reason left to maintain their native tongue – this was the point at which parents ceased transmitting it to their children. Language shift usually happens at a slower pace and requires many generations to reach its end (Romaine 1989), yet this is not true in the case of the Baṭāḥira, where only one generation was needed to nearly complete the process. Bəṭaḥrēt disappeared from daily life in the community, and young people now show clear signs of cultural shift toward an Arab identity, under the heavy influence of a dominant image of Arab–Bedouin heritage.

4.1 “Iḥna bēdu”: Identity replacement in the youngest generations

The discourse concerning cultural representation and self-identification has become one of real interest within both ethno-anthropological frameworks and daily discourse in the globalized world (Appadurai 1995; Hannerz 1997). It is striking to note how quickly the process of Arabization led the Baṭāḥira to adopt an Arab–Bedouin identity almost completely. The new generations are eager to present and identify themselves primarily as Bedouin and not necessarily as an ethnically separate group from the neighboring tribes. It must be noted that the term *bēdu* ‘Bedouin’ is usually used by the elders to refer to the lifestyle of the neighboring Arab, Ḥarsūsi, and Mahri tribes living inland (*bšēli əbādiyā*, ‘people of the desert’) who made a heavier use of camels (*bšēli ḥabēšar* ‘camel people’), as opposed to the fishermen living near the shore (*bšēli ərawnā*, ‘people of the sea’); it is never applied (to my understanding, at least) to the Baṭāḥira as a group, but seldomly to the few Baṭāḥira who owned animals and periodically lived inland as Bedouins. According to this use, the term simply acknowledges the fact that the Bedouins inhabited the desert and does not entail much more; on the contrary, the Bayt Kdaš, the group living in Salalah, seem to idealize the concept of the Bedouin lifestyle in a similar but stronger fashion to the younger generations living in Ashwaymiyah and frequently report it as a distinctive feature of their tribe’s past. Calling themselves Bedouins implies the adherence to an all-round identity, in stark contrast to the former use. This is not unexpected: on the one

hand, in a socially composite city such as Salalah, the Bedouin identity is best fitting for the Bayt Kdaš, since it allows them to proudly elaborate their history from a different perspective; on the other hand, the younger Baṭāḥira were already schooled within the framework of the national education system, which promotes the common Arab identity, where monolingual education in Arabic played a crucial role to the expense of minorities. This also raises the question of whether this widespread self-representation reflects the more or less overt feelings of their own households – possibly suggesting that a systematic process of identity replacement may have taken place in their parents' generation – or if cultural assimilation into their intertribal cohorts was induced through daily and continuous contact. Today, only the Baṭāḥira elders directly experienced the hunger and struggle of pre-unification times. The elders undoubtedly recognize the existence of a system of traditions once belonging and peculiar to the Baṭāḥira, such as material culture, fishing techniques, taboos, rituals, and so forth. They are also very aware that the Arab–Bedouin culture has now completely replaced the way the elders used to live, and that traditional cultural heritage has lost its vitality due to obsolescence. Most of Baṭḥari lore, in fact, stems from a way of life which has completely disappeared today. This knowledge has lost its immediate, practical utility for its community, and as such there is no apparent reason to keep it alive. One evident example is the whole practice of *ṣawf* 'taboo.' The imposition of temporary or permanent taboos on certain actions – such as fishing or food consumption – indirectly regulated the exploitation of natural resources by limiting or prohibiting the catching of specific fish species at a given moment of the month or the year, thus allowing for their ecologically sustainable reproduction. Since the Baṭāḥira have severed their relationship with the sea, passing on these habits to their descendants is no longer significant to the elders nor to their heirs. The value of traditional culture thus becomes evident if it is put in relation to the environment in which it developed. The disconnection between these two levels has had dramatic consequences for the survival of the Baṭḥari culture and language, and has been claimed with regard to the rest of the MSAL-speaking peoples (Watson & alMahri 2017). Another example of dramatic cultural loss relates to the field of traditional medicine and ethnobotanical knowledge, which can be of great importance for modern-day research nonetheless (see, for example, Leonti & Casu 2013).

On the part of the elders, the strong will to integrate into the new, modern Omani society and improve living conditions inevitably meant getting rid of any memory related to a past of hunger and poverty, intrinsically linked to the traditional way of life of the Baṭāḥira. With the language itself being a vestige and a constant reminder of those times, the need to get past their isolation was so

urgent that parents started to talk to their children only in Arabic – education and media did the rest. Before Morris' return to the field in 2014, the few speakers left reportedly had not spoken Bəṭāḥrēt for decades, and it was only thanks to Morris' continuous efforts that they managed to recall their long-unspoken mother tongue (Morris p.c.).

Now it is interesting to note the value of Bəṭāḥrēt within the Bayt Kdaš, a clan of the Baṭāhira tribe that moved to Salalah some time ago and has quite a different story from that of the rest of the tribe. It is not known at what point in history this part of the tribe moved to the city. This event might have taken place in a relatively distant time, since the members of the clan I was able to talk to generically refer to their ancestors when talking about who among them migrated to the mountains of Salalah. Despite belonging to the same tribe, the Bayt Kdaš had no contact with the rest of the tribe for decades, and connections have only been reestablished in recent years. Some of them have now bought or built new houses in the area of Ashwaymiyah and often spend time there. Their life in the lush area around Salalah was radically different from that of the rest of the tribe and they could enjoy much prosperous conditions. In the peculiar urban context of Salalah a kind of balanced diglossia (MSALs being used at home and as tribal languages while Arabic is used in other environments) was established over time, probably because of tribal prestige and power relations; differently from the situation in Ashwaymiyah, there never was a social demand to stop speaking MSALs and be fully integrated to the Arab identity, allowing for culture maintenance. In this scenario, the Bayt Kdaš fully switched to Šḥerēt and Arabic and integrated into the local community through marriages with the Šḥeró tribe. Interestingly, this group is the only one to have shown great interest in my research, seeing in it a way to rediscover their origins and build a strong tribal identity through the tokenization of Bəṭāḥrēt itself. However, I avoided being actively involved with this identity discourse, since I did not want to gaslight the local community in any way by forcing the debate over a very sensitive topic within local society.

4.2 The remaining MSALs

The sociolinguistic situation of the other ethnic groups living in Dhofar and speaking MSALs is considerably harder to evaluate, due to the higher numbers of speakers; nonetheless, a dedicated study on the topic would surely reveal a greater level of complexity, requiring many different layers of analysis. The identity discourse within other more numerous groups may be even more sensitive and, needless to say, would require the interested researcher to apply a certain degree of caution. In any case, the ongoing effects of Arabization are clear, with

younger generations of Dhofaris from MSAL-speaking families progressively losing their linguistic competence in favor of Arabic (Morris 2017) – in this regard, the extremely complex sociolinguistic situation of a city like Salalah would be a fascinating study in itself, as it was hinted in previous sections. Among other things, this loss of linguistic competence is due to the exclusive use of Arabic in education, a lack of written material, and very limited use of MSALs in the media. Arabic is considered to be the one and only language of Oman, necessary to secure a good job and to travel abroad (especially to the Emirates, whose charm and cultural influence over the younger generations is getting stronger). Meanwhile, MSALs are seen by many as the vernacular, dialectal medium and are therefore used in family or local contexts – mostly because they are unwritten, which seems to be a critical factor in determining speakers' opinions. However, things seem to be changing; some members of the local communities are now engaged in documentation and revitalization projects,⁴ and hopefully, their direct involvement will mitigate, if not invert, the process of language and culture loss.

5 Conclusion

Since the rise of modern nation-states and the ideologies behind them in Europe in the twenty-first century, the distinction between an official language and local or vernacular languages has been seen as crucial in order to enhance a shared identity from the point of view of central governments, often struggling against conflicting identities coexisting under the same political entity (Bonfiglio 2010). Apparently, this would not happen – or would happen to a lesser extent – within the pre-modern empires, where no strict language policy was set up and minorities would have a certain amount of independence from the central government (Tosco 2015), meaning the use and demise of a language was strictly dependent on the alternate fates of its community of speakers.

The situation of language minorities in the Arab world has been considered by many scholars. For example, Zaborski (1997), Owens (2007), and Tosco (2015) give general historical and synchronic accounts on the topic, while Miller (2003) focuses her analysis on the contemporary situation of Egypt and Sudan. Grandguillaume (1983) outlines the situation of the Maghreb and Tilmatine (2015) and Maddy-Weitzman (2011) focus on Algeria and the status of Berber. It becomes clear that most of the countries in the Arabic-speaking world historically adopted policies favoring the quasi-exclusive use of Arabic in its standard, official variety,

⁴See, for example, a Wikitongue-funded Language Revitalization Accelerator project by Said Baquir and Abdullah alMahri.

to the detriment not only of other linguistic minorities but also the local dialectal varieties of Arabic itself. From the speakers' perspective, the acquisition of the official language was considered a means of cultural redemption, granting individual acceptance and facilitating integration into society. A recurring concept in many definitions of what a modern nation is refers to an aggregation of people inhabiting a delimited territory and speaking the same language, and this same idea can be found at the basis of pan-Arabist movements from the late twenty-first century onward (Suleiman 2013). The construction of a shared, collective memory is thus seen as a necessary step in the building of a nation. Whatever might be detrimental to this vision is left behind – but this is a colonial legacy. Commenting on a draft version of this paper, Frisone (p.c.) posed the question: “If there were no borders or nations, would this link between identity and memory be so strong?” If they exist, it is because the European and sovereign social organization has spread the nationalistic idea of identification between the people and their history. When this concept is developed into a foundation myth, it builds a more or less coercive bond to a specific territory; borders, in turn, create a sense of belonging to an ethnicity by spatially delimiting the extension of a shared identity. It is entirely a caged macroconstruction of which all the systems of state power, inside and outside Europe, made use nonetheless: its effects are especially visible in the domain of colonial heritage.

It is safe to say that the influence of colonialism and the growth of Arab nationalism during the twentieth century played a crucial role in shaping the contemporary linguistic situation in the Middle Eastern area (Khalidi et al. 1991). The positivist idea of the need for an official language to unify such a wide area under the same macro-identity led to the rise of Modern Standard Arabic as a shared official language, which was undoubtedly beneficial in many respects. One major consequence, however, is that most of the minority languages, which were already struggling, and communities scattered across this extensive territory were put at risk under the pressure of culturally hegemonic (and sometimes violent) central governments. In fact, until very recent times, “[m]inorities' languages [were] almost totally excluded from public life and at best [were] accepted in their “folkloric” forms [... and] almost never taught” (Miller 2003: 8).

In this sense, Oman is a very peculiar case when compared to the rest of the Arabic world. Until its unification, Dhofar was a widely unknown land, where the lack of a central government prevented modern infrastructures from developing, leaving local lifestyles almost untouched by the outside world for centuries. What we do know about its history is still fragmented; unlike Northern Oman, the area inhabited by the MSAL-speaking people has precious few historical records, especially for the time between the fall of the ancient South Arabian

kingdoms and the late nineteenth century. Oman as we know it today was born after the Dhofari revolution, which took place between the 1960s and 1970s. This brought unification to this very diverse land, under the rule of the late Sultan Qaboos bin Said al Said, who rose to power after overthrowing his father, Said bin Taimur, in a palace coup in 1970 and transformed the newly unified nation from a poor, underdeveloped country into a modern state. A major consequence of the war and the socioeconomic change in the subsequent decades has been the blurring of traditional social distinctions (Peterson 2004: 266), with loyalty to Sultan Qaboos being put before tribal rivalries (although these were never completely extinguished).

If compared to the policies towards minorities adopted by other Middle Eastern countries, a noteworthy element of Sultan Qaboos' reign – and that of his successor, Haytham bin Tariq Al Sa'id, for now at least – is that they never directly acted against the heterogeneity of Omani ethnic composition by means of repression or forced cultural substitution, albeit the undeniable hegemonic position of Arabic and Arab identity within the Omani context. It must also be noted that Šherēt and MSAL-speaking groups in general in the western part of Dhofar were considered strictly linked to the separatist rebellion and looked at with suspicion by the government: in this perspective, to indirectly destabilize the maintenance of an important identity element such as language benefits the interests of the central power. It is also true that there are no ongoing safeguard programs addressing minority languages. The only part of the constitution dealing with this topic is Article 3, where it is stated that the official language of the state is the Arabic language.⁵ The preservation of minority languages is left to the enterprising spirit of individual speakers, but without any inclusive language policy, remaining competent in the traditional language often means being cut off from the contemporary world. The ramifications of the loss of language for identity and culture are likely to be enormous and will undoubtedly transform Dhofari society significantly (Peterson 2004: 260).

Like all aspects of a culture, the place of language as a core cultural value is culturally determined (Smolicz 1980) and negotiated within individual and community identities. The major frame of analysis postulates that “language features are the link which binds individual and social identities together” (Tabouret-Keller 2017: 317), but when this ceases, it might well be that other tools are used as an instrument of identity, be it tribal history or whatever is at hand or comes to mind. It seems that for the last remaining speakers of Bəṭaḥrēt, the only place where the full assumption of tribal identity and a real sense of unity for the whole social group can be found – that is, the only instances where they would say *iḥna*

⁵ “لغة الدولة الرسمية اللغة العربية” “the official language of the country is Arabic”.

l-Baṭāḥira ‘us, the Baṭāḥira’ – is connected to the memory of the past, of a world now completely vanished after the advent of modernity. Far from being idealized, this was a world where starvation and poverty were commonplace, and one that they are quite happy to have exchanged for survival and an easier way of life. What undoubtedly remains is a great sense of pride for having been able to survive to such conditions. This is in stark contrast with the part of the tribe in Salalah, which gave new meaning to their tribal belonging in order to better fit within the context of the city. As for the situation in Ashwaymiyah, the language and cultural shift swiftly took place within a couple of generations. There seems to be an unsaid schism between them and the younger generations that is not easy to address directly.

Acknowledgements

My thanks to the Baṭāḥira community for their kindness and for cooperating with me for the purposes of my research. I am also tremendously grateful to Miranda J. Morris, whose help has had an incredible impact on the quality of my work and my learning of the language and culture of the Baṭāḥira. I would like to thank Enrico Croce for producing the map used in Figure 1 and Gloria Frisone for her valuable comments on an early version of this paper. Finally, I also warmly thank Kamala Russell for her precious comments and suggestions, which greatly improved the paper. This research is part of the project Describing the Modern South Arabian Baṭhari language [ref. 40.20.0.007SL], funded by Fritz Thyssen Stiftung.

References

- Aikhenvald, Alexandra Y. 2007. Linguistic fieldwork: Setting the scene. *STUF - Language Typology and Universals* 60(1). 3–11.
- alTabuki, Salim B. 1982. Tribal structures in South Oman. *Arabian Studies* 6. 51–56.
- Appadurai, Arjun. 1995. The production of locality. In Richard Fardon (ed.), *Counterworks: Managing the diversity of knowledge*, 204–225. London: Routledge.
- Berruto, Gaetano. 2002. *Fondamenti di sociolinguistica*. Rome: Laterza.
- Bettega, Simone & Fabio Gasparini. 2020. Modern South Arabian languages. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 351–369. Berlin: Language Science Press.

- Bonfiglio, Thomas P. 2010. Deconstructing the native speaker. In Thomas P. Bonfiglio (ed.), *Mother tongues and nations*, 8–20. Berlin: De Gruyter Mouton.
- Cameron, Deborah, Elizabeth Frazer, Penelope Harvey, Michael B. H. Rampton & Kay Richardson. 1992. *Researching language: Issues of power and method*. 1st edn. London: Routledge.
- Dostal, Walter. 1960. Some remarks concerning the Baṭāḥira, a social inferior tribe in southern Arabia. *Archiv für Völkerkunde* 15. 7–9.
- Gasparini, Fabio. 2017. Phonetics of emphatics in Bathari. «QuadRi» – *Quaderni di RiCOGNIZIONI* 7. 69–85.
- Gasparini, Fabio. 2018. *The Baṭhari language of Oman. Towards a descriptive grammar*. Naples: University of Naples 'L'Orientale'. (Doctoral dissertation).
- Gasparini, Fabio. 2021a. Emphasis, glottalization and pharyngealization in Semitic and Afroasiatic. *Kervan. International Journal of African and Asian Studies* 25(2). 3–30.
- Gasparini, Fabio. 2021b. Nominalizzazione indipendente in sudarabico moderno. In Silvia Ballarè & Guglielmo Inglese (eds.), *Tipologia e sociolinguistica: Verso un approccio integrato allo studio della variazione. Atti del workshop della società di linguistica italiana, 10 settembre 2020*, 177–206. Milan: Officina ventuno.
- Grandguillaume, Gilbert. 1983. *Arabisation et politique linguistique au Maghreb*. Paris: Maisonneuve et Larose.
- Hannerz, Ulf. 1997. Fluxos, fronteiras, híbridos: Palavras-chave da antropologia transnacional. *Mana* 3(1). 7–39.
- Holes, Clive D. 2011. Language and identity in the Arabian Gulf. *Journal of Arabian Studies* 1(2). 129–145.
- Janzen, Jörg. 1967. *Die Nomaden Dhofars/Sultanat Oman: Traditionelle Lebensformen im Wandel*. Bamberg: Fach Geographie an der Universität Bamberg.
- Kaye, Alan S. 2001. Diglossia: The state of the art. *International Journal of the Sociology of Language* 2001(152). 117–129.
- Khalidi, Rashid, Lisa Anderson, Muhammad Muslih & Reeva S. Simon (eds.). 1991. *The origins of Arab nationalism*. New York: Columbia University Press.
- Leonti, Marco & Laura Casu. 2013. Traditional medicines and globalization: Current and future perspectives in ethnopharmacology. *Frontiers in Pharmacology* 4. DOI: <https://doi.org/10.3389/fphar.2013.00092>.
- Maddy-Weitzman, Bruce. 2011. *The Berber identity movement and the challenge to North African states*. 1st edn. Austin, Tex: University of Texas Press.
- Mearns, David L., David Parham & Bruno Frohlich. 2016. A Portuguese East Indiaman from the 1502–1503 Fleet of Vasco da Gama off Al Hallaniyah Island, Oman: An interim report. *International Journal of Nautical Archaeology* 45(2). 331–350.

- Miller, Catherine. 2003. Linguistic policies and language issues in the Middle East. In Akira Usuki & Hiroshi Katō (eds.), *Islam in the Middle Eastern studies: Muslims and minorities* (CAS Symposium Series), 149–174. Osaka: Japan Center for Area Studies.
- Morris, Miranda J. 1983. Some preliminary remarks on a collection of poems and songs of the Batahirah. *Journal of Oman Studies* 6(1). 129–144.
- Morris, Miranda J. 2016. *The documentation of Modern South Arabian languages: Bathari*. <http://hdl.handle.net/2196/00-0000-0000-000F-0429-1>. (20 November, 2021).
- Morris, Miranda J. 2017. Some thoughts on studying the endangered Modern South Arabian Languages. *Brill's Journal of Afroasiatic Languages and Linguistics* 9(1). 9–32.
- Morris, Miranda J. Forth. *A collection of Baṭḥari texts*.
- Morris, Miranda J. & Fabio Gasparini. Forth. *A descriptive grammar of Baṭaḥrēt (Baṭḥari)*.
- Moseley, Christopher (ed.). 2010. *Atlas of the world's languages in danger*. 3rd edn. Paris: UNESCO Publishing.
- Owens, Jonathan. 2007. Close encounters of a different kind: Two types of insertion in Nigerian Arabic codeswitching. In Catherine Miller, Enam alWer, Dominique Caubet & Janet C. E. Watson (eds.), *Arabic in the city. Issues in dialect contact and language variation*, 249–274. London: Routledge.
- Peterson, John E. 2004. Oman's diverse society: Southern Oman. *The Middle East Journal* 58(2). 254–269.
- Romaine, Suzanne. 1989. Pidgins, creoles, immigrant and dying languages. In Nancy C. Dorian (ed.), *Investigating obsolescence: Studies in language contraction and death*, 369–383. Cambridge: Cambridge University Press.
- Smolicz, Jerzy J. 1980. Language as a core value of culture. *RELC Journal* 11(1). 1–13.
- Suleiman, Yasir. 2013. *Language and identity in the Middle East and North Africa*. London, New York: Routledge.
- Tabouret-Keller, Andrée. 2017. Language and identity. In Florian Coulmas (ed.), *The handbook of sociolinguistics*, 315–326. Hoboken: John Wiley & Sons, Ltd.
- Thomas, Bertram. 1929. Among some unknown tribes of South Arabia. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland* 59. 97–111.
- Thomas, Bertram. 1937. Four strange tongues from central south Arabia - the Hadara group. *Proceedings of the British Academy* 23. 231–331.

- Tilmatine, Mohand. 2015. Arabization and linguistic domination: Berber and Arabic in the North of Africa. In Christel Stolz (ed.), *Language empires in comparative perspective*, 1–16. Berlin: De Gruyter Mouton.
- Tosco, Mauro. 2015. Arabic, and a few good words about empires (but not all of them). In Christel Stolz (ed.), *Language empires in comparative perspective*, 17–39. Berlin: De Gruyter Mouton.
- Watson, Janet C. E. & Abdullah M. alMahri. 2017. Language and nature in Dhofar. «*Quadri*» – *Quaderni di RiCOGNIZIONI* 7. 87–103.
- Zaborski, Andrzej. 1997. Minority languages in Arab countries. In Brunon Synak & Tomasz Wicherkeiewicz (eds.), *Language minorities and minority languages in the changing Europe*, 285–294. Gdansk: Wydawnictwo Uniwersytetu Gdańskiego.

Chapter 6

Creating a class of elite Palestinian multilinguals in Israel: Reflections on research in late capitalism

Nancy Hawker

Independent researcher

In this chapter, speeches from 2015 and 2022 by Palestinian politicians in Israel using Hebrew and Arabic are analysed for pragmatic functions, with a focus on politeness strategies, translation, humour, and inter-group boundaries. Language ideologies are related to the equation of one nation, one language, which speakers reference while practising multilingualism. The multilingual code is articulated with the emergence of an educated middle class that aspires to consumerism and liberal politics. In the context of late capitalism, it is the commodification of multilingual skills that circumscribes political aspirations in combination with considerations of class and conflict.

1 Introduction

Past research on Palestinians who live in Israel, and speak both Hebrew and Arabic, has shown that they are interactively guided in their language choices by certain tacit norms, namely

1. the principle of Arabic avoidance outside the in-group ('only speak Arabic with Arabs'); and
2. the principle of valuation of multilingualism ('display your multilingual assets') (Hawker 2019).



There are also explicit norms that are policed whenever Palestinian speakers speak Arabic on institutional state platforms that contribute, in a daily bureaucratic way, to the identification of Israel with Zionism (Handelman 2020). Here, Zionism is used in a narrow definition, as a specific aspect of an ideology that justifies and promotes particular political actions: a combination of ideas repeatedly put to practice in state-building and state-affirming rituals and symbols constituting Israel as the state of people who claim Jewish nationality, following 19th- and 20th-century European ideas of nation-statehood.¹ These norms are based on Hebrew dominance on Zionist platforms, and include

1. minimising the audibility of Arabic in Israeli institutions;
2. announcing and preparing the audience for short switches to Arabic; and
3. translating any Arabic speech immediately into Hebrew.

The task of this chapter is to trace the manifestation, negotiation, and contestation of these norms, and to see whether they are still relevant, using observations and analyses spanning several years. The chapter will argue that the ability to challenge norms increases with the speakers' relative social power (at intersections of class and other identities). The issue of relative power opens up specific platforms for Palestinian politicians in the Knesset (Israeli parliament), since they are institutionally given the floor to speak for set periods of time and gain confidence in their discursive skills by drawing on their multilingualism.

I will argue that within this generally elite circle, relative power differentials are sociolinguistically manifest in expressions of politeness, which organize social hierarchy. Class, as an economy-based form of social hierarchy, maps onto understanding the status of the Palestinian minority in Israel either as systematically discriminated against, or economically disadvantaged. Simply put, if the frame is discrimination, the speaker can politically side with the oppressed by using negative politeness. If the frame is disadvantage, the speaker can politically align with the privileged by using positive politeness. And yet all the speakers, regardless of their political politeness strategy, display their multilingual assets on a first order of indexicality, as a correlate of their middle-class education and their aspirations for economic opportunity and political engagement.

The evidence to test this argument is taken from audio-visual recordings and stenographer's minutes of Knesset debates. The selection of texts from 2015 and

¹This definition of Zionism is often qualified as secular political Zionism, distinguished from cultural and spiritual Zionism, and from Jewish religion (Raz-Krakovitzkin 2021: 34 in Rouhana & Shalhoub-Kevorkian 2021).

2022 highlights instances where there was extensive metalinguistic commentary on the use of Arabic on this platform, constituted as Zionist (in the narrow, nation-state-building sense). The analysis of this evidence centres on strategies of politeness, translation, and identity-definition, created interactionally. Since these pragmatic functions of speech can only be found in interactions, three long stretches of speech are reproduced as evidence, in my own idiomatic English translation.

1.1 Class in sociolinguistics

The idea that, in late capitalism, the service- and knowledge-based economies predispose speakers to repackaging language skills as economic and political resources will be expressed in the proposition that language itself becomes a ‘means of production’ (Heller & McElhinny 2017). The emergent Palestinian middle class in Israel are the owners of multilingual discursive means of production and, consequently, are relatively less exposed to precarity in a structural sense than those who do not own those resources. This conceptualization is different from the treatment of class as a habitus that forms language ideologies enabling judgments of taste (Bourdieu 1982). Nevertheless, these ideologies, which bear the scars of past (class) struggles over value (Myles 1999: 892), do affect the success of commodification and commercialization. The marketplace for Palestinians’ and other Arabs’ multilingualism is already skewed towards devaluing Arabic speakers’ Arabic. The linguistic offerings of late capitalism are ambivalent; multilingualism has cultural capital, but in a commodified way, geared towards individual material benefits rather than economic emancipation, freedom, and social justice (Davis 2008).

By the time of the 2015 elections, the second generation of Palestinians and other Arabs socialized under Israeli systems were in their prime, and the third generation had come of political age. The Palestinian and other Arab middle class in Israel has developed as a correlate of the neo-liberalization of the economy since the 1980s. With this limited middle class came the gradual emergence of a new voice: an Arabic and multilingual voice, claiming equality in the common good in the space that is controlled by the Israeli state. They deploy multilingual styles (Eckert 2004) appropriate to their educated middle-class in-group, for legal argumentation, humour, critical self-reflection, and sometimes earnest anger to call Israel’s bluff on its self-description as a democracy. These styles also function to stretch claims to equal rights from the strategies to survive in the territorial limits of Israel (as it was under military rule) and to be protected from precarity (as it was in the years of manual labour) to all aspects of liberal freedoms. These freedoms appeared to be within reach – such are the expectations that the

middle class is trained to see as a right (Dean 2014). This voice is skilled in its use of multilingualism and language-contact phenomena for different pragmatic functions, and co-constructions of political meanings.

I have commodified my own multilingualism for precarious employment in late-capitalist knowledge work, and I self-consciously display my skills in this volume, as does every contributing author. I am not of the Palestinian or Israeli middle class, but am perhaps of the cosmopolitan European middle class (Hartung et al. 2017; Nowicka & Rovisco 2016: 21–86). I currently work for Amnesty International in the Middle East and North Africa Office, which is an organization that campaigns for human rights and takes the inherent equality and dignity of all humans as a premise beyond question. The sociolinguistic analysis of the multilingual political discourse of middle-class Palestinians who claim equality is therefore, through several degrees of separation, also the analysis of my own discourse. Nevertheless, the execution of the analysis might be as scientific as any science (Bourdieu 2004: 76–78).

1.2 Method of data collection and selection

The evidence I present in Section 2 of this chapter is selected from recordings and observations at 25 locations across Israel during the 2015 parliamentary election campaigns. The recordings involved 18 Arabic-speaking politicians, 4 Hebrew-speaking politicians, aides, vocal supporters/critics, and members of the public. The locations included a range of situations from house visits to mass events. I also analysed pre-recorded sources, mostly on YouTube or the Knesset Channel, of Knesset debates and TV broadcasts. In total, I coded and analysed 57 hours of recordings, identifying switches between Hebrew and Arabic and their contexts. For a historical perspective, I examined 21 documents from the Knesset archives. These documents were official records of parliamentary and other party political debates, where Arabic was spoken in a Hebrew-dominant context, totalling more than 500 pages (Hawker 2019). I add new evidence to this in Section 3 from political speeches in 2022 – this is important due to contextual changes and changes in the styles of multilingual discourse. I chose the textual examples in Sections 2 and 3 as emblematic illustrations of general patterns that emerged in this research.

2 Arabic avoidance and Arabic promotion in Hebrew-dominant contexts: Contradictions?

On 23 December 2015, Ayman Odeh, a Palestinian politician in Israel, made a speech ahead of Christian and Muslim holidays, in which he switched from He-

brew to Arabic for a few sentences. Ayman Odeh leads the Joint List, which is a coalition of parties formed in 2015 that mostly attract the votes of Palestinian and other Arab citizens of Israel. The text analysed below was taken from the official Knesset record from 23 December 2015 (Knesset 2015: 130–131); the translation is my own. The Knesset records are all noted in Hebrew, in line with the Rules of Procedure (Knesset 2012), and occasionally – as seen below – in Arabic transcribed into Hebrew script. As in all examples for this chapter, what was Hebrew in the original (in this case, in the stenographer’s official record) is presented in italics, while Arabic is in roman font. There is no audio record of this event – that I could find – against which to compare the stenographer’s record. Therefore, the difference in the font as a language identifier for the translation is based on my recognition of how the Knesset stenographer would have distinguished the two languages as separate and bounded, providing the stenographer would have followed widespread attitudes in the chronotope of modern Israel/Palestine (Shohamy 2005: xii–xiii). The stenographer’s metalinguistic comment “speaks in Arabic; henceforth translation” is underlined. Whatever speech has been translated from Arabic for the records is enclosed in parentheses, as if literally bounded. Overlapping and inaudible speech – which connotes interruptions or heckling – is marked in the record with hyphens, reproduced here in the same way. The names of the speakers and their parliamentary parties are in bold and parentheses, as they are in the original; all of the speakers are Jewish citizens of Israel, apart from Ayman Odeh, who is a Palestinian citizen of Israel.

- (1) *Ayman Odeh wishing everyone happy holidays in the Knesset, 23 December 2015.*

Ayman Odeh (the Joint List):

*I have a request from the Chair, sir. Would you please
allow me to speak in Arabic for a minute, because today is
the Day of the Birth of the Prophet Muhammad, peace be
upon him, and the day after tomorrow is Christmas - - -*

Rachel Azariah (Kulanu):

But today is my birthday - - -

Ayman Odeh (the Joint List):

*- - - so with your permission, I would like to turn to my
people - - -*

Shelly Yachimovich (Zionist Camp):

- - - You don't need to ask for permission - - -

Ayman Odeh (the Joint List):

- - - and tell them - - -

15 **Shelly Yachimovich (Zionist Camp):**

You don't need to ask for permission.

Chair Yehiel Hilik Bar (Zionist Camp):

It is an official language according to the Rules of Procedure.

20 **Ayman Odeh (the Joint List):**

- - - I turn to my compatriots: (speaks in Arabic; henceforth translation: I turn to my compatriots, to our people, all our people, on the occasion of the Day of Prophet Muhammad's Birth and of Christmas. It is said, "A light shone in the
25 eastern sky, shining brightly in the dark night," and there are some who add, "A light shone in the eastern sky, to guide the wise men to the manger." Happy holidays to all our people, for the Prophet Muhammad's Birthday and for Christmas, and to anyone who gives for us to continue
30 struggling for a life in dignity in the land of our ancestors.)
In relation to the proposed bill - - -

Ksenia Svetlova (Zionist Camp): - - -

Ayman Odeh (the Joint List):

35 Thank you very much. Happy holidays to everyone. Happy holidays to everyone.

Shelly Yachimovich (Zionist Camp): - - -

Ayman Odeh (the Joint List):

Thank you, thank you, Shelly, bless you. [...] With regards to the proposed bill - - -

40 **Omer Bar-Lev (Zionist Camp):**

Tell us what you said.

Ayman Odeh (the Joint List):

Basically, what I said was, happy holidays to everyone, we shall continue to struggle together for life in dignity in our
45 historical homeland.

Anat Berko (Likud):

What "historical homeland"?²

²In the record: איזה היסטורית! (literally, "what historical!"), own translation.

Ayman Odeh (the Joint List):

Alright. So, we - - -

50 **Shouts:** - - -

Any length of speech in Arabic in the Knesset is a remarkable event, at least since the mid-1960s (Hawker 2019: 36–39).

Some of the Jewish interlocutors in the Knesset appear to be largely comfortable with the liberal tolerance of different religious holidays, multilingually expressed in short formulaic greetings. Ksenia Svetlova and Shelly Yachimovich respond to Ayman Odeh, possibly in Arabic and hence unrecordable to the stenographer (lines 32 and 36). In an ordinary context such as a short commercial interaction among strangers, such expressive speech acts would have the function of phatic communion and politeness (Amer et al. 2020). In the context of the Knesset, however, greetings in Arabic and across religious dividers are exceptional events. The speakers are enacting a liberal vision where citizens of all religions would be accorded mutual respect. This vision is what Palestinian and Jewish intellectuals have termed a hoped-for future *convivencia* (Shohat 2017: 24, 33), or conviviality. My interpretation of Ayman Odeh's staging of the performance of multicultural tolerance analyses it as a declarative speech act intended to bring about this liberal transformation. The multilingual exchange, in which Arabic phrases are immediately translated by an authoritative Palestinian speaker into Hebrew, is the discourse chosen for this performance.

However, the speech act presents a fluid, not to say ambiguous, relationship between religious identities and secular aspects of nationhood such as “ancestors” (line 30) and “homeland” (line 45). It becomes clearer, in Ayman Odeh's repetition and summary in Hebrew of what he had said in Arabic, that he sees religious holidays as elements of inherited traditions that are specific to Palestinians (lines 44–45). Further, this inheritance grows into political rights situated in contemporary “struggles”. The ambiguity allowed some interlocutors to react cooperatively to the religious element, and others confrontationally to the secular national one.³

While Ayman Odeh is aware that Arabic and Hebrew language policies strongly index one or the other national project (Ben-Rafael 1994; Suleiman 2019: 27), and indeed he relies on this order of indexicality, his practices are in fact multilingual. He is manipulating the indexicality for a more complex political message than

³The context for the linking of religion and politics is the fluid relationship between Judaism and secular national rights to land in Israel/Palestine. This relationship is structured in laws, state and parastate institutions, and symbols (Yadgar 2020: 86–101), and enables discourses of Jewish exclusivity in Israel (Rouhana & Shalhoub-Kevorkian 2021).

the equation of one nation, one language, and the logic of conflict that derives from this equation being applied twice in the same geographical space (Suleiman 2004). My argument is that on a higher indexical order, he is enregistering a multilingual code that he has mastered and co-created as member of an elite Palestinian class in Israel. I will now turn to the detailed analysis of the code-switching involved in this multilingual discourse in Example (1) .

Ayman Odeh's switch from Hebrew to Arabic is highly conspicuous, and issues of translation are also highlighted. He opens his speech from the speaker's dais with a request directed at the chair for permission to speak Arabic. He justifies this request on the basis that he is addressing "his" people, the Palestinians who celebrate the Prophet's Birthday and Christmas. What is significant is that he feels the need to justify his actions. The announcement of the switch to Arabic is interrupted first by an incongruous joke from Rachel Azariah and then by a legalistic framing from Shelly Yachimovich. I will analyse each element: the announcement, the joke, the legalistic framing, and the request for translation.

Announcing a switch from Hebrew to Arabic, when a Palestinian Arabic speaker is in the presence of Jewish Israeli interlocutors, is the norm in all my findings (Hawker 2019: 63–87). Conspicuous announcements, accompanied by markers of negative politeness, occurred in all cases that I recorded and observed where Palestinian speakers spoke Arabic on public platforms where at least part of the audience was Jewish, as in this case, or in small meetings where both Palestinian and Jewish citizens of Israel were present. In fact, in three instances, it was sufficient that only one person in the audience was Jewish Israeli among Palestinian citizens of Israel, all of them with knowledge of both languages, for a switch from Arabic to Hebrew to be announced. The way I have summarized these findings is that there is a norm among Arabic speakers in Israel to avoid speaking Arabic to anyone who does not want to be identified as an Arab in that interaction. It is a norm that has evolved in the habitus of colonized experiences since 1948 (Pappé 2011), resulting in the containment of Arabic within safely identified in-groups and the suppression of Arabic in any situation where there might be Jewish Israelis who do not identify as Arab. Violations of that norm – that is, directing any Arabic speech, even as an honest mistake, towards a person who does not identify as Arab – therefore must be justified in conspicuous announcements. Here, the announcement specifies that the addressees are Palestinians identified, and in fact constituted, as Arabs who are indexed by their language. The norm of avoiding Arabic outside the in-group is therefore not violated, despite the immediate Hebrew-speaking context; it is reinforced by the announcement.

Defiance is another aspect of the announcement of the switch to Arabic, and is mitigated by extensive negative politeness formulae, such as "would you please

allow me” and “with your permission” (lines 2–3 and 9). Announcing the switch in an apologetic manner acknowledges another challenge to established norms: the speech from the Knesset dais breaks the doxa of Arabic silence on Zionist platforms. Since approximately 2010, Palestinian politicians in Israel have started to claim Arabic audibility in public institutions where Hebrew monolingualism dominates (Hawker 2019: 37–52). These claims have taken the form of short rhetorical declarations in Arabic, immediately translated into Hebrew by the speakers themselves, as in this case. Usually, these declarations are met by metalinguistic comments that protest against Arabic audibility. Ayman Odeh is possibly pre-empting the protests by attending to his Hebrew-speaking interlocutors’ face (Eelen 2001).

The idea that it is ‘polite’ to speak in the dominant language to interlocutors outside the suppressed in-group is reminiscent of the sociolinguistic situation in Catalonia during its emergence from Franco’s dictatorship (Woolard 2012). Therefore, Ayman Odeh is being especially polite about being impolite; his request for permission both accepts the dominance of the Hebrew-speakers (which is ‘polite’) and asserts the right to speak Arabic (‘impolite’) in the face of the expected objection. The use of negative politeness strategies in particular create the layout of the conversation – negative politeness does not seek common ground between the interlocutors, but respects the distance and hierarchy between them (Eelen 2001: 25–29). The Hebrew-speaking Jewish Israeli chair of the Knesset plenum is not ‘in’ the in-group that Ayman Odeh is addressing, and this is reinforced by the announcement and the apologetic request. Both norms – the avoidance of Arabic in mixed company and the silencing of Arabic on Zionist institutional platforms – are equally restated and challenged.

The protest against the switch to Arabic comes in the form of a joke. Rachel Azariah’s joke is a kind of racist joke (Weaver 2016). The implication of saying, in effect, “If Muhammad and Jesus’s birthdays are worth speaking in Arabic, then my birthday is worth speaking in Hebrew,” (line 7) is the depreciation of the value of the Muslim and Christian holidays, because Rachel Azariah’s importance to herself and those around her is not of the same public nature as that of Muhammad and Jesus. This type of humour organizes affiliations – who is the object of the joke and who is the intended audience (Meyer 2000). It is not clear from the official record how well the joke is received; there is no indication of laughter, but there is no objection to it either. This particular joke might not have passed the thresholds of humour or offensiveness (Kuipers 2016).

There was a more effective protest, however, against Ayman Odeh’s politeness strategy: “You don’t need to ask for permission,” Shelly Yachimovich said twice, interrupting him (lines 12 and 16). The chair intervened to validate the legalistic

rule that is brought to bear against the tacit norm of silencing any Arabic speech: “It [Arabic] is an official language according to the Rules of Procedure” (lines 18–19). In fact, the letter of the Knesset Rules of Procedure does not mention Arabic as an official language, but rather as a second one.

The sittings are conducted in Hebrew. Members of Knesset have the right to speak in Arabic as well; speeches in Arabic shall be translated into Hebrew. Only when a guest from abroad delivers a speech in the Plenum are speeches in different languages optional therein. (Knesset undated[b])

The chair had to reiterate the rule allowing Arabic speech because an objection to Arabic was expected. Shelly Yachimovich neutralized that prejudiced objection before it could even materialize. “You don’t need to ask for permission [to speak Arabic]” was an assertion that Israeli politics is governed by rules, not prejudices. Shelly Yachimovich was defending the republican structure of the state institutions against the accusation, implicit in Ayman Odeh’s excessive politeness yet perfectly pragmatically understood, that Hebrew is dominant and Arabic suppressed. At the time of this speech, and until July 2018, Arabic was the second official language of Israel. Its demotion to “language with special status” in 2018’s Jewish Nation-State Law was a legal materialization of the prejudice that transformed the republican structure.

The republican structure informed the decision that all Knesset records be kept in Hebrew (Knesset 2012: articles 22 and 38).⁴ The only exceptions are short interjections transcribed into Hebrew script such as “happy holidays to everyone” and “thank you”, which remain in Arabic (lines 34 and 38). Professional interpreting was discontinued in the Knesset plenum in the 1960s (Hawker 2019: 32–37). The vagueness of the passive voice in the Rules of Procedure – “speeches in Arabic shall be translated into Hebrew” – leaves responsibility for the delivery of translation open. The possibility of a need to translate into Arabic, for the benefit of Arabic-speaking citizens of Israel, is not considered. Translation into Hebrew for the stenographer’s minutes happens separately from interpretation in the plenum, and therefore Omer Bar-Lev needed to explicitly ask Ayman Odeh to consecutively interpret his own words (line 41). Ayman Odeh’s translation is a summary which reveals the political message: “Basically, what I said was, happy holidays to everyone, we shall continue to struggle together for life in dignity in our historical homeland.” This revelation triggers unspecified “shouts” or heckling (line 50), which are not retrievable in the record. Shelly Yachimovich and

⁴The only speech recorded in Arabic in the Knesset minutes is the speech by Anwar Sadat on 20 November 1977, as he was a foreign guest (Knesset 1977).

the Chair Yehiel Hilik Bar were using rules and procedures to shore up the republican structure against precisely this type of communication breakdown. The breakdown was triggered by the Jewish Israeli ethnonationalist objection (“What ‘historical homeland’!”; line 47) to a Palestinian claim to life in dignity in their country. The content of the “shouts” can only be inferred from Ayman Odeh’s response to them. Ayman Odeh’s response centred on language, rather than on politics or religion.

After two more turns of heckling, Ayman Odeh said the following, in Hebrew-only speech:

That Arabs know Hebrew and most Jews don’t know Arabic is an added value to Arabs and not to Jews. Historically, Jews learnt more languages than other peoples. It just so happens that here, here specifically, the Arabs know more languages than Jews do, and that’s an added value in their favour and not in the Jews’ favour. I would like to say that knowing more than one culture enriches a person, it adds to a person, it allows a person to understand the history of another people, and to relate to their pains and longings. (Knesset 2015: 132–133; own translation)

The shouts and objections to multilingual discourse and its political message are countered by a spontaneous defence of language learning. I argue that this justification reads as a manifesto for the multilingual code of the emerging Palestinian middle class (Ghanem 2016). They are turning their multilingualism, once a burden of a minority contained by the state’s rules, into cultural capital (Block 2013). In turn, this cultural capital is a means of (self-)production in the context of late capitalism – they are producing their own material success as a class. It is a manifesto for a liberal future that the middle class see as their right (Dean 2014), but that sounds radical to the Knesset audience of 2015. The inaudible shouting in the debate was the rumblings of ethnonationalism (Peled 1992), overrunning the republican rules that permitted, within constraints, performances of radical visions of a liberal future.

Addressing Arabic only to Arabs, and displaying multilingualism on Zionist institutional platforms in Israel, was therefore not contradictory. The liberal political messages conveyed by the styles of the multilingual speech were in dynamic relation to two other language-ideological positions. One was ethnonationalism, which demanded the strict adherence to the equation of one nation, one language, with Israel being notionally the preserve of the Jewish Israeli nation, and the other was republicanism, which carefully managed the diversity of a linguistic minority by establishing rules, and heaped the duties of translation and language learning on the minority (Gal 2012).

3 New evidence from 2022: Palestinians enter the establishment?

The Knesset record for 23 December 2015 contains 531 pages of stenographer's minutes. Using the method of identifying Arabic passages thanks to the stenographer's metalinguistic comments, I was able to count that Arabic speech on that day added up to three of those pages, or less than 1 per cent of the day's record. Fast-forwarding to 4 January 2022 reveals a different picture. The record for that day amounts to 289 pages, of which 33 are in Arabic, and five more pages constitute a debate about the inadmissible audibility of Arabic in the Knesset and its translation, resulting in 13 per cent of the day's record. In this chapter, I conduct a time-lapse comparison. In 2015, Ayman Odeh spoke Arabic in a way that previously would have been unthinkable; seven years later, another significant event was recorded, begging the question whether the multilingual liberal future has indeed come about.

One of my data-collection methods since the 2015 election campaigns has been to follow the YouTube, Twitter, and Facebook accounts of Palestinian members of the Knesset. For the 2020 Israeli elections, I also conducted field visits on electoral meetings (Hawker 2020). When the speakers code-switch, I pay heightened attention. Usually, the multilingual Palestinian members of the Knesset speak Arabic to their Palestinian audiences, and Hebrew to their Jewish Israeli audiences and mixed audiences (such as the Knesset). This is what I would expect, in line with the two norms presented above: (1) the avoidance of Arabic in mixed company, and (2) the silencing of Arabic on Zionist institutional platforms. However, a video on the Facebook page of The Fans of Dr Mansour Abbas caught my eye, as it violated the second norm.

In the two-minute video, Palestinian members of the Knesset Mansour Abbas and Ahmad Tibi were having a debate on the speaker's podium in the Knesset, in Arabic, with no heckling or other interruptions, no request for permission, and no self-translation (Facebook 2022). In the Knesset minutes for that day, I added up everything that followed the stenographer's comment "speaks in Arabic; henceforth translation" in the parentheses denoting Arabic speech in Hebrew translation. I counted Arabic phrases transcribed into Hebrew script, without professional translation for the record. I noted metalinguistic comments by Knesset members, and the total added up to the 38 pages mentioned above. The video of the Abbas-Tibi exchange corresponds to page 58 of the Knesset minutes, in the middle of 27 pages of nearly uninterrupted Arabic (recorded in its official Hebrew translation; Knesset 2022c: 45–72). Owing to spatial constraints, I cannot

reproduce the entirety of the Arabic segment here, as I was able to for Example (1). This abundance in itself was astonishing.

I selected the data presented in this section based on the themes emerging from the analysis of the 2015 speech. The speeches asking for permission to speak Arabic and the accompanying politeness strategies, the protests against Arabic audibility and requests for translation into Hebrew, references to explicit institutional rules and to tacit norms, and usages of Arabic that define and redefine identities are all presented below. On the basis of this new evidence from 2022, I will argue strongly that the norm that had suppressed Arabic in the Knesset has been broken. The context for this change is that for the first time in Israeli history, a political party for which only Palestinian citizens of Israel vote,⁵ and that claims to represent the interests of that social group in particular, was included in the government coalition formed on 13 June 2021. The Joint List split in two in January 2021, and the Islamist party (named United Arab List) led by Mansour Abbas entered the government coalition. The rest of the Joint List, still led by Ayman Odeh, went down from 15 seats to 6 in 2021 (out of the 120 total in the Knesset) and remained in the opposition. The Arabic debate in the Knesset captured in the video was therefore a debate between a representative of the Israeli government and of opposition to it. That political divide, too, was astonishing.

I will also argue, equally strongly, that more (self-)reflection is needed in the field of language-contact studies. I have been arguing for a number of years that looking at contact between languages that index nationalities that are in apparently zero-sum conflict solely through the prism of that conflict entails the omission of important areas of contact such as economic activity, which cannot be reduced to inter-national conflict (Hawker 2018; Heller et al. 2015). Here is where the attention to class comes into play. Socioeconomic class intersects with nationality (and other identities) to create more subtle nuances of language indexicalities, in dynamic relation to the central equation of one nation, one language.

Nevertheless, the new evidence challenges some aspects of my earlier argument about the enregisterment of multilingual discourse as a marker of Palestinian middle-class public positioning, embraced with increasing confidence. The challenge is that the material aspects of class, and the link these aspects have to relative power, may play a greater role in destigmatizing marginal groups' discourse, and in enregistering new code patterns, than any other identity marker I had previously considered. It is not a liberal utopia come true; it is the late-capitalist materialism trump card on the table. The topic of debate on 4 January

⁵Though ballots are anonymous, the list of voting preferences by location on the Central Elections Committee website (<https://votes24.bechirov.gov.il/>) indicates that most voters for the Islamist party live in Palestinian towns and villages in Israel.

2022 was an amendment to a law regarding municipal planning, concerning approving the connection of marginalized Palestinian and other Arab communities in Israel to the electricity grid. The amendment was referred to as the “Electricity Law” (Knesset 2022b).

3.1 Permission and politeness

In Example (1), negative politeness strategies were a way of announcing a violation of the Arabic silencing in Israeli state institutions, and a way of highlighting distinctions of hierarchy that placed Hebrew and Hebrew speakers above Arabic and Arabic speakers in the Knesset. The announcement of switches to Arabic in Example (2) (all taken from the 4 January 2022 Knesset stenographer’s minutes, own translation) uses positive rather than negative politeness.

(2) *Announcing switches to Arabic in the Knesset, 4 January 2022.*

- a. **Walid Taha (United Arab List):** [after one and a half pages of minutes of his speech in Hebrew]

(Speaks in Arabic; henceforth translation: To the Arabs in society I say, the discussions around the Electricity Law have started now, and the Opposition with the exception of the Joint List have left the plenum, and I do not know if they will come back again to vote tomorrow morning or not. [...]) I want to thank the members of my committee again. I assume that you understood everything I said in Arabic. [...]

Chair Mansour Abbas (United Arab List):

But certainly the Minister Eli Avidar understood every word.

Polite attention to the Jewish Israeli interlocutors who might not have “understood everything [that was] said in Arabic” came in Hebrew at the end of the Arabic section, and there was no translation. Instead of negative politeness, there is a strategy of inclusion, or positive politeness. Mentioning “my committee” – Walid Taha was then the head of the Internal Affairs Committee, where most members are Jewish citizens of Israel – and making assumptions about their knowledge, is the opposite strategy to that of Ayman Odeh’s in Example (1), where he reinforced distance, hierarchy, and respect through negative politeness. In both cases, the effect of the politeness strategy is to mitigate the ‘impoliteness’ of using Arabic to address interlocutors who do not wish to be addressed as Arabs, as per the norm of Arabic avoidance in mixed company. Positive politeness strategies say, effectively, “We are in a team together, so you won’t be offended by the way I

speak.” Mansour Abbas furthered this strategy by including Cabinet Minister Eli Avidar in the team. Eli Avidar’s Knesset profile mentions his knowledge of Arabic (Knesset undated[a]). The pattern of launching into Arabic without preliminaries is then repeated by several Palestinian speakers in the Knesset (Knesset 2022c: 46–51). The Arabic speeches are interspersed with Hebrew speeches of similar length, from the same multilingual speakers. The pattern appears to be that when the speech is specifically targeted at fellow Palestinian Knesset members or Palestinian citizens of Israel beyond the immediate audience, Arabic is resorted to. Hebrew is for everyone.

(2) b. **Ayman Odeh (Joint List):**

Honourable Chair, my fellow members [masc.] and members [fem.] of the Knesset, allow me please to speak in the Arabic language, even though I am in favour of minority rights, including collective rights, and the Jews are in the minority here; but nevertheless, allow me this time.

Minister in the Prime Minister’s Office Eli Avidar:

Ayman Odeh (Joint List):

Listen, you are Jewish, but you are Egyptian, a speaker of Arabic.

(Speaks in Arabic; henceforth translation: I want to speak in Arabic out of a desire to hide. On the principle of “if you leave it in the heart it will torment, and if you pour it out it will embarrass”. So it’s better for us to speak to each other in the Arabic language, because there are some matters that are important. The purpose of these matters is not to air our dirty laundry, but rather to talk directly with our Arab public. [...])

Ayman Odeh returns to negative politeness strategies (“allow me please”), and the languages are again referenced as indexes of nationality: Hebrew for Jewish Israelis and Arabic for Palestinians. Normally, the Palestinians are in the minority, and Ayman Odeh is in the position of having to lobby for their rights. Yet at this session, most Jewish Israeli members of Knesset had left the room, and so the Palestinians were in the majority, and Ayman Odeh joked about his political principles of standing up for minorities. He was hoping that this joke would organize affiliations, as self-deprecating jokes tend to do, by bringing the audience together in alliance with him (Meyer 2000: 318–319). Later, he built on this idea of alliance by engaging in positive politeness, reinforcing the common ground, by

saying that Eli Avidar – who was born in Alexandria – was Egyptian and therefore included in the in-group that speaks Arabic. On the basis of speaking Arabic to Arabs, albeit from the Knesset dais, he then launched into Arabic for a substantial discussion – not for short greetings or slogans, but to agonize over the divisions in Palestinian society – which would have been unthinkable in 2015.

The Palestinian speakers found themselves in an unprecedented situation: they had one Knesset session to sort out political differences internal to Palestinian society in Israel, with the government – including Israeli right-wing and far-right parties, among them Eli Avidar's Yisrael Beiteniu – in some sort of agreement with the Islamists, and the opposition, Likud, out of the room. The issue of connecting excluded communities to the electricity grid, including 100,000 Bedouin living in officially unrecognized villages in the Negev/Naqab, is a human rights problem relating to the right to a decent standard of living (Negev Coexistence Forum 2020). The Palestinian representatives in the Knesset had been lobbying for years for these people's rights. However, the two Palestinian parties were now in disagreement over the framing: was the issue one of discrimination against a national minority (as the Joint List presented it in the debate), or was it to do with socioeconomic disadvantage (as the Islamists presented)? For those who claimed discrimination, negative politeness made sense; the powerful group that engaged in discrimination was kept at a respectful distance. For those who claimed disadvantage, positive politeness was apt, since the group which had a socioeconomic advantage, including the Palestinian Islamist politicians, was able to offer help to those in need.

(2) c. **Aida Touma-Suleiman (Joint List):**

[...] The ones who had prevented the planning and delayed the approval of the comprehensive plans were the state, the various governments, and the Ministry of the Interior. And why am I emphasizing this? Because the discourse must not descend to the level of: "We are willing to accept the narrative and the story and the discourse of those governments that oppressed the Arab population and deprived it and take the reckoning on ourselves, according to this discourse."

Chair Mansour Abbas (United Arab List):

Says in Arabic: Speak in Arabic, Aida. Everyone here is Arab.

Aida Touma-Suleiman (Joint List):

(Speaks in Arabic; henceforth translation: I speak in Arabic and I speak in Hebrew, and I was just going to switch to Arabic to try to understand this law accurately. [...])

Aida Touma-Suleiman gave her speech in Hebrew, though she claimed the “us” of the oppressed Arab national minority. She reiterated the framing of the electricity issue as one of discrimination. The Chair invited her to speak in Arabic, not on the basis of the official legalistic rule, but on the basis of the tacit norm: you can speak Arabic in the in-group, providing everyone wants to be addressed as Arabs. Mansour Abbas reassured her on the point of identification of interlocutors. Her reaction echoed the pro-multilingualism manifesto of the emergent Palestinian middle class: knowing several languages is cultural capital. Speaking and analysing in two languages allowed Aida Touma-Suleiman to “understand this law accurately”. There was no issue of politeness. Negative politeness is only for addressing the overlords, to let them know that you resent their dominance.

3.2 Protests against Arabic audibility

All is well, then: Palestinians can speak Arabic to Palestinians in the Knesset now? Not quite, since the lack of institutional translation betrayed a problem regarding the identity of Israeli state institutions, and how languages spoken therein create that identity. The problem of translation soon became manifest as protests against the very audibility of Arabic on 5 January 2022, as the Knesset session on the Electricity Law continued into its second day. The day started with Itamar Ben-Gvir, representative of the Jewish Power party (a Jewish supremacist organization), calling Walid Taha (United Arab List) a “terrorist”, and receiving a caution from the Chair Mickey Levy (Yesh Atid, a centrist secularist party). Walid Taha, as the sponsor of the bill, then explained the proposed law in Hebrew (Knesset 2022c: 82–89), and was interrupted multiple times by Jewish Israeli Knesset members of the National Religious party, the Likud party, and Shas, a party that represents Orthodox Jews of Middle Eastern and North African heritage. After yet another interruption, Walid Taha switched to Arabic, without preliminaries. The debate soon broke down, with multiple heckles hindering scheduled speeches, until Mansour Abbas took over chairing, and managed to return the debate to its apparently normal course, in Hebrew.

- (3) *Asking for translations into Hebrew in the Knesset, 5 January 2022.*

Walid Taha (Chair of the Committee on Internal Affairs and Environmental Protection):

(Speaks in Arabic; henceforth translation: I appeal to everyone who is following the discussions now around the Electricity

- 5 *Law, the law that they have tried in every way to thwart and prevent from arriving here today, to vote on it in both readings, the second and the third. [...]*)

Keti Kathrin Shitrit (Likud):

Walid, we don't understand, what's going on?

- 10 **Walid Taha:**

You didn't understand the needs of Arab society for decades either, Keti.

Keti Kathrin Shitrit (Likud):

Lies. This won't help you. It's a lie, a lie.

- 15 **Itamar Ben-Gvir (Jewish Power):**

- - - like in Syria. He thinks he's in Syria.

Walid Taha:

- (*Speaks in Arabic; henceforth translation: 96 hours were shortened – [about Itamar Ben-Gvir] that's a racist fascist, speaking in a vulgar way that corresponds to his vulgar manner – I was saying that the Committee shortened the debate time to 14 hours [...].*)
- 20

Keti Kathrin Shitrit (Likud):

- Really, Mickey, we don't have anything to do here, he is speaking in Arabic, let's leave. Really, why are you [plur.] here?*
- 25

Walid Taha:

- (*Speaks in Arabic; henceforth translation: Still, they were not satisfied, and declared a boycott of the Knesset debates, left the Knesset and went home yesterday [...]. Now that I am speaking, opposition Knesset members are protesting that I speak in Arabic.*)
- 30

Shouts: *This is crazy ... in the Knesset. Disgrace. Simply unbelievable.*

- 35 **Walid Taha:**

- (*Speaks in Arabic; henceforth translation: They are so angry, they could explode. They are angry that I am speaking in Arabic, and I will speak Arabic whenever I want. Avi Dichter, do you hear, I will speak Arabic and whoever doesn't like it can go and drink the water of the Dead Sea.*)
- 40

Avi Dichter (Likud):

(Says in Arabic: Walid, why didn't you speak Arabic when you voted against the peace agreements with the Emirates? Until today you didn't say anything in Arabic. [the back-and-forth
45 between Walid Taha and Avi Dichter, mostly in Arabic, continues for 10 pages (Knesset 2022c: 92–101)].)

Avi Dichter (Likud):

You are not a reasonable man. If you were a reasonable man, you would speak in Arabic and translate into Hebrew for the
50 98 per cent who don't understand you. As it is, you are not speaking to the Knesset. You are in the Knesset and you are not speaking to the Knesset.

Walid Taha:

Me, when you speak in Hebrew, I do understand. That you
55 don't understand when I speak Arabic, that's your problem. Why didn't you learn Arabic?

Itamar Ben-Gvir (Jewish Power):

- - - You are a guest.

David Amsalem (Likud):

60 You are right, it's our problem that they allowed you to speak Arabic here. You [plur.] are our problem. [...] It's a problem that you are still speaking here. [...] Here they are, the pair who are stealing the country, on the podium.

Chair Mansour Abbas:

65 - - -

Walid Taha:

(Speaks in Arabic; henceforth translation:)

David Amsalem (Likud):

Tell me, should we bring you coffee and baklava? Aren't you
70 ashamed? Look at this, in the Knesset of Israel, two Arabs are talking among themselves. Look at what we have come to. They are making a mockery of us.

Shouts: We want translation - - -

Chair Mansour Abbas:

75 Alright. Avi Dichter and Yariv Levin know Arabic, they know what we are talking about.

David Amsalem (Likud):

Soon you'll get electricity, you took 53 billion, and now you've turned the Knesset into Arabic, do you understand?

80 **Chair Mansour Abbas:**

David Amsalem, calm down, at the end of the day it's just an Electricity Law.

David Amsalem (Likud):

85 *At the end of the day it's the theft of a third of the State of Israel.*

Ofir Katz (Likud):

Think of the Jews here, that we are a minority, be considerate of us.

Chair Mansour Abbas:

90 *Ofir, alright. Well done. [...] Look, you can ask Avi Dichter; whatever he [Walid Taha] had said in Hebrew, he said in Arabic.*

Keti Kathrin Shitrit (Likud):

If he won't speak Hebrew, let them bring us earphones.

Walid Taha's insistence on speaking Arabic and refusing to self-translate appeared hostile, based on the lack of politeness markers. However, his ten-minute exchange in Arabic with Avi Dichter seemed to be pleasant to both of them, as they are smiling at each other, or to themselves, as observed in the video recording between minutes 36 and 47 (Knesset 2022a). They seemed to be on the same discursive team, using the Arabic they had in common for some form of positive politeness, enjoying the jousting. In contrast, around them the reactions to the Arabic became increasingly angry – David Amsalem looked apoplectic – until Kathrin Shitrit mentioned the need for earphones for simultaneous interpretation (Knesset 2022c: 131). Then, the discussion returned to Hebrew until the end of the session, when the law was passed.

The trigger for the switch to Arabic seems to be the turn immediately preceding it – a complaint by Moshe Abutbul that Walid Taha had ignored problems that Ultra-Orthodox Jews face which are supposedly similar to those faced by Palestinian and other citizens of Israel regarding the supply of electricity.⁶ Haredi Jews (as the Ultra-Orthodox prefer to call themselves) are among the poorest sectors of Israeli society for a combination of reasons, but mostly by communal choice (OECD 2020: 14). Faced with this confusion of the issue, and provoked by constant heckling from rightwing opposition to the Electricity Law, Walid

⁶Searching for evidence that Orthodox religious school (Yeshiva) dormitories were deprived of electricity, I found that some Yeshivas could not afford their electricity bills, and had their supply temporarily cut (Matzav.com Israel News Bureau 2014).

Taha turned abruptly to Arabic to address pointedly the members of the Joint List who had criticized the framing of the bill, and to appeal to them for support (lines 3–7). This switch prompted the first request for translation, from Kathrin Shitrit (line 9), which Walid Taha dismissed irritably with the quip, “You didn’t understand the needs of Arab society [for electricity] for decades, either” (lines 11–12), implying, “Well, too bad for you, not understanding Arabic.” The debate went downhill from there.

In this mood, Walid Taha adhered to Arabic for declarations of defiance (lines 37–40) which mimicked the manipulation of the indexicality of one nation, one language, explored in Example (1). However, he did not soften his defiance with polite greetings and self-translations, as seen in Example (1); apparently, politeness is not necessary when one is in government. What is meant by ‘manipulation’ is that the speaker speaks in a language that s/he claims as defining their national identity, in symbolic reference to the equation of one nation, one language. They are not speaking that language simply because they are of that nationality; in practice, they are speaking several languages as a skilled multilingual, for multiple pragmatic purposes. The symbolic reference to, rather than practical use of, the one national language is on a higher order of indexicality. The speech act of displaying now unapologetic multilingualism is intended to signify political power, or at least entitlement to the particular platform above the hecklers’ objections.

The provocation of having to listen to political theatre in Arabic in the Knesset incensed some of the interlocutors, but also overwhelmed the structures of Knesset debate. The Chair, Mickey Levy, apparently gave up on intervening for order after Kathrin Shetrit challenged him to restore the supremacy of Hebrew with, “he is speaking in Arabic, let’s leave. Really, why are you here?” (lines 24–26). There was no keeping to time; Walid Taha and Avi Dichter jousted away; heckling was unchecked for a while. Itamar Ben-Gvir’s racist “You are a guest” (line 58) – meaning that an Arab is barely tolerated in the Knesset (and in Itamar Ben-Gvir’s vision of the Land of Israel) – was not countered with a caution. Mansour Abbas arrived to start his shift as chair, amid the heckling, with an aside to the speaker at the dais, Walid Taha. He settled into his chair at minute 47 of the video recording of the debate (Knesset 2022a). The stenographer noted an effort to translate the Arabic speech, but the effort failed, and the record leaves the turns as blanks (lines 65–67). Even this minimal Knesset translation service, the last line of republican discursive defence, had broken down.

The day before, it had seemed acceptable for Arabic speakers to address other Arabic speakers in Arabic in the Knesset, since that had maintained the norm

of using Arabic within the in-group. However, this level of Arabic audibility underestimated the strength of the second norm: the silencing of Arabic on Zionist platforms. By having a brief exchange in Arabic on the Knesset podium, Mansour Abbas and Walid Taha undid the Zionist nature of that stage. For David Amsalem, they had transformed the Knesset into an Arab coffee house, offensive to the history of Zionist institution-building: “Tell me, should we bring you coffee and baklava? Aren’t you ashamed? Look at this, in the Knesset of Israel, two Arabs are talking among themselves. Look at what we have come to. They are making a mockery of us” (lines 69–72). Such is the power of discourse to transform situations – from Zionist to non-Zionist – on the back of language ideologies (Irvine & Gal 2009). Soon, the very existence of Israel was at stake; the idea that marginalized Bedouin villages could connect to electricity and might contemplate the reversal of decades of infrastructural neglect amounted to “the theft of a third of the State of Israel” (lines 63, 84–85). All this emotion was caused by Arabic in the Knesset.

Ofir Katz repeated the request for translation by returning to Ayman Odeh’s joke (Example 3.1), saying, “Think of the Jews here, that we are a minority, be considerate of us” (lines 87–88). The joke did not make Mansour Abbas laugh (he reacted with a sarcastic “Well done”; line 90), because the affiliations were not correct; Ofir Katz, as opposed to Ayman Odeh, was not actually fighting for minority rights, and could not therefore claim self-deprecation. Kathrin Shetrit finally requested institutional simultaneous interpreting, after nearly one hour of arguing (line 94). Institutional simultaneous interpreting has been provided in the Knesset on the annual “Arabic language day”, a performance of the liberal utopia periodically enacted by the Joint List since 2017 (Hawker 2019: 53–55). The rest of the year, the earphones stay in the cupboard.

Mansour Abbas reassured the angry interlocutors (lines 90–92) that despite the absence of explicit signposting, Walid Taha had indeed self-translated his Arabic, with Avi Dichter as his witness, whose reliability was underpinned by being Jewish Israeli. Hebrew dominance was reinstated until the end of the session, patiently chaired. It did not matter whether the Arabic had in fact been translated into Hebrew for comprehension purposes; what mattered was that the norm was explicitly re-established that Arabic speakers were responsible for their own interpreting. The law was passed, and Benjamin Netanyahu called it a “black day for Zionism and democracy” (Muallem 2022). At the time of writing, no villages had been newly connected to electricity, and the governing coalition looked fragile (Abu Arshid 2022).

3.3 Politeness that breaks norms

Islamists joining the Israeli government broke a political taboo, and they also inadvertently changed the norms of language choices for the Palestinians and other Arabs who speak Hebrew in addition to their native Arabic. The starkest description of that change in norms (and perhaps the most absurd description to a reader outside Israeli and Palestinian realities) is David Amsalem's lament: "Look at this, in the Knesset of Israel, two Arabs are talking among themselves. Look at what we have come to" (Example 3, lines 69–72). Look, indeed: since the Zionism of the Knesset platform is discursively constituted by Hebrew dominance, Arabs speaking Arabic raises questions as to the precise identity of the state institution that is being (re-)formed. Putting to one side the radical vision of a putative liberal state where every citizen would thrive, let's apply the language ideology test. If Israel were a republican state that manages minorities, then Arabic should have been translated in the Knesset. If it were an ethnonationalist state that excludes minorities, then Arabic speakers would be delegitimized. The litmus test of language ideologies, at least in January 2022, indicates an inclination towards the latter.

The discursive events of 4 and 5 January 2022 were not exceptions, but points in a trajectory of change. Palestinian politicians in Israel have been speaking defiantly in Arabic on Zionist platforms for ten years (Hawker 2019: 39–50), and the Islamists have simply taken it to another level (Leal 2022). The differences between 2015 and 2022 are in the newer lack of announcement of switches to Arabic, the positive rather than negative politeness strategies, and the occasional refusal to self-translate. Overall, I will submit that the changes are related to the interactive organizational work done by politeness. There has long been a stereotype that Jewish citizens of Israel, especially those born to families with a socialist-Zionist-pioneering ethos, will speak to the point, without polite adornments, while Palestinian citizens of Israel, especially those who are born to traditionalist patriarchal families, will speak indirectly, paying attention to interlocutors' face-saving needs (Katriel 1986). Mansour Abbas, described by Israeli commentators as "soft-spoken and affable" (Karsh 2022) is the personification of the traditionalist patriarch, and his performance of that type of politeness served him well in chairing difficult Knesset debates.

The stereotypes of contrasting Jewish Israeli and Palestinian politeness exist on the level of the first order of indexicality, in that they can be inherited, unselfconscious styles. Yet they can also be found in elaborations on higher orders of indexicality. These elaborations exist in Palestinians' self-conscious pride in being rude, such as when Walid Taha said, "Avi Dichter, do you hear, I will speak

Arabic and whoever doesn't like it can go and drink the water of the Dead Sea" (Example 3, lines 38–40). Another elaboration exists in Avi Dichter's policing of what is a "reasonable" way for a Palestinian to speak: "If you were a reasonable man, you would speak in Arabic and translate into Hebrew" (Example 3, lines 48–49). These are explicit affirmations of what is polite and impolite in the situation. My contribution to the analysis of politeness in political debates is to move away from stereotypes or elaborations on them, and towards organization of power relations. When Palestinian politicians use negative politeness strategies in asking permission to speak Arabic, they are showing regard for a raciolinguistic hierarchy of discursive legitimacy (Heller & McElhinny 2017: 251), which serves to expose that hierarchy where there is a pretence that there is none. When they include Jewish Arabic speakers such as Eli Avidar and Avi Dichter in the in-group that can be addressed in Arabic, using positive politeness strategies, they are showing that the cultural capital of speaking several languages is supranational. It is to this accumulation of cultural capital that I turn in the discussion of class formation that follows.

4 Discussion: Combining conflict and late capitalism in the multilingualism analysis

Israel has developed a knowledge- and technology-based economy, with some of the highest inequality and poverty rates among high-income nations (OECD 2021). Since the 1990s, in line with the neo-liberalization of the economy globally as well as in Israel, tertiary education has increased exponentially. At the same time, by 2015, the salary gaps between tertiary-educated and uneducated workers grew, and yet Palestinian and other Arab university graduates earned an average monthly income of USD 1,885, as compared to USD 3,149 among Jewish Israelis, and Palestinian and other Arab university graduates were more likely to be un- or under-employed (Ayalon & Mcdossi 2016). These graduates want political engagement, voting for the Joint List, but they also want the aspects of the good life that consumerist late capitalism can offer (Carmeli & Applbaum 2004). Clearly, this good life is not compatible with disconnection from the electricity grid. What is more, this educated, aspirational emergent middle class has been making multilingual political jokes about discrimination against them since 2009 (Henkin 2009). I argue that there are articulated links between the emergent middle class in consumerist late capitalism and enregistering multilingual discourse.

I still argue for those articulations, but would like to use this paragraph to reflect. I had thought that any increase in Arabic audibility on Zionist platforms,

discursively transforming the identity of those platforms, would be accompanied by some steps towards liberal multiculturalism, as the performances of the Palestinian politicians had augured over the past ten years. The manifesto on the cultural capital of multilingualism had been a promise that I had taken as a commissive speech act. In this liberal vision, there was supposed to be dignity, respect, and perhaps even resolution of conflict. The speakers themselves had said so, in the words of Ayman Odeh reported in Section 2 and Aida Touma-Suleiman in Section 3.1 above. In linguistic anthropology, we appreciate our participants' perspectives, yet I depart from my participants' self-perception to offer my own perspective on the evidence.

While participant perspectives are privileged, they are not necessarily made into the cornerstone of the analysis. Researchers take a critical stance with regard to all texts and appeal to critical ethnography, which allows them to triangulate the data, and to critical discourse analysis, which enables them to uncover hidden ideological meanings. (Pavlenko & Blackledge 2004: 25)

Combing through the evidence with the fine-toothed analysis of politeness, what I find does not resemble dignity, respect, or resolution of conflict. What I find is supranational opportunistic alliances for power maximization, which gives speakers in power the entitlement to dominate, in whatever language. I find selective invocation of reinvented norms and of fragile and unfair institutional rules. I also find suppression of both conflict and conflict resolution: "It's just an Electricity Law" (Example 3, line 81). Speakers are not making any steps towards liberal visions of freedom. What they are advancing towards is the offering of late capitalism: improvements in economic conditions are negotiated on behalf of the collective, yet for the benefit of individuals, both for politicians winning points at elections and voters seeing specific material gains. Wider political issues regarding the definition of the common good and conceptions of freedom are suspended (Davis 2008). Late capitalism and its commodification of language, including of multilingualism's cultural capital, is put to the service not of the liberal side of the emergent middle class, but of the consumerist side. And yet, the most painful realization is knowing that this offering is an improvement on the apartheid experienced by Palestinians at present (Amnesty International 2022). The push towards Palestinian access to the bounties of consumerism might undo some aspects of apartheid and co-opt others (Taha & Hawker 2020). Meanwhile, to simplify debates a little, institutions experiencing bouts of multilingual discourse could take the interpreters' earphones out of the cupboard where they have been gathering dust – the precarious jobs of interpreters are indeed a harbinger of the late-capitalist future.

5 Conclusion

I have brought together evidence from political speeches that code-switch between Arabic and Hebrew to show that there is a trajectory of change in multilingual styles. The commodification of multilingualism that comes with late capitalism has appreciated the cultural capital of middle-class speakers who use several languages. They use their languages skilfully for pragmatic purposes of politeness and of group identification, and to convey political messages. Already departed by several indexical orders from the nationalist conflict patterns of one nation, one language, the new patterns do not align neatly with liberal values of dignity, respect, and conflict resolution. Rather, the multilingual discourse is the code of the emergent middle class that sees its relative power as an opportunity to push for greater access to consumerist material benefits. This opportunity challenges an oppressive reality but might not lead to the type of freedom that the speakers themselves had articulated with multilingual discourse.

Acknowledgements

My sincere thanks go to the members of the CoTiSp team, Valentina Serreli, Katrin Pfadenhauer, and Sofia Rüdiger, who encouraged me with patience and kindness as I wrote this chapter. My greatest intellectual debts go to Deborah Cameron, who mentored my academic career and taught me to identify structural biases against speeches by marginalized speakers, to Alexandre Duchêne, who has trusted my expert judgment enough to make me trust myself and who introduced me to the scholarship on late capitalism, and to Clive Holes, who has defended my focus on Israel/Palestine and taught me to discern pragmatic purposes of Arabic political speeches. I am grateful to politicians and their constituents in Israel/Palestine who allowed me to attend their meetings and learn from them how to be a citizen. I thank my friends who have guided me through the Israel/Palestine political landscape: Alaa, Helal, Maha, Noa, Saleh, Tal, and Yael.

References

- Abu Arshid, S. 2022. المشتركة بين شبكة الأمان وإسقاط حكومة بينيت [The joint list...between safety net and overthrowing the Bennett government]. Arabs 48 News .48 موقع عرب <https://www.arab48.com/>.

- Amer, Faten, Dipima Buragohain & Ina Suryani. 2020. Responding strategies to greetings and invocations in Jordanian Arabic in CCCS interactions: A socio-pragmatic study. *European Scientific Journal* 16(8). 62–74.
- Amnesty International. 2022. *Israel's apartheid against Palestinians: Cruel system of domination and crime against humanity*. <https://www.amnesty.org/en/documents/mde15/5141/2022/en/>.
- Ayalon, Hanna & Oded Mcdossi. 2016. First-generation college students in an expanded and diversified higher education system: The case of Israel. In Nabil Khattab, Sami Miaari & Haya Stier (eds.), *Socioeconomic inequality in Israel: A theoretical and empirical analysis*, 75–96. New York: Palgrave Macmillan US.
- Ben-Rafael, Eliezer. 1994. *Language, identity, and social division: The case of Israel*. Oxford: Oxford University Press.
- Block, David. 2013. *Social class in applied linguistics*. London & New York: Routledge.
- Bourdieu, Pierre. 1982. *Ce que parler veut dire: L'économie des échanges linguistiques*. Paris: Fayard.
- Bourdieu, Pierre. 2004. *Science of science and reflexivity*. Cambridge, Oxford, New York, Boston: Polity.
- Carmeli, Yoram S. & Kalman Applbaum. 2004. *Consumption and market society in Israel*. Oxford & New York: Berg.
- Davis, Mark. 2008. *Freedom and consumerism: A critique of Zygmunt Bauman's sociology*. London & New York: Routledge.
- Dean, Kathryn. 2014. *Capitalism, citizenship and the arts of thinking: A Marxian-Aristotelian linguistic account*. London & New York: Routledge.
- Eckert, Penelope. 2004. The meaning of style. *Texas Linguistic Forum* 47. 41–53.
- Eelen, Gino. 2001. *A critique of politeness theory*. London & New York: Routledge.
- Facebook. 2022. *محبو الدكتور منصور عباس* [The Fans of Dr Mansour Abbas] on Facebook Watch. <https://tinyurl.com/5y33m72s>.
- Gal, Susan. 2012. Sociolinguistic regimes and the management of “diversity”. In Alexandre Duchêne & Monica Heller (eds.), *Language in late capitalism: Pride and profit*, 22–42. London & New York: Routledge.
- Ghanem, Asad. 2016. Israel's second-class citizens: Arabs in Israel and the struggle for equal rights the struggle for Israel. *Foreign Affairs* 95(4). n.p.
- Handelman, Don. 2020. *Nationalism and the Israeli state: Bureaucratic logic in public events*. London & New York: Routledge.
- Hartung, Cathrine, Nicoli Barnes, Rosie Kate Welch, Gabrielle H. O'Flynn, Jonnell Uptin & Samantha McMahon. 2017. Beyond the academic precariat: A collective biography of poetic subjectivities in the neoliberal university. *Sport, Education and Society* 22(1). 40–57.

- Hawker, Nancy. 2018. The mirage of 'Arabrew': Ideologies for understanding Arabic-Hebrew contact. *Language in Society* 47(2). 219–244.
- Hawker, Nancy. 2019. *The politics of Palestinian multilingualism: Speaking for citizenship*. London & New York: Routledge.
- Hawker, Nancy. 2020. Who will Palestinians vote for in the Israeli elections? *Open Democracy*. <https://www.opendemocracy.net/en/north-africa-west-asia/who-will-palestinians-vote-israeli-elections/>.
- Heller, Monica, Lindsay A. Bell, Michelle Daveluy, Mireille McLaughlin & Hubert Noel. 2015. *Sustaining the nation: The making and moving of language and nation*. Oxford: Oxford University Press.
- Heller, Monica & Bonnie McElhinny. 2017. *Language, capitalism, colonialism: Toward a critical history*. Toronto: University of Toronto Press.
- Henkin, Roni. 2009. The cognate curse in Negev Arabic: From playful punning to coexistence conflicts. *Israel Studies in Language and Society* 2(2). 169–206.
- Irvine, Judith T. & Susan Gal. 2009. Language ideology and linguistic differentiation. In Alessandro Duranti (ed.), *Linguistic anthropology: A reader*, 402–434. Hoboken: Wiley-Blackwell.
- Karsh, Efraim. 2022. Why Israel's Arabs are its biggest threat. *Middle East Quarterly* 29(2). <https://www.meforum.org/62991/why-israel-arabs-are-its-biggest-threat>.
- Katriel, Tamar. 1986. *Talking straight: Dugri speech in Israeli Sabra culture*. Cambridge: Cambridge University Press.
- Knesset. 1977. *The 43rd session of the ninth Knesset הישיבה הארבעים ושלוש של הכנסת התשיעית*.
- Knesset. 2012. *Knesset rules of procedure. The Knesset*. <https://main.knesset.gov.il/EN/activity/Pages/RulesOfProcedure.aspx>.
- Knesset. 2015. *Knesset records, the 80th meeting of the twentieth Knesset הישיבה ה-80 של הכנסת העשרים*. https://fs.knesset.gov.il/20/Plenum/20_ptm_318579.doc.
- Knesset. 2022a. דיון במליאת הכנסת –05.01.22– חלק ב [Debate in the Knesset Plenum –5 January 2022–Part B]. ערוץ כנסת [Knesset Channel]. <https://www.knesset.tv/parliament/programs/37229/>.
- Knesset. 2022b. *Following a heated debate, bill for connecting houses built without a permit to the electric grid passes final readings in the Knesset Plenum*. <https://main.knesset.gov.il/EN/News/PressReleases/Pages/press5122r.aspx>.
- Knesset. 2022c. הישיבה התשעים וחמש של הכנסת העשרים וארבע [The 95th session of the 24th Knesset].
- Knesset. undated(a). *Knesset Member Eli Avidar*. <https://main.knesset.gov.il/en/MK/APPS/mk/mk-personal-details/983>.

- Knesset. undated(b). *The Knesset Plenum at work*. <https://m.knesset.gov.il/en/activity/pages/plenumwork.aspx>.
- Kuipers, Giseline. 2016. Follow the joke: Humour and ethnography. *Etnofoor* 28(2). 125–129.
- Leal, Iris. 2022. *An Arab lawmaker speaks his language, and the right loses its mind*. Haaretz. <https://www.haaretz.com/opinion/.premium-an-arab-lawmaker-speaks-his-language-and-the-right-loses-its-mind-1.10526233>.
- Matzav.com Israel News Bureau. 2014. No electricity at Yeshiva Be'er Ya'akov. <https://matzav.com/no-electricity-at-yeshiva-beer-yaakov/>.
- Meyer, John C. 2000. Humor as a double-edged sword: Four functions of humor in communication. *Communication Theory* 10(3). 310–331.
- Mualem, Mazal. 2022. *Electricity law connects Bedouin homes to Israel's power grid*. Al-Monitor: The Pulse of the Middle East. <https://www.al-monitor.com/originals/2022/01/electricity-law-connects-bedouin-homes-israels-power-grid>.
- Myles, John. 1999. From habitus to mouth: Language and class in Bourdieu's sociology of language. *Theory and Society* 28(6). 879–901.
- Negev Coexistence Forum. 2020. *Palestinian-Bedouin communities living in Israel more vulnerable than ever during COVID-19*. <https://www.amnesty.org.uk/blogs/campaigns-blog/palestinian-bedouin-communities-living-israel-more-vulnerable-ever-during>.
- Nowicka, Magdalena & Maria Rovisco (eds.). 2016. *Cosmopolitanism in practice*. London & New York: Routledge.
- OECD. 2020. *OECD Economic surveys israel overview*. <https://www.oecd.org/economy/surveys/Israel-2020-OECD-economic-survey-overview.pdf>.
- OECD. 2021. *Israel economic snapshot: Economic policy reform 2021 going for growth*. <https://www.oecd.org/economy/growth/Israel-country-note-going-for-growth-2021.pdf>.
- Pappé, Ilan. 2011. *The forgotten Palestinians: A history of the Palestinians in Israel*. New Haven: Yale University Press.
- Pavlenko, Aneta & Adrian Blackledge. 2004. Introduction: New theoretical approaches to the study of negotiation of identities in multilingual contexts. In Aneta Pavlenko & Adrian Blackledge (eds.), *Negotiation of identities in multilingual contexts*, 1–33. Bristol: Multilingual Matters.
- Peled, Yoav. 1992. Ethnic democracy and the legal construction of citizenship: Arab citizens of the Jewish state. *American Political Science Review* 86(02). 432–443.

- Rouhana, Nadim N. & Nadera Shalhoub-Kevorkian. 2021. *When politics are sacralized: Comparative perspectives on religious claims and nationalism*. Cambridge: Cambridge University Press.
- Shohamy, Elana. 2005. *Language policy: Hidden agendas and new approaches*. London & New York: Routledge.
- Shohat, Ella. 2017. *On the Arab-Jew, Palestine, and other displacements: Selected writings of Ella Shohat*. London: Pluto Press.
- Suleiman, Yasir. 2004. *A war of words: Language and conflict in the Middle East*. Cambridge: Cambridge University Press.
- Suleiman, Yasir. 2019. *Arabic language and national identity: A study in ideology*. Edinburgh: Edinburgh University Press.
- Taha, Habatalla & Nancy Hawker. 2020. Freedom and peace at the shopping centre: The politics of consumerism in Israel/Palestine. *Journal of Political Ideologies* 25(3). 294–315.
- Weaver, Simon. 2016. *The rhetoric of racist humour: US, UK and global race joking*. London & New York: Routledge.
- Woolard, Kathryn. 2012. Catalonia: The dilemma of language rights. In Nessa Wolfson & Joan Manes (eds.), *Language of inequality*, 91–109. Berlin: Walter de Gruyter.
- Yadgar, Yaacov. 2020. *Israel's Jewish identity crisis: State and politics in the Middle East*. Cambridge: Cambridge University Press.

Part III

Migration and language contact

Chapter 7

Mobile and complex: A West African linguistic repertoire

Klaudia Dombrowsky-Hahn^a & Axel Fanego Palat^b

^aUniversity of Bayreuth ^bGoethe University Frankfurt

Migration is one of the sources of individual multilingualism. Patterns of mobility are typically more complex than a simple move from an original home to a new residence; they can involve trajectories including internal, rural–urban, south–south, south–north, and circular migration. An individual’s experience of migration is reflected in their linguistic repertoire. Migrants commonly acquire new linguistic resources, expanding their repertoire throughout their itinerary. This is especially true of mobile people from West Africa, where urban and rural multilingualism is common in many regions.

In our project entitled “African people in the Rhine-Main region – a project on linguistic integration”, we study language repertoires of speakers from different African countries. Through multimodal methods, including the collection of language portraits, accompanying narratives, and interviews, we get to know mobile people’s biographies and their histories of language acquisition. The data can also be analysed with a view to contact phenomena.

In this chapter, we take a close look at the use of German during an extended interview conversation with one speaker, Kajatu, a woman born in Guinea. We focus on three examples from different tiers of language structure: the semantics of the spatial preposition *in*, morphosyntactic properties of genitive constructions, and phonetic–phonological details of nasalization processes. In all three, we find evidence that the speaker draws on her entire linguistic repertoire, marked by several West African and European languages. Differences between Kajatu’s use of German and standard norms cannot simply be attributed to ‘automatic’ processes of native language interference. Instead, individual usage patterns emerge and stabilize that can sometimes be traced back to one of the various other languages in her repertoire. In this sense, the linguistic forms on the levels of phonetics, morphosyntax, and (lexical) semantics index the individual’s biography and identity.



1 Introduction: Mitigating structural imprints and fluid repertoires

Many migrants from West Africa are highly multilingual. They learn some languages along their paths of migration, while their proficiency in others results from multilingualism in the places where they used to live, places they may consider home. We will focus on one speaker, Kajatu [ka'jatu],¹ born in Guinea. At the time of the interview in 2019, she had been living in Germany for over twenty years. Kajatu lived in several countries before settling in Germany. Her linguistic biography is therefore remarkable, but not exceptional – which is all the more reason for us to deem her story worth documenting. Kajatu's story, the narrative she builds in the interview, is told mostly in German, and her German is clearly characterized by her complex linguistic biography. The lessons we learned from engaging with Kajatu and her way of using German have several layers, and conveying these layers is the main goal of this chapter.

First of all, one significant insight is that Kajatu's linguistic competence is based on a repertoire that draws on several languages on a constant basis. Yet her linguistic competence constitutes one repertoire, not an 'assortment of languages'. By this we mean that she does not constantly shift back and forth between languages, and she does not insert chunks from one language into another in a patchwork fashion; instead, her communication is holistic. Secondly, a closer look at Kajatu's German reveals that it is not a random mixture of input from whatever language first comes to her mind when she speaks German, or a lack of certain lexical or grammatical means. Instead, many non-standard features of her German partially indicate stabilized usage patterns. These draw on her full repertoire, including languages other than German, which can be traced through structural analysis. Thirdly, where there is variation in the way a particular structure is instantiated, there is the potential for indexical significance, for example when Manding-like structures appear more frequently when Kajatu talks about her Maninka-speaking grandmother. Variation is therefore not a sign of insufficient competence or insecurity in terms of knowledge of German.

We looked into examples from three different sub-domains of grammar, because we wanted to scrutinize conventional ideas of imperfect learning and ('native language') interference – to the extent this is possible on the basis of a fairly limited set of data, as in our case study. Phonetics/phonology, morphosyntax, and (lexical) semantics appear to differ in how grammatically entrenched they

¹Kajatu is the speaker's fictitious name.

are, arguably mirrored (and perhaps caused) by their order in first language acquisition. This would suggest that interference effects in second/later language learning are not equally likely for all domains of grammar. A functional motivation for this could be that phonetic and phonological features are more strongly driven by language-specific norms and conventions, deeply engrained in early first language acquisition. In contrast, lexical semantics pertains to the construal of meaning. This field is, in principle, infinite and constantly addresses new communicative needs, which can arise irrespective of an individual's alleged first language. The speaker's lexical semantic repertoire seems more likely to be affected by linguistic experiences beyond first language acquisition, and not as immutable as their phonetic-phonological habitus, for instance. Our examples do not provide straightforward corroboration for this. Instead, with regard to all three features (the semantics of the German preposition *in* 'in', genitive constructions, and processes affecting the pronunciation of vowel-nasal combinations), non-standard usage patterns attested in Kajatu's speech defy such a clear ordering. The explanation, we believe, lies in a refined understanding of linguistic repertoires and speaker agency.

We will address some theoretical notions in Section 2, which follows this introduction. Section 3 will then provide some useful background information on the linguistic setting in parts of West Africa relevant to Kajatu's experiences. These will be presented in greater detail in Section 4, before we turn to specific evidence of language contact in her speech and use of German in Section 5. This will be followed by a discussion and conclusions in Section 6.

2 Repertoires and languages: Theoretical notions and methodological approaches

We align with Matras (2020: 4ff) in choosing a repertoire approach to multilingualism. According to Matras, a multilingual speaker's repertoire is made primarily of linguistic structures, among others word forms, phonological rules, and constructions. During language socialization, these structures become associated with social activities, including factors such as interlocutors, institutional settings, and conversational topics. Their separation into languages or classification as linguistic systems and their labelling, as well as conventions and constraints as to when and how to use and mix them, is learned through metalinguistic activities.

Multilingual individuals, especially if they learn languages in an unguided way, draw on *all* the linguistic resources in their complex repertoires to achieve

successful communication, even if these choices are not always “conscious, deliberate, or strategic” (Matras 2020: 7). A multilingual speaker’s complex repertoire is constantly present and available, and the speaker draws on it rather than switching single “language systems” on and off (Matras 2020: 9). When drawing on their entire repertoire and exploiting its full expressive potential while complying with interlocutors’ expectations, speakers become creative communicators. A similar approach is advocated for multilingualism in African languages by Lüpke & Storch (2013) and Storch (2016). Theirs is a refined linguistic understanding of agency – a crucial concept when emphasizing language as an activity rather than as a structure (see Pennycook 2010). Interestingly, post-structuralist and critical approaches to language (Makoni & Pennycook 2007, 2012, Pennycook 2018) converge in some of these views, with cognitive linguistic positions insisting on the significance of usage-based models, radical construction, or emergent grammar, although they differ in others, for example when they emphasize that linguistic repertoires and practices are inherently messy (Storch 2016).

Overall, these influential voices emphasize ideas of real-time construal of meaning, language as an activity, and fluid repertoires. How then can the emergence of those stable structures and usage patterns that we see in Kajatu’s use of German be accounted for? We suggest that her use of formal substance, structures, and construction types across conventional language boundaries is a good starting point for finding out about how language impacts us, and how we in turn mould language for our own purposes. The task we have set for ourselves is to highlight how, in the case of one speaker’s German, a broad range of languages is drawn upon when speaking German. This does, in fact, lead to certain structural choices in Kajatu’s use of German, afforded by other languages – for example varieties of Fula and Manding – that are not ‘switched off’, but remain available to Kajatu at all times. When a speaker does this in a regular fashion, a traditional system-based view of language would describe this as one system ‘interfering’ with another.

Matras (2020: 76ff) rejects the concept of interference as a meaningful explanation for why speakers rely on structures of one of their languages when, in fact, they speak another. The most common interference scenario that springs to mind is that of native language structures surfacing in a ‘foreign’ language. The term ‘interference’ suggests that the ‘correct’ acquisition of a target language by an individual learner is hampered by categories and structures of their native language (see Weinreich 1953). Notions of ‘imperfect learning’ possibly leading to substrate effects in language-contact scenarios, specifically with a view to language shift, is but one example which can be extrapolated from individual language competence to sociolinguistic effects at a macro level.

A repertoire approach emphasizes achieving successful communication as the primary goal of using language. From that perspective, ‘interference’ is not a meaningful term. Rather, structures of one language used when speaking another are viewed “as enabling factors that allow language users to build bridges between different subgroups within their overall repertoire of linguistic forms and to use these bridges to maintain communication” (Matras 2020: 78).

This paves the way to a contact-linguistic approach to migrants’ language-learning practices that overcomes notions of (inextricably deficient) second language acquisition. The speaker whose communicative creativity is studied in this chapter uses several languages, and we will show that resources from all of them are creatively employed in the interaction. Capturing these phenomena systematically can be a daunting task, which is why a few words about how we went about this are apt.

The interview which forms the core of our study was conducted by the first author of this chapter. She met Kajatu through a mutual acquaintance in the context of a research project on linguistic practices of Africans living in the Frankfurt area in Germany. In this particular project, a range of methods were used to obtain data. Central to all of them was a multimodal approach that aimed to uncover the speakers’ heteroglossic practices and their own interpretations of the experience of these practices, as described by Busch (2018). Our interlocutors were first asked to visualize their linguistic repertoire using a drawing of a body silhouette representing them. They were requested to insert languages, varieties, or registers in or around the body silhouette using coloured pens at their convenience (see Figure 1 for an example). These visual images produced by our interlocutors formed the starting point for interviews that we conducted with 25 people in 2019 and 2020. Based on the drawings and supplemented with ethnographic observations, these conversations about the speakers’ languages and varieties and how they learned them were often quite long; the interview with Kajatu lasted an hour. In addition to being rich and informative personal narratives centring on the theme of mobility and language learning, they are also textual data that lend themselves to structural analysis of language-contact phenomena. However, the structural analysis is not detached from contextual information.

Taken together, the different perspectives result in a linguistic ethnography approach that combines the analysis of language structures with insights drawn from observation and the interpretation of personal narratives as recorded in the interview. Before we examine Kajatu’s case, we will provide some information on language settings – both in terms of sociolinguistic scenarios and with regard to significant typological information – to help the reader situate Kajatu’s trajectory and experiences.

3 Language settings in West Africa as a backdrop to Kajatu's repertoire

Pular, Maninka, and Susu, the Guinean languages in Kajatu's repertoire, are the most important lingua francas in the capital, Conakry, and in three of the four regions of the country respectively. Susu is the lingua franca of Basse Guinée (Lower Guinea) in western Guinea and the main language in the capital. Pular is widely spoken in Moyenne Guinée, the central region, and Maninka in Haute Guinée, the north-eastern part of the country and in urban centres of Région Forestière in south-eastern Guinea. Although 35 languages are spoken in the country, many inhabitants speak the three listed above (Barry 2014); thus, Kajatu's multilingualism can be considered as typical of Guineans.

Bambara (or Bamanankan) and Fulfulde are widespread as an L1 and L2 in Mali, a country where 63 languages are spoken according to *Ethnologue* (Eberhard et al. 2021). Both have the status of national languages in Mali. In addition to being spoken as an L1 in the regions of Ségou, Koulikoro, Sikaso, and Kayes and by many inhabitants of the capital, Bamako, Bambara is the most common lingua franca, gaining prominence throughout the country (Cissé 2020, Dumestre 1994b, 1998). Approximately 80 per cent of the Malian population use it to a greater or lesser extent (Eberhard et al. 2021). Bambara is the de facto (albeit not de jure) language of Malian politics (Cissé 2020, Dumestre 1994a), but it hardly plays a role as language of instruction in schools. Bambara has become an important language in other West African countries as well as in France and other European countries through migration. Members of the diaspora commonly regard it as the national language of Mali, but associate it with a status of West African lingua franca at the same time (Galtier 1995, Van Den Avenne 1998, 2001, 2004).

Fulfulde is one of the five regional lingua francas in Mali. It is spoken as an L1 and L2 in the regions of Mopti and Kayes. It is also present as an L1 in other parts of Mali, especially since 2012, when terrorism forced numerous people from central Mali to flee their home regions.

Pular and Malian (or Massina) Fulfulde are varieties of Fula, also known as Fulfulde (in a wider, more generic sense) or Peul, a language originally spoken in 19 countries between Mauritania, Senegal, Gambia, the two Guineas, Sierra Leone in the west, and Sudan in the east and which includes a huge range of varieties. Guinean Pular, also known as Pular of Futa Jallon, the main area of distribution, is the Fula variety best represented in the diaspora (Mohamadou 2017) – whether in other African countries, such as Angola (Niedrig 2003: 336,

339, A. Diallo pers. comm.), or in Europe, for instance among African refugees in Hamburg, Germany (Niedrig 2003: 336, 339).²

French is the only official language in West African countries with a French colonial past, including Guinea, Mali, Senegal, and Côte d'Ivoire, all places where Kajatu spent extended periods of time. While there was continuity in maintaining the colonizer's language in the administrative and education systems in the other former French colonies after independence, Guinea suspended French as the main language of education between 1968 and 1984 in an effort to decolonize (Barry 2014). Since its reintroduction as the only official language in administration and education, the importance of French has been growing, especially among young people planning to migrate, study, or travel abroad. They consider a knowledge of French to be a great advantage (Diallo 2021). Only an estimated 25 per cent of the Guinean population and a minority of Malians speak French (Cissé 2020).³ When Kajatu lived in Guinea and Mali, French was taught and learned almost exclusively at school. Kajatu did not have much schooling, if any. She learned French during her stays abroad, especially in Senegal and Côte d'Ivoire. Accordingly, she considers it an international language rather than associating it with her home country.

Table 1: Genetic classification of African languages relevant to this chapter.

Language	Cluster	Language family	Mainly spoken in
Pular	Fula	Atlantic	Guinea
Fulfulde	Fula	Atlantic	Mali
Susu		Mande	Guinea
Maninka	Manding	Mande	Guinea
Bambara	Manding	Mande	Mali

Table 1 summarizes the genetic classification of the African languages in Kajatu's repertoire. They belong to the Mande and Atlantic language families; Pular and Malian Fulfulde are geographical varieties of Fula, which belongs to the Atlantic languages, while Susu, Maninka, and Bambara are members of the Mande family.

²Niedrig (2003) reports on a research project with refugee youths from West Africa and other parts of the continent living in Hamburg at that time. A total of 32 out of the 73 people she consulted were proficient in Pular (Fula). Some of them had started to learn the language only after their arrival in Germany.

³According to *Ethnologue*, French is spoken by about 16 per cent of the Malian population (Eberhard et al. 2021).

Maninka and Bambara are closely related dialects of the Manding⁴ cluster. Susu is less closely related to Manding. Mande and Atlantic languages have long been in contact in many parts of West Africa, but they differ considerably from each other. The main typological differences between them are summarized in Table 2.

Table 2: Typological features of Mande and Atlantic languages (cf. Güldemann 2018: 151, Williamson & Blench 2000: 18–22).

Typological feature	Fula	Manding
word order in transitive clause	S-V-O-X	S-AUX-O-V-X
noun phrase (genitive modifier)	N+Gen	Gen+N
nominal morphology	noun class suffixes on the noun, initial consonant mutation as remnant of prefixes	absence of noun classes
adpositions	prepositions	postpositions/ prepositions

Mande word order in transitive clauses differs from the word order in Atlantic languages by the preverbal position of the object; adjuncts follow the verb. In Atlantic languages, the object is in the postverbal position. In a noun phrase with a genitive modifier, the head is initial in Atlantic languages, and final in Mande. Noun class systems are a characteristic feature of Atlantic languages. Fula marks noun classes on lexical nouns by means of suffixes, whereas nouns in Mande do not have any similar morphological marking. Adpositions in Atlantic languages precede the noun (i.e., they are prepositions), while Mande languages have mostly postpositions. Speakers proficient in languages of both families, Atlantic and Mande, have quite a rich repertoire of constructions at their disposal when they start to learn a typologically distant language such as German. Kajatu is clearly one such case.

4 Kajatu and her language-learning experiences

Kajatu, about 50 years old, draws seven points in the silhouette that represents her body (Figure 1) and attributes them to the following languages: the Guinean

⁴<https://www.ethnologue.com/subgroup/286/>

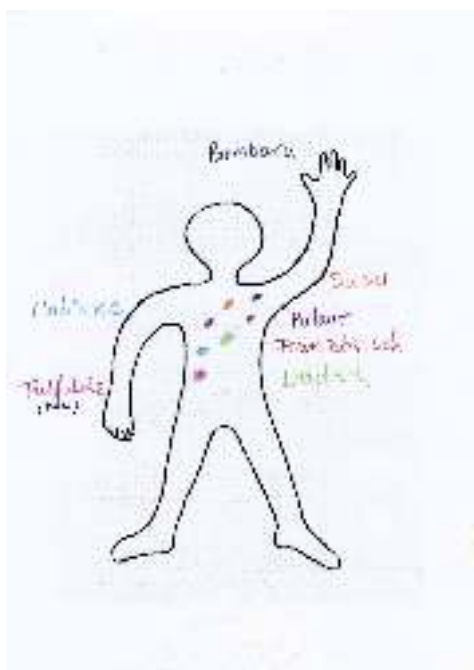


Figure 1: Kajatu's language portrait.

languages Pular, Susu, and Maninka (which, according to the French tradition, she calls “Malinke”), the Malian languages Bambara and Fulfulde, French, and German. Kajatu was born and partly raised in Guinea's capital, Conakry. She considers Guinean Pular her “Muttersprache” (‘mother tongue’), as it is the language that was used at home when she was a child. Pular is both her parents' main language. They were Fulbe who originally came from Dalaba, a town situated in the Mamou region in central Guinea. Kajatu learned her grandmother's language, Maninka, during the family's stays in Dalaba while on holiday. As a child, she moved to Bamako, the capital of Mali, to accompany her elder sister, who married a Malian Fula man. She grew up in Bamako from then on and spent 15 years there. She learned Bambara and the Malian Fula variety Fulfulde, which Kajatu refers to as “Fulfulde Bamako” or “Pular von Bamako”. She also spent several years in Senegal and in Côte d'Ivoire, moving there with her elder sister and her husband repeatedly for extended periods of time. There, the language she used outside the family was French. As a young woman, she travelled a lot for trading, among other places to Senegal and Nigeria, before she came to Germany more than 20 years ago and married a German. Kajatu has three grown-up

children and three younger foster children. German is the only family language. She uses Pular with her family in Guinea, Pular, Maninka, and Susu with her Guinean friends on a daily basis, Bambara with her Malian friends, and French with other friends and colleagues from West Africa outside Guinea. Kajatu travels frequently to Guinea and other countries in West Africa, and even when she does not travel, she speaks all the languages she mentions in connection with her language portrait with her relatives and friends in the Rhine-Main region and on the phone on a daily basis.

In Klein's (1986) terms, Kajatu's learning of Pular and Maninka can be classified as bilingual first language acquisition (Klein 1986: 15) with a clear dominance of Pular, which she learned from birth within her family environment. In contrast, the acquisition of Maninka was temporary, since it took place only in interactions with her grandmother during short stays in Dalaba. Kajatu started to learn Bambara and the Malian variety of Fula when she moved to Bamako. She does not mention a date or her age at the time of the move to Bamako, but some elements in the narrative suggest that she was at least four years old at the time. This would mark the earliest contact with and possible beginning of the learning process with regard to Bambara and Malian Fulfulde. According to Kajatu, she grew up in Bamako, where she also learned Bambara: "Bambanan, das habe ich in Bamako [...] weil ich bin gewachsen in Bamako." ('Bambara, I have [learned it] in Bamako, since I grew [up] in Bamako').

Her learning of Bambara and Malian Fula would qualify as child second language acquisition in Klein's terms (1986: 15), ranging between the age of three to four and puberty – the critical period for language acquisition (see also Matras 2020: 72ff).

It is difficult to tell whether Kajatu started to learn the additional languages Susu and French before or after puberty. The acquisition of German, in any case, would count as a clear instance of adult second language acquisition.

Kajatu learned all her languages, including French, spontaneously in everyday communication without systematic guidance. In Guinea and Mali, French is taught almost exclusively in the formal context of school lessons,⁵ but there are several (indirect) indications that Kajatu had little or no schooling. When asked how she learned French, she replied using the impersonal indefinite pronoun *man* in (1).

⁵This has changed in the last few years, as noted in Diallo (2021).

(1) German

Französisch, man lern das in die Schule.

French INDF learn that in DEF.SG.F.ACC? school

‘French, one learns it at school.’ [Kajatu 03:13]

The strategy of a generic account instead of a personal one is often used as a mechanism to avoid talking about personal experience when it comes to sensitive topics in personal narrations (de Fina 2019: 35). By saying that French is learned at school, Kajatu remains vague and perhaps purposefully leaves an insinuation of having attended school in place, without committing to it. Given that she spent time in Côte d’Ivoire, un-monitored learning by communicating in French is likely to have taken place. As Kajatu states concerning Côte d’Ivoire, “fast alle spricht Französisch” (‘Almost everybody speaks French’) [Kajatu 12:40].

We only know of one context of formal education that she underwent: adult education courses in Germany, which she attended irregularly and half-heartedly. Therefore, in her own view, she achieved little success in the acquisition of German.

Next to Pular, her “mother tongue”, Kajatu categorizes Susu and Maninka as languages of her homeland. She considers Susu Guinea’s main language: “Susu ist Hauptsprache von Guinea, [...] alle spricht das” (‘Susu is the main language of Guinea; everybody speaks it’), but also asserts, “Französisch [...], alle Herrn muss das verstehen” (‘French, all people have to understand it’) [Kajatu 16:35]. She does not evaluate her own proficiency in these languages, except for Malian Fulfulde, which she believes she speaks very well, and German, which she considers to be “not so good”.

(2) German

War nicht so lange [ref. der Besuch der Volkshochschulkurse], aber die Reste hab ich in die unterwegs genommen, überall ein bisschen aber wegen das mein Deutsch ist nicht so gut, aber mein Mann ist Deutscher.

([It] was not so long [referring to her attendance of German classes at evening school], but the rest I grasped along the way, everywhere a little bit, but for that reason my German is not so great. But my husband is German.) [Kajatu 14:28]

She appears critical of herself in this regard, but we should bear in mind other factors that come into play. For example, the interview was conducted in German, and down-playing her competence is certainly affected by politeness strategies

in the research interaction. At the same time, she is likely to measure her own German skills against the expectations of others and her own expectations in a context that emphasizes formal aspects of ‘correctness’ in language use. For all practical purposes, she is a proficient user of German, her husband’s language used constantly in the family home.

5 Tiers of language structure in a repertoire understanding

Building on the information (including typological properties) concerning the languages that characterize Kajatu’s repertoire (in Section 3) and the overview of her biography and mobility in West Africa (in Section 4), we will now look into three different sets of constructions more closely. Each pertains to a different level of linguistic analysis. We will move from the meanings of *in* in Section 5.1 to genitive constructions in Section 5.2 and, finally, to nasalization phenomena in Section 5.3. This order follows an (often implicit) assumption concerning speakers’ deliberate leverage, arguably decreasing from lexical choices to morphosyntax and, in particular, phonology that are learned early on, not used or influenced consciously, non-deliberate, and, therefore, hard to unlearn.

5.1 The meanings of *in*

In the interview, Kajatu mostly uses the preposition *in* to express static containment. In this sense, her use is no different from standard German practice – “to live in Germany; to be in the house, etc.”, as illustrated in (3).

- (3) German
ich bin **in** Guinea ge-bor-en
ʔis bin **?in** ginea geboren
I be.1SG.PRS **in** Guinea PTCP-be.born-PTCP
‘I was born **in** Guinea.’ [Kajatu 02:20]

This is typically the case with notions of genuine stable containment. In certain cases, she uses the German *in* in metaphorical containment settings. In (4), we see *in die Arbeit* (‘in work’) instead of either *bei der Arbeit* or *auf der Arbeit*, both of which are more common expressions in standard German meaning ‘at work’.

(4) German

Ich bin die Einzig-e in die Arbeit,
 ʔis bin di ʔäisige ʔin di ʔarbatə
 I be.1SG.PRS DEF.SG.F.NOM only.one-SG.F.NOM in DEF.F.ACC? work
 wer in Guinea komm-t
 vɛ ʔin ginea kɔm-tə
 REL in Guinea come-3SG.PRS
 'I am the only one at work who comes from Guinea.' [Kajatu 18:22]

This instance of a less common use of the German *in* probably relates to the fact that the German prepositions *bei* (for loose spatial association, like the French *chez* and English *by*) and *auf* 'on' (for – mostly vertical – support/attachment, as in the English *on*) are in fact not very intuitive in this particular context.

The semantic range of *in* in Kajatu's German systematically extends beyond its standard German scope of meaning in sentences that contain expressions of translational motion – not just "into" but, more strikingly and surprisingly, also "out of" a containing GROUND as in (5).

(5) German

mein-e Eltern komm-t in ein-e Stadt
 main-ə ʔɛltɛn kɔm-t ʔin ʔainə sat
 1SG.POSS-PL.NOM parents come-3SG.PRS in INDF-SG.F.ACC? town
 heiß-t Dalaba
 haɪs dalaba
 be.called-3SG.PRS Dalaba
 '...my parents come from a town called Dalaba.' [Kajatu 02:32]

This is not a slip of the tongue or a nonce occurrence. In (4) to (6), we find comparable examples of translational motion expressions containing *in*, but with a reversal in the directional meaning.

(6) German

Haben.wir ge-sag ok, aber wiesowieso in ...bei uns in Heimat wir
 ʔāmə gəzak okɛɪ ʔabɐ vizovizo ʔin ...baɪ ũs ʔin haɪmatə vɪɐ
 have.1PL PTCP-say ok but anyhow in ...at ours in home we
 wohn-en immer zusammen bis man heirate is, man
 vo:nənə ʔimɐ sʊza:mən bis man haɪratə ʔis man
 live-1PL.PRS always together until INDF marry.PTCP? be.3SG.PRS INDF

geht raus in Wohnung
 ge:t raus ?in vo:nūŋ
 go.3SG.PRS out in flat

‘We said, ok, but in any case, in our home country, we live always together, until one gets married, one moves out **from** home.’ [Kajatu 11:02]

In part, Kajatu’s use of the German spatial preposition *in* in (4) to (6) may go back to a systematic difference between verb- and satellite-framing languages (Croft et al. 2010, Talmy 1975, 1985). All languages Kajatu used prior to learning German are characterized as verb-framing in that theoretical approach. They express PATH notions including directionality by lexical verbs. In contrast, predominantly satellite-framing languages like German express them through prepositional phrases (including choice of preposition and case). In contrast, the locative phrases with *in* as used by Kajatu represent (container-like) GROUNDS, not PATH-denoting expressions.

Kajatu may at times struggle to express motion events with their conceptual components GROUND and PATH arranged differently in basic construction types in German.

In Pular, the meaning of ‘being/coming from’ corresponding to Example (5) would be expressed as in Example (7):

- (7) Fula (Pular, Futa Jallon)
 maw-be an ben ko saare inn-etee-nde Dalabaa
 parent-BE POSS1SG DEF.BE FOC town.NDE be.called-PTCP.PASS-NDE Dalaba
 nde iw-i
 DEF.NDE come.from-PFV
 ‘My parents come from a town called Dalaba.’ [M. Diallo, elicited example]

The phrase “from a town called Dalaba” is a locative argument licensed by the (PATH-containing) motion verb *iwugol* ‘to come (from)’. It is not introduced by a preposition. The three constructed examples in (8) show how PATH-related notions depend entirely on the lexical meaning of the verb. Again, there are no prepositions specifying directional (*from/to*; 8a, 8b) or boundary-related notions (such as semantic nuances distinguishing *into/to/towards*; 8b, 8c).

(8) Fula (Pular, Futa Jallon), constructed examples

- a. Ko Labe mi iw-i.
FOC Labe I come.from-REL.PFV
'I come from Labe.'
- b. Mi yah-ay Labe.
I go-IPFV Labe
'I will go to Labe.'
- c. Ko Labe mi iw-t-i.
FOC Labe I come.from-REV-REL.PFV
'It is to Labe that I returned.'

In Bambara, such locative phrases without an overt postposition exist, too. A place name carries inherent locative noun properties as illustrated in (9a).

(9) Bambara, constructed examples

- a. ù bé bó Dalaba
they IPFV.AFF exit Dalaba
'They come from Dalaba.'
- b. í bó-ra só-` kónɔ
I exit-PFV house-ART inside
'I came out of/went out of the house/left (from inside) the house.'
- c. í dòn-na só-` kónɔ
I enter-PFV house-ART inside
'I entered/went into the house.'

In contrast, (9b) and (9c) contain a postposition expressing interior space. In Bambara, this is the semantically transparent word *kónɔ* whose full lexical meaning 'belly' is available alongside the use as a postposition with the (approximate) meaning 'in(side)'. It is important to note that the postposition profiles a region in space that relates to the GROUND-expressing noun, not the PATH component of the translational motion; the latter is contained within the verb. The postpositional locative phrase does not change in shape, irrespective of whether, as in (9c), the inside of the house is entered into, or as in (9b), moved out from (see Dombrowsky-Hahn 2012 for a comprehensive account of the construal and expression of motion events in Bambara).

We should also draw attention to a general-purpose locative preposition, *ka*, in Pular. It serves to construe nouns as locative nouns in very general terms, as in

ka kammu ‘at, in, towards the sky’. Such locative expressions introduced by the preposition *ka* occur in a wide range of contexts. A restriction exists concerning the noun denoting the GROUND, which has to be unequivocally identifiable, expressed by a definite noun introduced in earlier discourse, or a unique or salient place (e.g., a town’s [main] square or market, or a settlement’s [only] river). The semantic range of the preposition is not restricted in terms of topological notion.

(10) Fula (Futa Jallon)

- a. o yett-i⁶ horde ka tenkere
 she take-PFV calabash ka shelf
 ‘She took a calabash **from** the shelf.’
- b. o yah-i ka caangol
 she go-PFV ka river
 ‘Elle alla à la rivière [She went to the river.]’
- c. o jas-i ngayka ka njaareendi
 she dig-PFV pit ka sand
 ‘She dug a pit **in** the sand.’ [Labatut n.d.: 84, confirmed by A. Diallo, pers. comm.]

As (10b) illustrates, prepositional phrases introduced by *ka* occur also in sentences referring to translational motion. As with the postpositional phrases in Bambara (9b, 9c), such locatives are GROUND phrases; they do not express PATH notions. These are part of the lexical make-up of motion verbs in relevant expressions, as illustrated in (11). In these examples, the directionality of translational motion does not affect the GROUND phrase *ka suudu*. The few examples illustrate that *ka suudu* can mean ‘into the house’, ‘out of the house’, or ‘to the house’, depending on the verb’s meaning.

(11) Fula (Pular, Futa Jallon)

- a. mi yalt-ay ka suudu
 I move.out.of-IPFV ka house
 ‘I will leave/move **out of** the house.’
- b. mi naat-i ka suudu
 I enter-PFV ka house
 ‘I entered/went **into** the house.’

⁶<y> in the orthography of Fula corresponds to the implosive palatal consonant noted in the IPA as [ɟ]. It differs from the approximant <y> [j]. y [ɟ] is one of three implosive stops in the language, next to [ɓ] and [ɗ].

- c. mi ar-t-i ka suudu
 I come-REV-PFV ka house
 'I came back (in)to the house.' [M. Diallo, elicited examples]

As in Bambara, place names do not usually require a preposition (see also (7) and (8) above).

In contrast to Bambara, the preposition *ka* refers not only to interior space or containing GROUNDS. In this regard, the two languages differ, but as to how they express PATH in motion events, Fula and Bambara both rely on the same conceptual pattern. It is the verb that expresses PATH, not a (spatial) adposition phrase. In Fula, we see this very clearly in the contrast between (8a) and (8c), where the only difference is the use of a related (derived) verb stem in (8c).

Conceptually, the verb-framing pattern prevails in both Bambara and Fula. Despite this similarity, Fula (with its multi-purpose preposition *ka* introducing general locative nouns) seems more likely than Bambara (with its semantically very transparent postposition *kónɔ* 'in[side]') to have served Kajatu as a conceptual model for the construal and expression of GROUNDS in German. The choice of the German *in* and its broadening in semantic scope may be a plain frequency effect, given that *in* is a common preposition in German.

Another strategy would be to omit the use of any preposition with German locatives. In fact, such examples – structurally similar to the Fula example in (8) – are not uncommon when Kajatu uses German.

(12) German

- a. habe ge-flogen Senegal, Elfenbeinküste, Lomé, Togo
 have.1SG.PRS PTCP-fly.PTCP Senegal Côte d'Ivoire Lomé Togo
 'I flew [to] Senegal, Côte d'Ivoire, Lomé, Togo.'
- b. aber wir waren Bamako zusammen
 but we be.1PL.PST Bamako together
 'But we've been [to?/in?] Bamako together.' [Kajatu 30:12]

Most examples of this type contain place names, and most express locative being in/at a place rather than translational motion, with some exceptions, like (12a), or ambiguous cases (12b). A larger sample of Kajatu's German could help us understand better when she uses the preposition *in* in the non-standard manner and when she resorts to not using any preposition at all for the GROUND phrase. Whatever triggers this distinction, it goes to show that the distribution of both constructions in Kajatu's German is not random but indicative of conventional

and recurrent patterns. It may represent a genuine functional split in her individual grammatical usage, impacted, but not predetermined, by the equivalent expressions in Fula and Bambara. We are not dealing with a plain ‘transfer’ of structures here that are, in principle, foreign to German. The emergence of usage patterns in Kajatu’s German rather draws equally on the various linguistic resources which we may associate with separate languages, but which are at Kajatu’s disposal simultaneously.

In the following sections, we will scrutinize this idea in other tiers of language structure – those that are arguably characterized more strongly by conventional norms, usage patterns, and formal means usually taken to belong to an individual language in a systemic perspective.

5.2 Noun phrases with a genitive modifier

Noun phrases with a genitive modifier present several differences in Pular (as representative of Fula), Bambara (representing Manding), and German. Among others, the languages differ with respect to the order of head and modifier. Roughly speaking, in Bambara, all noun phrases with a genitive modifier show the order of genitive modifier and head. In Pular, they show the reverse word order, that is, the head precedes the genitive modifier. Examples (13) to (16) illustrate the constructions in Pular and Bambara.

- (13) Manding (Bambara)

modifier head

ń fà-`

1SG father-ART

‘my father’

- (14) Manding (Bambara)

modifier head

Adama fà-`

Adama father-ART

‘Adama’s father’

- (15) Fula (Pular, Futa Jallon)

head modifier

baaba an

father POSS.1SG

‘my father’

(16) Fula (Pular, Futa Jallon)

head modifier
baaba Adama
father Adama
'Adama's father'

Neither Pular nor Bambara mark case on lexical nouns, contrary to German. In German, noun phrases with a genitive modifier vary according to case, dedicated possessive pronouns that agree in gender with the head, and the changing word order depending on the presence or absence of the preposition *von*, which functions as a link between head and modifier.

There are two types of construction which do not have the same distribution in German. One is a construction with the genitive case illustrated in (17). It pertains to a more formal style, but occurs in colloquial speech as well, in particular with pronominal possessive modifiers (*mein Vater* 'my father').

(17) German

- a. modifier head
mein Vater
POSS.1SG.M.NOM father.M.SG.NOM
'my father'
- b. modifier head
Adama-s Vater
Adama-GEN father.M.SG.NOM
'Adama's father'

The other construction has, as illustrated in (18), an overt prepositional linker. It generally belongs to a more colloquial register.

(18) German

- a. head link modifier
der Vater von mir
DEF.M.SG.NOM father.M.SG.NOM of 1SG.DAT
'my father (lit. the father of me)'
- b. head link modifier
der Vater von Adama
DEF.M.SG.NOM father.M.SG.NOM of Adama.DAT
'Adama's father (lit. the father of Adama)'

In the first type, the modifier is the possessive pronoun, agreeing in gender, case, and number with the noun it modifies (17a), or a modifying noun in the genitive case (17b). In both cases, the modifier precedes the head as in *mein Vater* ‘my father’ or *Adamas Vater* ‘Adama’s father’. In the second type, illustrated in (18a) and (18b), the preposition *von* links head and modifier, the latter being a pronoun or a noun in dative case.

Kajatu often uses the modifier–head genitive construction type in German. This order corresponds to the Bambara model. At times, however, expressions used by Kajatu are reminiscent of the Pular pattern, for instance in *Fulfulde Bamako* ‘Fulfulde of Bamako’. In standard speech, it would have required the linking preposition *von* ‘of’ in German (*Fulfulde von Bamako*).

During the interview, Kajatu uses a few complex constructions whose modifier is itself a noun phrase with a genitive modifier. One of them, shown in (19), conforms to the Manding (Bambara and Maninka) model,⁷ as shown in (20), which differs from both the Pular (21) and the German constructions (22).

- (19) German
 [[‘mein-e ‘Papa] ‘Mutter]
 POSS.1SG-? dad mother.F.NOM
 ‘my dad’s mother’ [Kajatu 11:30]

Both the noun phrase, which is itself a modifier, and the entire noun phrase are head-final like the corresponding Bambara construction (20). They differ from the head-initial Pular counterpart (21).

- (20) Manding (Bambara)
 [[‘ń fà-`] bá-`]
 1SG father-ART mother-ART
 ‘my father’s mother’
- (21) Fula (Pular, Futa Jallon)
 [neene [baaba an]]
 mother father POSS.1SG
 ‘my father’s mother’

In the corresponding standard German, the genitive modifier is head-final, but the overall construction is head-initial and displays the preposition *von* as a link between head and modifier, see (22).

⁷The prosody of the nominal phrase *meine ‘Papa ‘Mutter* ‘my dad mother’ – distinctly audible in the recording – clearly distinguishes the genitive construction from the compound *meine ‘Papamutter* ‘my dad-mother’, which, theoretically, would also be possible.

(22) German

a. Standard German

[die Mutter von [mein-em Papa]]
 DEF.F.SG.NOM mother.SG.F.NOM of POSS.1SG-M.DAT father
 ‘my dad’s mother, i.e., my paternal grandmother’

b. German literary style

[[mein-es Vater-s] Mutter]
 POSS.1SG-M.GEN father.SG-M.GEN mother.SG.F.NOM
 ‘my father’s mother’

Literary style as in (22b) is unlikely to have triggered Kajatu’s utterance in (20) despite the same word order. Note that using the colloquial form of address *Papa* ‘Daddy, Pa’ is stylistically awkward in this construction type (*meines Papas Mutter* ‘my Daddy’s mother’).

The other two complex genitive constructions of the three Kajatu uses in the interview correspond to standard German grammar (“der Mann von meiner Schwester” ‘my sister’s husband’ [Kajatu 05:31], “die Leiche von seine Papa” ‘the corpse of his father’ [Kajatu 42:22]).⁸ The third, “meine Papa Mutter” ‘my father’s mother’, which follows the Bambara/Manding model, is uttered when Kajatu relates that she learned Maninka from her paternal grandmother when visiting her. This suggests that the conversation topic triggered this particular construction.

5.3 Phonology: Nasal consonants and nasalized vowels

The realization of syllables with a nasal coda in German words provides evidence that Kajatu makes use of her entire repertoire. Table 3 provides a list of Kajatu’s realizations of the monosyllabic word *dann* ‘then’, a temporal adverb that occurs frequently in the biographical narrative. The examples show the following realizations:

- A nasalized vowel and a final velar–nasal consonant: [d̃aŋ]
- A nasalized vowel without a syllable coda: [d̃a]
- A nasalized vowel and a coda homorganic with the following consonant: [d̃ām] before [b] or [m], in the latter case resulting in the gemination of the final nasal as in [d̃ām:āmə]

⁸Standard German requires the dative in this syntactic frame: *die Leiche von seinem Papa* ‘the corpse of his father’.

- An oral vowel and the labiodental nasal [dam] before a labiodental [v]
- An oral vowel and the final [n], corresponding to the German standard pronunciation: [dan]
- The corresponding form with a nasalized vowel [dān]

Table 3: Kajatu's realizations of *dann* 'then'.

a.	dāŋ ... va nis me	dann war nicht mehr (06:19)
b.	dāŋ ... viē sin fehairatətə	dann wir sind verheiratet (09:30)
c.	dā binis	dann bin ich (08:37)
d.	dā ven di kində ʔis gəkōmə	dann wenn die Kinder ist gekommen (22:50)
e.	dām:ānə swestə va gəhairatət in Bamako	dann meine Schwester war geheiratet in Bamako (03:32)
f.	dām ... binis	dann bin ich (08:41)
g.	dam ʋa: ʔis di:zə.	dann war ich diese (07:37)
h.	dān habis haīmve:	dann hab ich Heimweh (32:26)
i.	dan mūsən viē mit das maxən	dann müssen wir mit das machen (33:33)

In careful speech, the standard German *dann* is realized as [dan] with a final alveolar nasal consonant; in less careful speech in informal German, different instances of regressive assimilation are encountered as well, for instance the geminated labial nasal resulting from a following [m]. Therefore, the assimilations perceived in Kajatu's speech do not necessarily result from phonological patterns in languages acquired during childhood. Some of her realizations, however, suggest that they are constructions inspired by her previously learned languages, among others the velar coda [dān] and the realization with a nasalized vowel with or without a coda, for instance [dām] and [dā]. A look at nasal consonants and nasal vowels in Pular and Manding allows us to test the hypothesis.

Unlike Pular and Manding, German has no nasal vowels. In Pular and the Manding varieties, nasal vowels interact with nasal consonants in a syllable coda. The processes observed suggest that nasal consonants in a syllable coda are underspecified for the place of articulation. The specification of the place of articulation is conditioned by the context to its right.

Depending on the context, in Pular there are three possible processes, P1, P2, and P3, schematized in Figure 2.

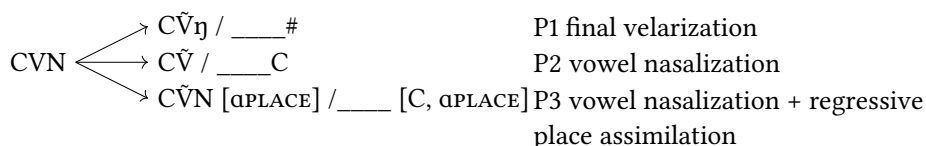


Figure 2: Phonological processes concerning nasal consonants as a syllable coda in Pular.

P1 (CVN → CṼŋ / ____ #) stands for the process in Pular by which a vowel preceding a coda N becomes nasalized, and the coda itself is realized as a velar nasal. It applies when the coda N occurs at the end of a respiratory unit, as in (23), or when the speaker hesitates before proceeding with the sentence.

(23) Fula (Pular), realization according to P1

a yar-ii ndiyan?
 [ʔa jar-i: ndijãŋ]
 2SG drink-PFV.ABS water
 ‘Did you drink water?’

When a nasal syllable coda N is followed by a consonant (e.g., *won-de* ‘live-INF’), either process P2 or P3 applies.

P2, (CVN → CṼ / ____ C), describes the nasalization of the vowel preceding a nasal consonant in a syllable coda, and the deletion of the nasal consonant, when followed by another consonant (Diallo 2000: 44).

P3, schematized as CṼN [αPLACE] / ____ [C, αPLACE], is similar to P2 in that the vowel becomes nasalized. The nasal consonant is not deleted but assimilates to the place of articulation of the following consonant.

The infinitive *won-de* is thus realized as [wõde] in (24) or as [wõnde] in (25) (A. Diallo, pers. comm.).

(24) Fula (Pular), realization according to P2

[wõde]
 won-de
 be-INF
 ‘be, live’

(25) Fula (Pular), realization according to P3

[wõnde]
 won-de
 be-INF
 ‘be, live’

When the consonant following the N of the coda is a nasal consonant, P3 results in the gemination of the consonant, as in (26).

- (26) Fula (Pular), realization according to P3
 [ˈwĩndãm:ɔˈlɛːtɛr]
 winnd-an mo leeter
 write-BEN him/her letter

There are different views on nasality for the varieties of Manding, Bambara, and Maninka (cf. among others Creissels 2009, Vydrin 2020). According to the most widely accepted view, the Bambara system of phonemes includes nasalized vowels \tilde{V} , and there are no closed syllables in the language (Dumestre 2003: 18, Vydrin 2019). According to Vydrin (2019: 17), when nasalized vowels are followed by a consonant, $/\tilde{V}/$ is realized $[VN]$, that is, as an oral vowel followed by a nasal consonant homorganic with the adjacent consonant. However, this is not the only realization possible, as will be shown below.

The analysis of nasality differs for Kita Maninka: Creissels (2009: 16–17) posits the existence of CVN syllables, where N is an underspecified nasal consonant manifest only through the nasalization of the vowel $[\tilde{V}]$ and, when it is followed by a plosive, through the realization of a nasal segment homorganic with the plosive. When applied to the Manding varieties learned and used by Kajatu, Creissels' analysis brings to the fore significant differences, as well as certain similarities, between Pular and Manding. According to this analysis, we have to admit a (C)VN syllable with an underspecified nasal coda next to the usual (C)V syllable. The realization of the (C)VN syllable is (C) \tilde{V} before a pause (cf. B1, illustrated in (27)), but when followed by a consonant, several processes are at work according to the speaker's dialect and idiolect, and, possibly, according to the context of occurrence (cf. Figure 3 and Examples 27 to 30).

- (27) Manding (Bambara), realization according to B1: \tilde{V} before sentence boundary
 à ká bòn [bò]
 3SG QUAL.AFF big
 'it is big'
- (28) Manding (Bambara), realization according to B2: nasalized vowel \tilde{V} before consonant
 à bìnà ókázôn sòrò [ókázɔ̃ː sòrò]
 3SG FUT.AFF opportunity find
 'he will get an opportunity (to come by car)' (Diallo 2003: 19)

- (29) Manding (Bambara), realization according to B3: \tilde{V} and regressive assimilation
- a. án má à yé [ám:á!jé]
1PL PFV.NEG 3SG see
‘we have not seen her’ (Sauvant & Molin 1956: 68, retranscribed by Diallo 2003: 15)
- b. ókázôn gale [ókázɔŋgálé]
opportunity first
‘the first opportunity’ (Diallo 2003: 22)
- (30) Manding (Bambara), realization according to B4: V and regressive assimilation
- bón-ba [bómbá] bón-dá [bóndá]
house-AUGM house-door
‘big house’ ‘house door’ (Vydrin 2019: 17)

Some processes correspond in Bambara and Pular. When a consonant follows a (C)VN syllable, the vowel becomes nasalized and the nasal consonant is either deleted (B2[=P2]) or its place of articulation assimilates to that of the following consonant (B3[=P3]). Bambara has an additional realization when a consonant follows, called B4 in Figure 3, in that the nasal assimilates to the following consonant and the preceding vowel is not nasalized.

Bambara and Pular also differ with respect to the realization of (C)VN syllables before a pause; they are realized with a final nasal vowel (C) \tilde{V} in Bambara (B1), whereas they end on a velar nasal in Pular (P1).

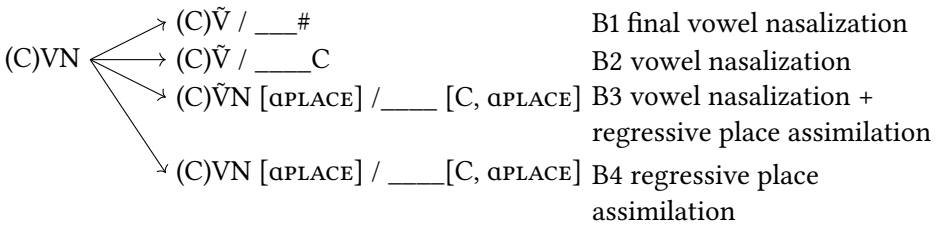


Figure 3: Phonological processes concerning nasal consonants as a syllable coda in Bambara.

Kajatu’s realizations of the German *dann* are equivalent to all possible realizations of CVN syllables in German, Pular, and Bambara/Maninka, with the exception of *dā* before a pause (which would correspond to B1). The adverb ‘then’ does

not occur in isolation or at the end of an utterance in Kajatu's speech, and before a suspensive pause, Kajatu realizes it corresponding to P1. Next to the Pular-specific process P1 and the Bambara-specific process B4, the processes equivalent in both languages, P2/B2 and P3/B3, are attested as well. It is thus impossible to assign a more important role to either the L1 or the L2. On the basis of the data at our disposal, it is difficult to discover a regular distribution pattern and, consequently, to say whether we are dealing with free variation or whether this repertoire of realizations indexes something other than the speaker's individual multilingualism.

6 Discussion and conclusion

We have dissected Kajatu's German, choosing three features, and built an argument based on a description of linguistic data that make Kajatu's German look very non-standard and exceptional. Her use of German is certainly quite unique, as the result of her particular biography and her mobility throughout her life. The linguistic codes and experiences with various language ecologies are specific to Kajatu's case. But the mechanisms at work, we are convinced, are not unusual. A multilingual individual moving between spaces and among societies in West Africa, each characterized by various languages in different ways, will necessarily lead to a complex language repertoire. The German language data from the interview that we have focused on bear witness to this. Kajatu communicates proficiently, and that is sufficient to warrant a rich description and a systematic analysis, including that of structural properties and regular patterns.

Kajatu's personal narrative emphasizes continuities rather than ruptures between first and second language acquisition. There are no clear first language interference effects, but a broad range of linguistic resources, originating in the various codes at her disposal, which she relies on when speaking German.

What shapes and determines Kajatu's choices and possibly directions of transfer (for want of a better term) when moving between languages and drawing on different languages cannot be predicted easily. Different hypotheses have been launched and discussed at length with regard to how transfer or interference take place in speakers' minds. In addition to the putatively compelling effects of an individual's native language, it has been suggested that other languages learned afterwards may be equally significant and leave an imprint on languages learned even later in life – perhaps because the effort is more conscious, or perhaps because the learning processes occur closer to each other in time (see Vildomec 1963, Gass & Selinker 2008: 154).

Typological and other similarities could be expected to play a role as well, possibly facilitating certain ‘routes of transfer’ of particular structures. The fact that certain areas of grammar are more highly structured than others, organized into strict paradigms and deeply entrenched categories (phonology more so than morphosyntax, and both outranking the lexicon), suggests a sequence from early to later in first language acquisition, which may reflect their stability and likelihood to be maintained when an individual speaker learns and uses a ‘non-native’ language.

Overall then, tracing possible ‘interference’ scenarios is anything but simple or straightforward. And yet, choices of constructions, linguistic items, and structural properties are not purely chaotic either, even in the case of speakers like Kajatu, where notions of native competence, sequencing language acquisition into categories like those offered by Klein (1986), are pushed to their limits.

In Kajatu’s German, we find structures that do not commonly form part of the repertoire of ‘native’ speakers. Some of these relate to comparable structures in one of her other languages. The complex genitive constructions that correspond to the word order they have in Maninka or Bambara are a Manding feature. The CŨŋ realization of syllables with a nasal coda at the end of a respiratory unit can be found in Pular, but not other languages. The same is true for Kajatu’s generalization of the German preposition *in*, which reflects the wide range of functions of *ka* in Pular when compared to the more restricted meaning of the Bambara *kóno* ‘in’.

Not all instances of transfer have to be associated with a language of origin, though. Some typological features are common to all Manding or even all Mande languages in Kajatu’s repertoire (i.e., Susu, in addition to Maninka and Bambara), and a specific ‘loan trajectory’ associated with a donor language is not a reasonable assumption. The same is true for those features that cannot be attributed to either Fula or Manding. Both converge in the realization of syllable codas as nasal vowels (with the nasal consonant assimilating or being deleted). We should also bear in mind that these are partly found in informal German as well.

To return to the bigger picture, we can trace features in Kajatu’s German back to different linguistic codes in her repertoire. We do not assume that there are various grammars in her mind that compete or impact on each other in real-time speech production. This, however, does not imply the absence of clear usage patterns, stabilizing choices that characterize Kajatu’s German in a particular way.

Structural features that are associated with one language or linguistic code, but occur in another (in this case, German), can be found in all tiers of linguistic

structure. They are not ranked from phonetics to lexical semantics with a decreasing likelihood of ‘native-language impact’, nor are they ordered in terms of which language is ‘most native-like’ for Kajatu.

Contact linguistics offers a number of fundamental generalizations concerning the likelihood of transfers of language structures across varieties. The logic of borrowing scales derives from assumptions about what language structures lend themselves to being transferred between languages in contact. Situations of language shift may also lead to similarities across languages, as when phonetic–phonological features are preserved and introduced into the languages shifted towards. Our observation that various sub-domains of language structure do not differ systematically in terms of how stable and native-language-proportioned they are appears to go against these – quite widely accepted – tenets. How can they be reconciled with the case under study?

We believe that different kinds of language ecologies need to be distinguished. A predominant tendency towards monolingual ideologies and practices in a given area or community may favour mechanisms captured adequately, for example by the ‘orderly’ notion of a borrowing scale.

Other settings and scenarios may be substantially different. It could be argued that this is simply a matter of scale, assuming that extremely multilingual scenarios may complicate the picture, but do not force us to assume different qualitative mechanisms (and limitations) concerning the transfer of language structures. Most importantly, the logic of first language acquisition as essentially distinct from (inherently imperfect) learning by (adult) second language learners could and would be upheld in this view, which relates individual processes of language learning to macro-sociolinguistic effects on entire speech communities (cf. Thomason & Kaufman 1988, see also Trudgill 2011).

On a more micro-level scale, the case discussed in this chapter cautions us against merely extrapolating patterns from individual language-learning dynamics to processes at a macro level. It also illustrates the significance of what has long been known, but still has not informed linguistic theory-building widely enough; namely, the possibility of fundamentally different processes and mechanisms which are highly diverse in a linguistic context.

That is one reason we regret not having had more chances to attend situations in which Kajatu interacts with fully fledged speakers of Fula, Bambara, or other languages. If we are right in assuming that she communicates by speaking according to necessary context cues, but drawing on her entire repertoire, we might expect structures from various languages to be used in such situations. The standard view might propose lexical borrowing from German as a possible,

perhaps likely feature of African languages used in the local diaspora, but not others. We are not so sure about the latter, but that remains to be seen.

Abbreviations

AFF	Affirmative
AUGM	Augmentative
BE	noun class morpheme – class BE (human plural)
NDE	noun class morpheme – class NDE
QUAL	auxiliary in quality expressing clause
RELPFV	relative perfective

Acknowledgements

First of all, we would like to express our gratitude to Kajatu for her willingness to share with us her experiences about language learning and her linguistic repertoire. A jaaraama! We would also like to express our great appreciation to our colleague, Abdourahmane Diallo, for his assistance with the Pular examples and to Klaudia's friend M. Diallo for her help in translating selected sentences into Pular. We are grateful to the Alliance of the Rhine-Main-Universities for the financial support that made this research possible. Finally, we wish to thank the organizers and the participants of the CoTiSp conference for their valuable questions and suggestions.

References

- Barry, Alpha O. 2014. Le français et les langues nationales en Guinée: complémentarité ou accommodation. In Musanji Nglasso-Mwatha (ed.), *Le français et les langues partenaires: Convivialité et compétitivité*, 45–64. Pessac: Presses Universitaires de Bordeaux.
- Busch, Brigitta. 2018. The language portrait in multilingualism research: Theoretical and methodological considerations. *Working Papers in Urban Language & Literacies* 236. 1–13.
- Cissé, Ibrahima A. H. 2020. *Langues et pouvoirs en période de crises au Mali. Cours donné au Darmouth College (États Unis) et discuté à University of Helsinki (Finlande)*. Powerpoint presentation. University of Helsinki.
- Creissels, Denis. 2009. *Le malinké de Kita: Un parler mandingue de l'ouest du Mali*. Cologne: Köppe.

- Croft, William A., Jóhanna Barðdal, Willem Hollmann, Violeta Sotirova & Chiaki Taoka. 2010. Revising Talmy's typological classification of complex event constructions. In Hans C. Boas (ed.), *Contrastive studies in construction grammar*, 201–236. Amsterdam: John Benjamins.
- de Fina, Anna. 2019. The interview as an interactional event. In Peter L. Patrick, Monika S. Schmid & Karin Zwaan (eds.), *Language analysis for the determination of origin. Current perspectives and new directions*, 21–40. Cham: Springer.
- Diallo, Abdourahmane. 2000. *Grammaire descriptive du pular du Fuuta Jaloo, Guinée*. Frankfurt am Main; New York: P. Lang.
- Diallo, Abdourahmane. 2021. Quelques repères historiques et sociolinguistiques sur l'implantation du français à Conakry. *The Mouth* 8. 65–84.
- Diallo, Mohamed Larabi. 2003. L'assimilation en bamanankan. *Mandenkan* 38. 15–45.
- Dombrowsky-Hahn, Klaudia. 2012. Motion events in Bambara (Mande). *Journal of African Languages and Linguistics* 33(1). 37–65.
- Dumestre, Gérard. 1994a. La dynamique des langues au Mali: Le trinôme langues régionales - bambara - français. In Gérard Dumestre (ed.), *Stratégies communicatives au Mali: Langues régionales, bambara, français*, 3–11. Paris: Didier Érudition.
- Dumestre, Gérard (ed.). 1994b. *Stratégies communicatives au Mali: Langues régionales, bambara, français*. Paris: Didier Érudition.
- Dumestre, Gérard. 1998. Le bambara véhiculaire du Mali. *Faits de Langues* 6. 121–127.
- Dumestre, Gérard. 2003. *Grammaire fondamentale du bambara*. Paris: Karthala.
- Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2021. *Ethnologue: Languages of the world*. 24th edn. Dallas, Texas: SIL International.
- Galtier, Gérard. 1995. La dynamique des langues africaines dans la communauté malienne de Paris. *Mandenkan* 30. 1–40.
- Gass, Susan M. & Larry Selinker. 2008. *Second language acquisition: An introductory course*. 3rd edn. New York: Routledge/Taylor & Francis Group.
- Güldemann, Tom (ed.). 2018. *The languages and linguistics of Africa*. Berlin: De Gruyter.
- Klein, Wolfgang. 1986. *Second language acquisition*. Cambridge; New York: Cambridge University Press.
- Labatut, Roger. N.d. Initiation au peul. (with collaboration of Mme Balde, Samassekou, Adama; Ousmanou, Bouba; M'bodj, Djibi; Diallo, Hama; Aliou, Mohamadou). Paris: INALCO.
- Lüpke, Friederike & Anne Storch. 2013. *Repertoires and choices in African languages*. Boston: De Gruyter Mouton.

- Makoni, Sinfree & Alastair Pennycook (eds.). 2007. *Disinventing and reconstituting languages*. Clevedon; Buffalo: Multilingual Matters.
- Makoni, Sinfree & Alastair Pennycook. 2012. Disinventing multilingualism: From monological multilingualism to multilingua francas. In Marilyn Martin-Jones, Adrian Blackledge & Angela Creese (eds.), *The Routledge handbook of multilingualism*, 439–453. New York: Routledge.
- Matras, Yaron. 2020. *Language contact*. 2nd edn. Cambridge: Cambridge University Press.
- Mohamadou, Aliou. 2017. *Présentation du peul*. Powerpoint presentation. https://www.canal-u.tv/video/site_pouchet_cnrs/presentation_du_peul.39075.
- Niedrig, Heike. 2003. Bildungsinstitutionen im Spiegel der sprachlichen Ressourcen von afrikanischen Flüchtlingsjugendlichen. In Ursula Neumann, Heike Niedrig, Joachim Schroeder & Louis Henri Seukwa (eds.), *Lernen am Rande der Gesellschaft. Bildungsinstitutionen im Spiegel von Flüchtlingsbiographien*, 303–346. Münster: Waxmann.
- Pennycook, Alastair. 2010. *Language as a local practice*. Milton Park: Taylor & Francis.
- Pennycook, Alastair. 2018. Repertoires, registers and linguistic diversity. In Angela Creese & Adrian Blackledge (eds.), *The Routledge handbook of language and superdiversity: An interdisciplinary perspective*, 3–15. Abingdon & New York: Routledge.
- Sauvant, Emile & Paul Marie Molin. 1956. *Grammaire Bambara*. Issy-les-Moulineaux: Les Presses Missionnaires.
- Storch, Anne. 2016. Communicative repertoires in African languages. In *Oxford research encyclopedia of linguistics*. DOI: <https://doi.org/10.1093/acrefore/9780199384655.013.155>.
- Talmy, Leonard. 1975. Semantics and syntax of motion. In John P. Kimball (ed.), *Syntax and semantics*, vol. 4, 181–238. New York: Academic Press.
- Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (ed.), *Language typology and lexical descriptions*. Vol. III: Grammatical categories and the lexicon, 36–149. Cambridge: Cambridge University Press.
- Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford; New York: Oxford University Press.
- Van Den Avenne, Cécile. 1998. Le bambara dans la migration: identité et véhicularité. In Suzy Platiel & Raphaël Kabore (eds.), *Les langues d'Afrique subsaharienne*, 129–134. Paris: Ophrys.

- Van Den Avenne, Cécile. 2001. De l'expérience plurilingue à l'expérience diglossique. Migrants maliens en France (From multilingualism to diglossia: Malian migrants in France). *Cahiers d'Études Africaines* 41(163/164). 619–636.
- Van Den Avenne, Cécile. 2004. *Changer de vie, changer de langues: paroles de migrants entre le Mali et Marseille*. Paris: L'Harmattan.
- Vildomec, Veroboj. 1963. *Multilingualism*. Leiden: A.W. Sythoff.
- Vydrin, Valentin. 2019. *Cours de grammaire bambara*. Paris: Presses de l'Inalco.
- Vydrin, Valentin. 2020. Vowel elision and reduction in Bambara. *Italian Journal of Linguistics* 32(1). 103–124.
- Weinreich, Uriel. 1953. *Languages in contact: Findings and problems*. The Hague: Mouton.
- Williamson, Kay & Roger Blench. 2000. Niger-Congo. In Bernd Heine & Derek Nurse (eds.), *African languages. An introduction*, 11–42. Cambridge: Cambridge University Press.

Chapter 8

Asymmetrical power relations between languages of Equatorial Guinea: Views from the migration context in Madrid

Sandra Schlumpf

University of Basel

Equatorial Guinea is a small yet culturally and linguistically diverse country in Central Africa. It is characterized by complex contact situations between languages that differ widely in terms of status, prestige, and official recognition: three official European languages (Spanish, French, and Portuguese), several African languages (which all belong to the Bantu family), and two languages of mixed origins (the Portuguese-based Creole Fá d'Ambô and Pichi, a local variety of English). This linguistic setting shows several asymmetrical power relations, for example between European and African languages, between the African languages of different ethnic groups, and between standard English and Pichi. In this chapter, I focus on language attitudes of Spanish-speaking Equatoguineans toward three languages spoken in their country: Bubi, Fang, and Pichi. I work with a corpus of 24 semi-structured life-story interviews conducted between 2017 and 2018 with Equatoguineans living in Madrid, Spain. I explain the observed language attitudes in this interview corpus with the help of language ideologies and the sociohistorical and cultural background of the speakers. This allows me to recognize specific implications of their language attitudes, which sometimes show clear asymmetries in the distribution of prestige and power between the languages analyzed. Finally, I mention specific problems relevant to the topic of this chapter, such as the maintenance of colonial hierarchies in postcolonial times, the transmission of conservative and often Eurocentric language ideologies and classifications of language phenomena, the projection of Western standards on African realities, and the lack of understanding of completely individual linguistic settings.



1 Introduction

Equatorial Guinea is the only Spanish-speaking country in Africa today, aside from the sensitive case of Western Sahara. It is located on the geographical and ideological periphery of the Spanish-speaking map and is often forgotten or treated as a small, exotic appendix in linguistic handbooks on the Spanish language. This is also true of Equatorial Guinea's history, cultures, and, in particular, literature (e.g., Mbomio Bacheng 2011; Trujillo 2012; Repinecz 2019; Riochí Siafá 2021). This situation of invisibility has been intensified by historical and political circumstances since the country gained independence in 1968, such as both Equatoguinean regimes (especially the first dictatorship of Francisco Macías Nguema Bidyogo from 1968 to 1979), the Francoist dictatorship during the first wave of mass emigration from Equatorial Guinea to Spain, and the categorization of Equatorial Guinea as “classified material” by the Spanish Law of Official Secrets in the 1970s (Schlumpf 2019: 290-292). Even today, people in Spain show little interest in this former colony in Africa, and as such, they have very limited knowledge about it. The country and its history are hardly ever part of the school or university curriculum, and it is only recently that a more critical, postcolonial, or even decolonial perspective on Spain's colonial role in Equatorial Guinea has emerged (e.g., Aixelà-Cabré 2020).

Equatoguinean society is characterized by its multiethnicity and multilingualism, representing a good example of the African polyglossia pyramid described by Wolff (2016: 210). Despite the small size of the country (approximately 28,000 km²), it presents a diverse linguistic setting with languages of different origins, levels of official status, and functionalities. Several asymmetrical power relations can be observed (cf. Martín Rojo 2017), for example between the African languages of the Bantu family and the official European languages, between the African languages of different ethnic groups, between the local English variety known as Pichi¹ and standard English, and between Equatoguinean Spanish and other varieties of Spanish, especially the Peninsular variety of the former colonial power.

In this chapter, I focus on Equatoguineans' language attitudes toward different languages spoken in their country, a topic which has barely been analyzed to

¹There are different hypotheses about the linguistic status of Pichi, the most frequent are those who describe it as a Pidgin English (Lipski 1992), an Afro-Caribbean English Lexifier Creole (Yakpo 2013a; Yakpo 2013b) or as a (new) language (Lipski 2012). As the aim of this paper is not the discussion of the linguistic classification of Pichi, I use the term *variety* as a more neutral option.

date.² What language perceptions and attitudes can be found today within the Equatoguinean community living in Madrid? Which functions, values, and levels of prestige do Equatoguineans connect to different languages? Which positive and negative associations arise? I will look at these questions with regard to three languages spoken in Equatorial Guinea: two Bantu languages (Bubi and Fang) and Pichi. None of them has the status of an official language in contemporary Equatorial Guinea.

In the following section, I present my corpus of sociolinguistic interviews, which forms the core of this study. Then, after a review of the importance of investigating language attitudes in postcolonial contexts (Section 3.1), I will continue with a description of the attitudes of Equatoguineans toward two African Bantu languages: Bubi and Fang (Section 3.2). I will then offer an excursus concerning the use of the terms *lengua* ‘language’ and *dialecto* ‘dialect’ in my corpus (Section 3.3). Finally, I will present the language attitudes toward Pichi (Section 3.4). The chapter finishes with a short conclusion (Section 4).

2 Corpus of sociolinguistic interviews with Equatoguineans in Madrid

The results that I present in the following sections are based on my corpus of semi-directed, sociolinguistic interviews, which I conducted with Equatoguineans in Madrid, Spain, from 2017 to 2018. All the interviewees were born in Equatorial Guinea but were living in the Autonomous Community of Madrid at the time of their interview. The Equatoguinean community in Spain is of special relevance because it is the largest outside the country. As of January 1, 2021, approximately 23,000 people born in Equatorial Guinea were living in Spain. Almost 40 percent (8,791 people as of January 1, 2021) were registered in the Autonomous Community of Madrid, hence the interest of this central region around the Spanish capital for this survey.³

The corpus of interviews with Equatoguineans on which this study is based was collected in a migration context. The interviews are composed of different

²Some exceptions are the classical studies of the 1980s and 1990s done by Antonio Quilis and Celia Casado-Fresnillo (Quilis 1983; 1988; Quilis & Casado-Fresnillo 1993), the very short follow-up study by Chirilă (2015), and the publications by Schlumpf (2018; 2020a; 2020b).

³For more information about the interview corpus and the Equatoguinean community in Spain, see Schlumpf (2021b: 349-358). The latest demographic data are drawn from the Spanish National Institute of Statistics: <https://www.ine.es/dynt3/inebase/es/index.htm?type=pcaxis&path=/t20/e245/p08/&file=pcaxis&dh=0&capsel=1> (accessed in March 2022).

thematic modules around the following main topics: the interviewees' arrival in Spain, their past in their country of origin (Equatorial Guinea), their adaptation to life in Madrid and differences to their home country, their education and professional situations, their families, and their plans and expectations for the future. All the modules include linguistic and sociolinguistic questions, for example concerning the use and functions of different languages, the prestige and values of languages and varieties, and specific dialectological features.

In total, I carried out 46 interviews, all of which were between 40 and 90 minutes in length. From these, I selected 24 interviews that form my main corpus. These 24 interviews are evenly distributed among three sociolinguistic variables (see Table 1): ethnic group (B = Bubi/F = Fang), sex (H = male [*hombre*]/M = female [*mujer*]), and duration of residence in Spain (−8 = up to eight years/+8 = more than eight years). For each possible combination of the three variables, the main corpus contains three participants.⁴

Table 1: Equatoguineans in Madrid: composition of the main corpus (code of participant number 1: 01_−8HB).

	Bubi (B)		Fang (F)	
	Male (H) [<i>hombre</i>]	Female (M) [<i>mujer</i>]	Male (H) [<i>hombre</i>]	Female (M) [<i>mujer</i>]
Up to 8 years in Spain (−8)	3 [01–03]	3 [04–06]	3 [13–15]	3 [16–18]
More than 8 years in Spain (+8)	3 [07–09]	3 [10–12]	3 [19–21]	3 [22–24]

All interviews were audio recorded, after which the 24 interviews for the main corpus were transcribed entirely.⁵ In addition, the main corpus was prepared for

⁴The same three variables are used to establish the participants' codes, meaning the code of the first participant is 01_−8HB (= number of the participant + time spent in Spain + sex + ethnic group). These codes identify each speaker providing the quotes presented in this chapter.

⁵For the most part, the transcriptions follow the recommendations created for the *Proyecto para el Estudio Sociolingüístico del Español de España y de América* (PRESEEA), summarized in Moreno Fernández (2021). However, in this chapter, I present the examples in their unlabeled version (i.e., without metalinguistic annotations) and using some additional marks: “/” = short pause; “//” = longer pause; “-” = interruption; “:” = lengthening; “h:” = exhalation; “.h” = short inhalation; “.h:” = longer inhalation. The codes “Ie.” and “Ir.” refer to the interviewee and the interviewer, respectively.

digital processing and analysis. Firstly, all utterances of the Equatoguinean interviewees were digitized and lemmatized using SketchEngine (corpus size: 185,185 words).⁶ Secondly, they were introduced into the online spreadsheet–database hybrid Airtable,⁷ which allowed me to create concordances and annotate and analyze the linguistic, discursive, and thematic phenomena.

The results presented in the following section are based on the qualitative analysis of the 24 interviews and on the concordances in Airtable, which contain all the occurrences of the names of the languages studied, and certain related linguistic terms: *lengua* ‘language’, *idioma* ‘language’, *dialecto* ‘dialect’, and *lenguaje* ‘language’.

3 Language attitudes toward different languages spoken in Equatorial Guinea

3.1 On the importance of language attitudes in postcolonial and decolonial studies

In this chapter, I understand *attitudes* as follows, according to Gallois et al. (2007: 596, emphasis in original):

This concept [...] represents the judgements that people tend to make and generalise about an object (social or otherwise) outside themselves. Attitudes are formed toward a particular entity, called the attitude object, and are generally theorised to contain three main components: cognitive, conative (behavioural), and affective. The *cognitive component* in attitude formation and maintenance represents a person’s beliefs and thoughts about the attitude object, without any positive or negative tone. The *conative component* is a predisposition to behave in accordance with the beliefs. Finally, the *affective component* represents an emotional reaction, positive or negative, that accompanies the beliefs.

⁶SketchEngine (<https://www.sketchengine.eu/>) is a platform that provides both corpus management and advanced text analysis. It was developed for linguistic researchers, lexicographers, translators, language teachers, and students. Tagging of texts in Spanish is supported by the FreeLing morphological tagger with a morphological dictionary obtained from different open-source projects with over 555,000 word forms. More details about SketchEngine can be found in Kilgariff et al. (2014).

⁷Airtable (<https://airtable.com/>) is a free online spreadsheet–database hybrid. It enables the user to annotate all concordances in a database, making them easy to retrieve, change, or extend, and includes all the functionalities common in spreadsheet tools, such as straightforward importing or exporting and intuitive filtering and sorting of entries. In addition, Airtable supports easy annotation schema definition with checkboxes, drop-down lists, and comments.

More precisely, according to Moreno Fernández (2015: 177-178), a *language attitude* refers to the following:

[A] manifestation of the social attitude of individuals, distinguished by focusing on and referring specifically to both the language and the use made of it in society [...]. The attitude toward language and its use becomes especially attractive when one appreciates to its full extent the fact that languages are not only carriers of certain linguistic forms and attributes, but are also capable of transmitting social meanings or connotations, as well as sentimental values. Cultural norms and marks of a group are transmitted or emphasized through language.⁸

Therefore, the analysis of language attitudes gives important insights into the values that certain languages and linguistic varieties have in a specific context, and about what roles languages play in the identity of individuals and speech communities. I consider this a very important question, especially in postcolonial contexts, where usually several, very different languages – and different communities with their specific cultural and linguistic backgrounds – exist together in a certain space and come into contact on a daily basis.

This approach to language attitudes is strongly connected to the study of language ideologies, understood as systems of ideas that connect linguistic phenomena with concrete cultural, political and/or social contexts (Del Valle 2007: 19–20). As already stated by Blommaert (2006: 518):

The study of language ideology grew out of linguistic anthropology and shares the basic preoccupation in this tradition of investigating the nexus of language and culture. It does so by introducing another level of cultural structuring in language: the language-ideological, indexical metalinguistic level. This level drives the development of linguistic structure [...] and it organizes the social, political, and historical framing of language and language use.

⁸Own translation. The original quote is: “La *actitud lingüística* es una manifestación de la actitud social de los individuos, distinguida por centrarse y referirse específicamente tanto a la lengua como al uso que de ella se hace en sociedad [...]. La actitud ante la lengua y su uso se convierte en especialmente atractiva cuando se aprecia en su justa magnitud el hecho de que las lenguas no sólo son portadoras de unas formas y unos atributos lingüísticos determinados, sino que también son capaces de transmitir significados o connotaciones sociales, además de valores sentimentales. Las normas y marcas culturales de un grupo se transmiten o enfatizan por medio de la lengua.”

Both language attitudes and language ideologies are always bound to a particular cultural context, which is “structured across intersubjective and institutional scales” (Rosa & Burdick 2017: 117). Therefore, language attitudes, which also reflect the language ideologies of the speakers, allow us to understand what functions and values different languages in multilingual contexts have at a precise moment in time. This enables us to uncover possible asymmetries and power relations between languages and language communities, and how these are reflected in the beliefs of the speakers (Bouchard 2022: 1). Terminology is also important in this regard (see Section 3.3).

Finally, the study of language attitudes of speakers of a lesser-known variety of the Spanish language responds to the call to decolonize current knowledge in the sense of making visible other ways of seeing that have not been shown before. Lara Delgado (2015) calls this the *empowerment of the epistemologies* which traditionally have been silenced, whereas Shepherd (2020: 318) talks about the “willingness to take seriously the knowledge traditions of the dispossessed”. This should lead to the revision and correction of established knowledge. In other words, “it is necessary to incorporate other analyses that allow us to have a general view and visibilize other components that are established through: epistemic, spiritual, racial/ethnic and gender/sexual hierarchies; products of [...] the European/Euro-North American capitalist/patriarchal modern/colonial system” (Gómez Vélez et al. 2017: 49–50).⁹ In fact, placing the focus on the Spanish-speaking Equatoguineans affords insights into a community hardly ever studied from a sociolinguistic perspective. Thus, underrepresented voices and beliefs become visible. Moreover, a decolonial perspective makes it possible to disclose colonial legacies in the description and recognition of languages and communicative practices. In fact, until today: “Representations and meanings attached to languages, linguistic forms, and practices structure and stratify social spaces in a way that reproduces colonial hierarchies” (Bouchard 2022: 5). These consequences of coloniality and colonial power asymmetries in linguistic contexts are what Veronelli calls the *coloniality of language* (see Veronelli 2015; 2016; 2019).

In order to find out about possible relations, disencounters, and asymmetrical power relations between different languages spoken in Equatorial Guinea, I will look at the language attitudes toward Bubi, Fang, and Pichi expressed by Equatoguineans who live in the Autonomous Community of Madrid, the political heart of the former colonial power.

⁹Own translation. The original quote is: “Es necesario incorporar otros análisis que permitan tener una mirada general y visibilizar otros componentes que se establecen a través de: las jerarquías epistémicas, espirituales, raciales/étnicas y de género/sexualidad; productos del [...] sistema europeo/euro-norteamericano capitalista/patriarcal moderno/colonial.”

3.2 Equatoguineans' attitudes toward two Bantu languages: Bubi and Fang

Bubi and Fang are the two most frequently spoken Bantu languages in Equatorial Guinea. Both are part of the Bantu languages of Zone A according to Guthrie's famous classification of the Bantu languages, and therefore are included in the so-called Western Bantu languages (cf. Guthrie 1953; 1971).

In my 24 interview transcripts, the term *bubi* appears 388 times; 191 occurrences refer to the Bubi language, while the other 197 refer to a Bubi person or the Bubi ethnic group. In the case of *fang*, the total number of occurrences in the interviews is 528, of which 283 refer to the Fang language and 240 to the Fang ethnic group or a Fang person; a few cases are not entirely clear. I focus mainly on the occurrences that refer to the Bubi and Fang languages, although in some cases, I also rely on instances that appear in relation to the Bubi and Fang people.

In what follows, I will examine the positive and negative topics that appear in relation to the two languages. The clearly positive opinions stand out as proof that Bubi and Fang are cultural elements of great importance for their speakers. The following four points can be highlighted.

Firstly, several Equatoguinean interviewees associate Bubi and Fang with personal and cultural values such as identity and pride, roots, origin, customs, local music, and food (see Example 1).

- (1) *[hablando de sus hijos hipotéticos:]* Ie.: [...] intentaré / si puedo económicamente // que todas las vacaciones se vayan a Guinea // porque creo que hay valores de Guinea que no tienen que perder [...] en casa les intentaré educar / a la manera de Guinea / entre medias [...] intentaré que hablen bubi [...] cuando consiga yo aprender [...] y que sepan de Guinea / que sepan sus costumbres su cultura // para que estén orgullosos de dónde son que in- intentar que no lo renieguen / eso: y: eso es lo que intentaré (03_-8HB)
[talking about his hypothetical children:] Ie.: [...] I would try / if I can afford it // that they go for every vacation to Guinea // because I believe that there are Guinean values that don't have to be lost [...] at home I would try to raise them / the Guinean way / in part [...] I would try to make them speak Bubi [...] when I manage to learn it [...] and that they know about Guinea and its customs its culture so that they can be proud of where they are from try that they don't deny it / that and that's what I would try (03_-8HB).

The two Bantu languages represent a connection with the speakers' country of origin and relatives who still live in Equatorial Guinea. The Bubi language is a

way for the speakers to move mentally and emotionally to Equatorial Guinea, as a form of therapy and to remember; it is associated with the Bubi culture, music, and traditional food. The Fang language is an important “mother tongue” and therefore something peculiar and important (e.g., “I am Fang and being Fang I consider that I have to learn and know Fang”, 23_+8MF). Without any doubt, both languages are integral parts of the multicultural tradition of Equatorial Guinea.

Secondly, some speakers cannot express certain things in a language other than the African languages (see Example 2). These languages are described by the speakers as something that is their own (*yours, theirs*), as opposed to Spanish, which is after all a language imposed during the time of colonization.

- (2) *Ie.: [...] [el bubi] tiene otros componentes que no tiene el español entonces eeh tiene: / .h o: o sea te sientes un poco mejor expresándote en bubi que en español porque .h: el bubi lo sientes como una lengua propia tuya [...] lo otro como una lengua: .h eeh extraña porque hay palabras muchas veces que quieres decir y: y no puedes pero en bubi sí: (08_+8HB)*

Ie.: [...] [Bubi] has other components that Spanish doesn't have so it has ooh that is you feel a bit better expressing yourself in Bubi than in Spanish because you feel Bubi as your own language [...] the other as a ooh strange language because there are words you want to say many times and and you can't but in Bubi you can (08_+8HB)

In fact, interviewees from both ethnic groups describe the communicative situations in which they are able to use their African languages in the Spanish migration context as joyful moments, since they feel comfortable speaking and listening to them.

Thirdly, both Bubi and Fang interviewees express their desire to transmit their African languages to their children, also in the context of migration, even though this is rather difficult, as I will analyze later.

Finally, in some interviews, I noted a certain awareness of language revitalization (e.g., rap in Bubi). Individual attempts to recover or even learn African languages are described, particularly in the context of migration. For instance, one young Bubi interviewee always tries to spend more time with her grandmother in order to practice Bubi, and others say that they speak and learn more Bubi or Fang in Spain than when they were living in Equatorial Guinea. It seems, though, that the awareness of the cultural value of these languages is more apparent, at least for some of the Equatoguinean migrants, when the distance to their homeland and own communities is greater.

At first sight, the negative opinions associated with Bubi and Fang seem to outweigh the positive ones. However, it soon becomes clear that many of these

comments in fact express a positive attitude toward these African languages, albeit indirectly. The following six topics can be summarized.

The first topic is mentioned by many of the interviewees: they regret the problem of the intergenerational transmission of Bubi and Fang, mainly in the Spanish migratory context, but to a lesser extent also in their country of origin. They point to this as the main challenge related to their African languages and express, almost unanimously, their desire to be able to transmit this linguistic legacy to their children. They find it painful and sad that young people no longer speak their own languages (some still understand them, but do not speak them); as such, several of the interviewees fear that the Bubi language in particular will be lost in the near future (not only in the Spanish diaspora, but maybe also in Equatorial Guinea itself). In the case of Fang, negative practical implications are also noted, since when returning to Equatorial Guinea, a lack of knowledge of Fang can cause communication problems or lead to discrimination. Others, however, think that Spanish is sufficient to get by in Equatorial Guinea. The causes that explain the problem of linguistic transmission in Spain are multiple: lack of contact with other people who speak the same language; mixed, multilingual couples; the different situation of women in Spain with regard to work (and, therefore, the relatively limited time they spend with their children); and the parents' own lack of linguistic proficiency. Especially in the case of Bubi, it seems that the grandparents' generation, and grandmothers in particular, are fundamental for the transmission of the language to the younger generations. This same fact is also implicitly present in a number of opinions that associate the Bubi language with the village, the rural environment, and the elderly or the aged.

Whereas this first topic – the problem of linguistic transmission to subsequent generations – indirectly expresses a very positive attitude toward the African languages, the second point is clearly negative. Two Fang interviewees question the usefulness of the Fang language (see Examples 3 and 4); they believe that speaking Spanish is enough to live in Equatorial Guinea, and that in Spain or in other countries, Fang does not offer any advantages to them or their children. Another person adds that, since Fang is not an official language in Equatorial Guinea, it has no value outside the country (17_-8MF).

- (3) *Ie.: no h: no me parece importante de mi cultura [...] me parece más importante otras cosas de cultura .h como enseñar al que no sabe // como enseña: pues las tradiciones no: un: no un dialecto [...] no me parece importante un dialecto*

Ir.: vale / .h si tuvieras hijos por ejemplo: ¿les enseñarías hablar fang también?

Ie.: ¿si tuviese hijos aquí: [en España]?

Ir.: sí aquí

Ie.: no / no les enseñaría hablar fang [...] porque el fang pff / ¿qué les va a aportar? [...] cero [...] si va a Guinea: / y tal no va a tener problema: .h / hablando español // lógicamente sí va a tener problema / que no hable fang // porque dirá alguno “pero si tu padre es fang” y ¿a mí qué? (19_+8HF)

Ie.: no I don't think it's important in my culture [...] I think it's more important other things of culture like teaching the one who doesn't know // like to teach well traditions not a not a dialect [...] I don't think that a dialect is important

Ir.: okay / if you had children for example would you teach them to speak Fang?

Ie.: if I had children here [in Spain]?

Ir.: yes here

Ie.: no / I wouldn't teach them to speak Fang [...] because Fang pff / what will it bring to him/her? [...] zero [...] if he/she goes to Guinea / and so he/she won't have a problem / speaking Spanish // of course he/she will have a problem / with not speaking Fang // because someone will say “but your father is Fang” and so what? (19_+8HF)

- (4) *Ie.: [...] no no no o sea [el fang] no me aportaría nada [...] solo se habla en Guinea en Camerún y en Gabón / y cuando estoy aquí: / no no me aporta nada el fang [...] porque: en la calle yo tengo que: defenderme / y si: hablo fang ¿quién me va a entender? ¿quién me va a ayudar? .h o sea que: nada*

Ir.: mhm mhm / .h // a pesar de ello ¿el fang para ti es algo: / importante?

Ie.: sí: claro / porque también forma parte de: / de mi vida [...] me identifica que yo soy guineana y soy fang (18_-8MF)

Ie.: [...] no no no I mean [Fang] would not bring me anything [...] it is only spoken in Guinea in Cameroon and in Gabon / and when I am here / Fang does not bring me anything [...] because I have to defend myself in the street / and if I speak Fang who will understand me? who is going to help me? so nothing

Ir.: mhm mhm // despite this Fang for you is something / important?

Ie.: yes of course / because it also forms part of / of my life [...] it identifies me that I am Guinean and I am Fang (18_-8MF)

A third point triggering negative comments that comes up in more than one interview has to do with the linguistic interferences between the African languages Bubi and Fang and other languages, especially Spanish. This is linked to a long-lasting tradition of negative language attitudes concerning the multilin-

gual repertoires of Spanish-speaking Equatoguineans. To cite just one example of an early publication about the “problem of polylingualism”:

To all the vernacular languages [...] we must add the very particular problem of bilingualism. This, the bilingualism, and, by analogy, polylingualism, is one of the most pressing and complex problems, and that until now has not been given the pedagogical importance it requires, not only because of its psychological effects, but also because of the political and social effects it entails. (Castillo Barril 1969: 52)¹⁰

Indeed, this opinion is quite widespread, both among certain linguists and among the multilingual speakers themselves, and it is well known from very different linguistic contexts. Morales de Walter (2008: 307–308), for example, describes the language attitudes of Puerto Ricans toward their daily code-switching between Spanish and English as follows (*italics added*):

Code-switching within or between sentences has received *much criticism*, especially from educators who see in this use *possible deficiencies in one or both languages* or, in any case, carelessness in speaking. The bilinguals themselves are hesitant in their assessment, although in Brentwood 82 percent of the adults alternated Spanish and English, 39 percent of them pointed out that this was *not correct*, and especially criticized the alternation within the same sentence. 77 percent of the students alternated codes and, like their parents, a good proportion of them believed that this was *not appropriate*. Users of code-switching think it is a *problem*, but they justify it. [...] In general, parents and children, although they use it, are quite critical when they have to give their opinion on language switching; they think that this is one of the causes that makes their way of speaking *stigmatized*.¹¹

¹⁰Own translation. The original quote is: “A todas las lenguas vernáculas [...] habrá que sumar el problema particularísimo del bilingüismo. Este, el bilingüismo, y, por analogía, el polilingüismo, es uno de los problemas más acuciantes y complejos y que hasta ahora no se le ha dado la importancia pedagógica que requiere, no sólo por sus efectos psicológicos, sino por los políticos y sociales que implican.”

¹¹Own translation. The original quote is: “La alternancia de códigos dentro de una oración o entre oraciones ha recibido muchas críticas, especialmente de los educadores que ven en ese uso posibles deficiencias en una o ambas lenguas o, en todo caso, descuido al hablar. Los propios bilingües dudan en su valoración, aunque en Brentwood un 82% de los adultos alternaba español e inglés, el 39% de ellos señalaba que eso no era correcto, y criticaban especialmente la alternancia dentro de la misma oración. El 77% de los estudiantes también alternaban códigos y, como sus padres, una buena proporción de ellos creía que esto no era adecuado. Los usuarios

In a similar way, in my interview corpus, it is stated that Bubi and Fang are being “corrupted” (i.e., mixed), that the Fang language “hinders” Equatoguinean Spanish, and that the mix of languages in discourse is something negative. Although some argue that frequent code switching is simply a natural expression of their multilingual identity, negative attitudes are more frequent.

A fourth negative language attitude is mentioned about the Bubi language. One Bubi interviewee thinks that the Bubi feel inferior when they speak Bubi (they have a “mental problem,” 09_+8HB). He argues that other people mock them when they speak Bubi, a situation that contrasts with what he observes in other Equatoguinean ethnic groups. Indeed, several people believe the Bubi are submissive and conformist, and that they have been discriminated against since Equatorial Guinea’s independence. This is explained by the fact that both presidents of independent Equatorial Guinea have been Fang, which is the reason for the clear dominance of Fang in politics and practically all areas of life, a phenomenon known as *Fanguization* (cf. Aixelà-Cabré 2013: 62–69). As some interviewees explain, the Fang are in the government, in the bureaucracy, in everything. They constitute the majority and dominant ethnic group, and they feel somehow superior. The Bubi, on the contrary, are a minority ethnic group with fewer opportunities, they have been largely excluded from power, and they have been persecuted and marginalized for decades. It is interesting to note that, according to Bolekia Boleká (2003: 93), this division between Bubi and Fang has its roots in colonial times. He explains that the Spanish governors and businessmen told the Bubi negative stories about the Fang (e.g., portraying the Fang as savage, brutal invaders) and vice versa (e.g., portraying the Bubi as weak, lazy, and inferior). During the process of independence, this ethnic conflict was exacerbated.

A fifth negative topic mentioned in some interviews is connected with the previous point. Several interviewees, both Bubi and Fang, confirm and criticize the fact that the Fang always speak Fang or speak it too much. As they explain, today, the Fang are not only the majority ethnic group in the country as a whole – and especially in mainland Equatorial Guinea (Río Muni), which is their original home – but also in the capital Malabo, on the island of Bioko, the traditional territory of the Bubi ethnic group. According to some of the interviewees, the Fang feel that they are already the owners of the capital, and there are some Bubi who want the Fang to return to the continent. It is because of this situation that one Fang participant describes the Fang as “invaders” and “braggarts,” and

del cambio de códigos piensan que es un problema, pero lo justifican. [...] En general, padres e hijos, a pesar de que lo usan, son bastante críticos cuando tienen que dar su opinión sobre la alternancia de lenguas; piensan que esa es una de las causas que hace que su modo de hablar esté estigmatizado.”

he adds – laughing – that in Equatorial Guinea, they are called “Japanese” because there are so many of them (20_+8HF). As a result of this Fang presence all over the country, they apparently always address people in Fang automatically, regardless of whether the other person speaks Fang or not; some even add that the Fang want to impose their language on other ethnic groups. Hence, there is a clear difference in power and status between Fang as the dominant language and the other autochthonous languages of Equatorial Guinea, which makes it more difficult to use the latter in public and formal domains.

This last point has another negative consequence: in the opinion of several interviewees, both Bubi and Fang, the Fang find it difficult to speak Spanish well, precisely because they speak Fang so often. They think that the Fang maintain their language more than the Bubi, that they speak it more often at home and everywhere (both in Equatorial Guinea and in Spain), and that, for this reason, the Fang are not as interested in speaking Spanish as the Bubi. Others like to qualify this idea and specify that linguistic proficiency depends on the education of each person and that in all the ethnic groups, there are some who speak Spanish better than others. Interestingly, the widespread idea that the Fang speak Spanish worse than the Bubi – a kind of shared mental space, or shared image of linguistic reality built on individual experiences and common beliefs – is based not primarily on linguistic facts, but above all on the ethnic, cultural, social, and political differences in Equatorial Guinea. The inferior position of the Bubi in Equatorial Guinea is compensated via the image of their linguistic superiority in Spanish. This shows the importance of taking the background of a linguistic community into account when describing and interpreting their language attitudes. In this case, it is necessary to recognize how the ethnic conflicts in the speakers’ home country are reflected in the language attitudes of Equatoguineans in the Spanish migration context toward both Spanish and the Bantu languages spoken by the two ethnic groups, Bubi and Fang (cf. Schlumpf 2020a).

In summary, it is important to bear the following points in mind regarding attitudes toward the African languages Bubi and Fang:

- Firstly, it is necessary to emphasize the great cultural value that these languages carry. They are a sign of ethnic identity, hence the frequent allusions to subjects such as culture, customs, and origin, and the idea that they represent a means for Equatoguineans abroad to connect with their home country. Altogether, a positive internal prestige is shown here, which allows identification with the group (*in-group language*).
- Secondly, and in a more negative sense, it is worth recalling the frequent critiques of the multilingual, mixed repertoires of the Equatoguineans,

who naturally switch between different languages in their spontaneous discourse with their compatriots, both in Madrid and in Equatorial Guinea itself.¹² This reflects a general lack of recognition of intense language-contact situations and the linguistic implications this has, especially in (post)colonial contexts.

- Thirdly, when comparing the two Bantu languages studied, clear differences in power are noted. Fang constitutes the dominant language in Equatorial Guinea, because the Fang are the majority group and due to the political circumstances in Equatorial Guinea since its independence. Bubi, on the contrary, is a dominated language. Within the Bubi community, this situation has produced a certain feeling of inferiority on the one hand, and a shared idea of linguistic superiority in Spanish on the other.
- Finally, in addition to this internal difference between Bubi and Fang concerning status and prestige, I have highlighted that these African languages are in constant opposition to other languages. Spanish is the first official language and the lingua franca of Equatorial Guinea, as well as the dominant language in the migration context in Madrid. The Bantu languages of Equatorial Guinea, in spite of their cultural and linguistic relevance, especially in informal contexts, do not enjoy any form of official status in the country; in some instances, this fact is openly criticized by the Equatoguineans interviewed, but in other cases, it is used to demonstrate the perceived uselessness of these languages, at least outside Equatorial Guinea.

3.3 The use of linguistic terminology as a reflection of colonial power structures: *Lengua* versus *dialecto*

Some of the aspects described in the previous section can be corroborated by how the Equatoguinean interviewees denominate the different languages spoken in their country of origin. In particular, the use of the terms *lengua* and *dialecto* to refer to different languages is highly revealing for the purpose of this topic and exemplifies a widespread phenomenon observed in various contact situations with a hierarchical structure (e.g., in Latin America).¹³

¹²More data on the maintenance of the African languages in the migration context can be found in Schlumpf (2020a).

¹³For some further details on this topic and examples of interview passages, see Schlumpf (2021a).

The term *lengua*¹⁴ is used to refer to both the African languages and Spanish. However, the most frequent semantic and morphological combinations are different. On the one hand, the references to the African languages (in this case, Bubi and Fang) highlight, above all, their cultural and identity values (e.g., *lengua materna* ‘mother tongue,’ *lengua local* ‘local language,’ *lengua vernácula* ‘vernacular language,’ *mi/tu lengua* ‘my/your language,’ *nuestras lenguas* ‘our languages,’ *sus lenguas propias* ‘their own languages’). On the other hand, the largest number of occurrences that refer to Spanish show the combination *lengua oficial* ‘official language’ (16 of 29 occurrences). Furthermore, on a few occasions, Spanish in Equatorial Guinea is described as a *lengua extranjera* or *lengua extraña* ‘foreign/strange language’ and as *una lengua impuesta por los colonizadores* ‘a language imposed by the colonizers.’ This clearly shows the main association of Spanish with its *official* status in Equatorial Guinea, while the Bantu languages Bubi and Fang represent a symbol of the *local* traditions, whose main communicative settings are Equatoguinean households. As such, Spanish enjoys overt prestige, whereas the two African languages represent important in-group languages and enjoy a kind of covert prestige, according to Labov’s (2006) and Trudgill’s (1972) terminology. This difference reproduces a clearly colonialist view on the sociocultural realities in Equatorial Guinea, a view that over history and in many colonized countries has led to the imposition of foreign (European) languages to the detriment of local languages. The examples of my interviews demonstrate that this colonial and Eurocentric discourse continues to circulate today, and has even been converted into a properly African discourse.

If we look at the occurrences of the term *dialecto*¹⁵ that appear in the interview corpus, most of them refer to one of the Bantu languages of Equatorial Guinea. Here, the syntactic combinations with possessives stand out (e.g., *mi/su/nuestro dialecto* ‘my/his or her/our dialect’), as does the expression *dialecto de casa* ‘home dialect.’ Spanish and other European languages are never labelled as dialects, except when referring explicitly to a certain diatopic variety. This is highly revealing when considering that “the name ‘dialect’ has often served to mask a situation of linguistic subordination and reinforce power relations, not only between

¹⁴The term *lengua* is used a total of 131 times, of which 55 refer to local Bantu languages and 29 to Spanish. The remaining occurrences refer to other languages or to *lengua* in general.

¹⁵The term *dialecto* is used a total of 33 times and shows a very clear semantic distribution, since 26 of the 33 cases refer to an African language. The remaining cases are not very clear; they refer to *dialecto* in general or to different dialects of a specific language (e.g., dialects of Spanish).

the linguistic varieties, but also between groups of speakers and language communities” (Moustaoui Srhir 2016: 56).¹⁶

The passages in the interviews in which a clear and direct terminological opposition is established between *lengua* and *dialecto* are of particular interest. All these examples show and, at the same time, reproduce linguistic hierarchies between European *languages* and African *dialects*. Even today, in most African countries, European languages occupy the position of official languages (the exceptions are found in the north of the continent, cf. Barbosa da Silva 2011: 65) and represent “the languages of public and intellectual discourse” (Zezeza 2006: 21). On the other hand, the colonial period saw autochthonous languages relegated to non-official and informal contexts, and therefore to certain speech registers – a situation that largely continues to this day.

This terminological and conceptual problem is another reminder of the objectives of decolonial theories. In fact, Lara Delgado (2015) points out the need to eliminate the “epistemic racism which refers to a hierarchy of colonial domination where the knowledge produced by Western (imperial and oppressed) subjects within the zone of being is considered a priori as superior to the knowledge produced by the non-Western colonial subjects in the zone of non-being.”¹⁷ The dichotomy between *languages* and *dialects* is a good example of these hierarchies of colonial domination that continue to be transmitted and consolidated today. In this context, “the theories of decoloniality go beyond decolonization, and propose ‘other’ alternatives that seek to subvert the hegemonic power, in order to make visible the effects that colonialization and coloniality have brought in power, in knowledge and in being” (Gómez Vélez et al. 2017: 51).¹⁸ One very important source of these hierarchies constructed in colonial times is the predominant, Eurocentric perspective on history, culture, and society. I understand this concept as “the cognitive perspective produced in the long time of the whole Eurocentered world of colonial/modern capitalism, and which *naturalizes* the experience of the people in this pattern of power. That is, it makes them perceive

¹⁶The same author explains the following: “attitudes towards ‘dialects’ considered socially as non-standard, are attitudes that reflect the structure of the society where they are spoken and the power relations that exist between the various groups that compose that society – its speakers and their social classes” (Moustaoui Srhir 2016: 51).

¹⁷Own translation. The original quote is: “Racismo epistémico que se refiere a una jerarquía de dominación colonial donde los conocimientos producidos por los sujetos occidentales (imperiales y oprimidos) dentro de la zona del ser es considerada a priori como superior a los conocimientos producidos por los sujetos coloniales no-occidentales en la zona del no-ser.”

¹⁸Own translation. The original quote is: “Así las cosas, las teorías de la decolonialidad van más allá de la descolonización, y plantean alternativas ‘otras’ que buscan subvertir el poder hegemónico, para visibilizar los efectos que ha traído la colonización y la colonialidad en el poder, en el saber y en el ser.”

it as *natural*, consequently, as given, not susceptible to be questioned” (Quijano 2007: 94, emphasis in original).¹⁹ This Eurocentric system of norms is thus perceived as “a superior and universal pattern of reference,” and consequently, “[t]he other forms of being, the other forms of organization of society, the other forms of knowledge are transformed not only into different, but into lacking, archaic, primitive, traditional, premodern” (Lander 2000: 10).²⁰ In order to decolonialize our knowledge, it is important to recognize the problematic consequences of this Eurocentric vision and to invert established power structures via the critical analysis of traditional discourses, concepts, and use of terminology. Upgrading the Bantu languages of Equatorial Guinea to the category of *language* (as opposed to *dialect*) would be an important step toward the positive recognition of these linguistic codes – both within the scientific community and among the speakers themselves.

3.4 Equatoguineans’ attitudes toward Pichi

Having analyzed the speakers’ attitudes toward Bubi and Fang (Section 3.2) and their usage of the terms *lengua* and *dialecto* (Section 3.3), I will now turn to the Equatoguineans’ language attitudes toward Pichi, also known as Pichinglis(h) or Pidgin English.²¹ This local variety of English is mainly spoken in the capital Malabo and its surroundings, yet it can also be heard in other areas of Bioko and, increasingly, in the city of Bata. According to Yakpo (2013b: 276–278), 70 percent of Bioko’s population use Pichi regularly. One of the main functions of Pichi is that of serving as the lingua franca between speakers of different ethnic groups. However, it is also used in informal conversations between speakers who share the same ethnic background, especially Bubi. Moreover, Pichi allows communication with Africans from other countries, where similar varieties of English can be found (e.g., Cameroon, Ghana, Nigeria, and Sierra Leona). Yet despite all this,

¹⁹Own translation. The original quote is: “Se trata de la perspectiva cognitiva producida en el largo tiempo del conjunto del mundo eurocentrado del capitalismo colonial/moderno, y que *naturaliza* la experiencia de las gentes en este patrón de poder. Esto es, la hace percibir como *natural*, en consecuencia, como dada, no susceptible de ser cuestionada.”

²⁰Own translation. The original quote is: “Esta es una construcción eurocéntrica, que piensa y organiza a la totalidad del tiempo y del espacio, a toda la humanidad, a partir de su propia experiencia, colocando su especificidad histórico-cultural como patrón de referencia superior y universal. [...] Las otras formas de ser, las otras formas de organización de la sociedad, las otras formas del saber, son transformadas no sólo en diferentes, sino en carentes, en arcaicas, primitivas, tradicionales, premodernas.”

²¹For more information about Pichi and its linguistic features, history, and functions in Equatorial Guinea, see Lipski (1992) and Yakpo (2010; 2013a; 2013b; 2016a; 2016b; 2018; 2019).

Pichi enjoys neither prestige nor official recognition in Equatorial Guinea (Yakpo 2016b: 25–26).

In my interview corpus, a total of 164 references to Pichi were found: *pichi* (122 occurrences), *inglés* (28 occurrences), *pichinglish/pichinglis* (5 and 6 occurrences respectively) and *broken English* (3 occurrences). As far as general statements are concerned, and although opinions also differ, most interviewees say that it is mainly the Bubi who speak Pichi, especially people from the island (Bioko) and people from Malabo. The main contexts in which Pichi is used are the street and conversations with friends, among young people. Pichi is also spoken in some Equatoguinean households, although in other families it is never used. By contrast, Pichi is not represented at all in the school environment – where Spanish is the only language used – or in other (semi-)official contexts.

Looking at the positive and negative evaluations that appear in relation to Pichi, there is a large imbalance between approximately 60 occurrences of negative content compared to fewer than 30 occurrences of content presented as something positive (cf. Yakpo 2016a: 218ff; 2016b). In summary, Pichi is seen as something positive in terms of its communicative utility and its particular value as a dynamic and ever-changing code, used in particular by young people in Malabo/on Bioko. Some young Bubi participants even describe it as a “secret code” of their generation, although elder generations also use a more traditional version of Pichi. Some interviewees point out its basic value as an additional language next to Spanish and the Bantu languages spoken in the country; according to another person, knowledge of Pichi makes it easier to learn English at school (06_-8MB).

On the other hand, the negative opinion predominates that Pichi is somewhat incorrect, impure, and non-standard (cf. Yakpo 2016b: 35–36). It is bad English and therefore an incomplete or “broken” code of a complete language: a kind of English, but not English; something like English, but not the “real” one (cf. Woolard & Schieffelin 1994: 63); a cheap replica of English; a corrupted version of English; a mixed code (*dialect*) with influences of other (local) languages (see Example 5). Similar linguistic terminology and descriptions are documented in other studies on Pichi, including “broken English” (Bolekia Boleká 2001: 36), “corrupted English” (Castillo Rodríguez 2013: 356), “broken-inglis” (Lipski 1992; 2004: 117), “inglés-africano,” “broken English,” and “inglés roto” (De Zarco 1937: 5). According to two of my interviewees, knowledge of Pichi can make the acquisition of standard English at school difficult (see Example 6), which contrasts with the interviewee who said that it helps with English acquisition. In addition, Pichi is sometimes associated with delinquency and bandits.

- (5) *Ie.: [...] [el pichinglish] es una réplica del: bueno una réplica por decirlo así del inglés pero muy barato o sea: es .h: sí se ha modificado totalmente o sea: no es inglés ni es nada es una // a lo mejor algo que suena inglés pero: / eh la: la gran mayoría de palabra:s se han inventao: se: se han modifcao (05_-8MB)*
Ie.: [...] [Pichinglish] is a replica of well a replica so to speak of English but very cheap I mean it is yes it has been totally modified I mean it is not English nor is it anything it is a // maybe something that sounds English but / eh the the great majority of words has been invented has has been modified (05_-8MB)
- (6) *Ie.: no yo no quise aprender pichi porque yo: el pichi me parece algo cutre [...] porque es el inglés mal hablado ¿sabes? .h una persona que ya ha aprendido pichi y que / habla pichi / es / difícil que sepa hablar el inglés (16_-8MF)*
Ie.: no I didn't want to learn Pichi because I Pichi seems to me somewhat shabby [...] because it is badly spoken English you know? a person who has already learnt Pichi and who / speaks Pichi / it's / difficult that he/she knows how to speak English (16_-8MF)

Clearly, the fact that most negatively influences the evaluations of Pichi – despite its common use in Equatorial Guinea as a kind of second lingua franca – is its close cognation with English (cf. Yakpo 2016b: 218–220). This is the reason why it is almost always evaluated *in comparison to* English, rather than as an independent linguistic code. All this despite the fact that standard English is not even currently spoken in Equatorial Guinea (except in the oil camps). This can be compared with typical diglossic situations according to the original definition of diglossia by Charles A. Ferguson (1959): the coexistence of a clearly superposed and prestigious variety of a language (*high variety*) and another variety of the same language, more regional, usually a more informal and oral code (*low variety*). Although standard English is not part of the linguistic setting of Equatorial Guinea, its existence and prestige (and the Equatoguineans' awareness of it) are enough to consolidate the *low variety* status of Pichi. This direct comparison to “real” English leads to an even more negative interpretation of Pichi than, for example, the Bantu languages Bubi and Fang.²² Although these two languages

²² A similar comparison can be found between different languages and varieties in the Maghreb. Similar to Pichi in Equatorial Guinea, the local Arabic variety known as Darija is characterized by its long-standing inferiority vis-à-vis Modern Standard Arabic and has only recently gained some recognition and sociolinguistic prestige among its speakers (see Moustauoui Srhir 2019; Moustauoui Srhir & Moscoso García 2019). On the other hand, Tamazight, although only officially recognized in Morocco in 2011, has always been a sign of the cultural and linguistic identity of the Berber community vis-à-vis the majority Arab population. In the two Spanish exclaves in Northern Morocco, it can also be observed that Tamazight has a stronger integrat-

also seem to be somehow evaluated as less prestigious than Spanish, they symbolize the local traditions of the different ethnic groups. In summary, they assume important cultural values for the speakers and are part of the speakers' ethnic identity, which is not the case for Pichi (cf. Yakpo 2016a: 26ff.).

4 Conclusion

In this final section, I summarize some major points in this chapter and highlight a number of specific problems relevant to the topic. These problems can be used to understand, analyze, and question traditional power relations between languages and linguistic varieties in colonial and postcolonial settings.

The description of the language attitudes of the Equatoguineans in Madrid showed that Bubi and Fang enjoy a high level of prestige as part of the cultural and ethnic identity of the speakers. They both fulfill particular communicative functions, in interactions with elder generations or when talking about specific topics, for instance. Pichi, on the other hand, has an important function as Equatorial Guinea's second lingua franca and as an in-group code for younger generations. It also enables speakers to communicate with Africans from other countries. However, both the two Bantu languages and Pichi find themselves in an uncomfortable situation of comparison and inferiority vis-à-vis standardized European languages – namely, the official Romance languages (especially Spanish, the language of the former colonial power) and standard English. Bubi and Fang, in the opinion of most of the interviewees, are somehow less valuable for official purposes or as languages used in the educational system than Spanish. As for Pichi, it is considered to be less prestigious than English and is frequently described as “poor,” “corrupted,” or “broken” English. These power relations between European and African languages and varieties are reflected in the terminological opposition between *lengua* and *dialecto*, a legacy of colonial domination and hierarchizations. Whereas the former is used with very different associations when relating to Spanish and to Bubi and Fang (the official status of Spanish versus the local values of the two African languages), the latter almost exclusively refers to the Bantu languages. All this has serious implications for the speakers themselves. Negative evaluations of their African mother tongues and the commonly used Pichi turn into a negative self-perception and a critical opinion toward an integral part of their own cultures and communities.

ing function in Melilla than Darija has in Ceuta, because the latter is seen as a “deformed” and “impure” form of Classical Arabic (Tilmatine 2011: 39). Similar to Tamazight in the Maghreb, the Bantu languages Bubi and Fang in Equatorial Guinea represent important cultural symbols for their ethnic groups, although they do not enjoy any official status in the country.

When it comes to differences between the two ethnic groups studied, Bubi and Fang, I have shown that feelings of inferiority in the case of the former have been propagated since Equatorial Guinea's independence in 1968 under both Fang regimes (Francisco Macías Nguema Bidyogo, 1968 to 1979, and Teodoro Obiang Nguema Mbasogo, since 1979), a process known as *Fanguization*. Until today, most influential positions in the country are held by Fang.

Considering the specific context of my study – the migration context in Madrid – and the daily experiences of my interviewees as Equatoguineans in this Spanish setting, it can be observed that the negative evaluations (and self-evaluations) are fostered by external forces such as social exclusion, limited access to the labor market, and even racism. The widespread lack of knowledge about Equatorial Guinea and the languages spoken in the country, as well as the non-recognition of the Equatoguinean Spanish as part of the Hispanophone world, do not improve the attitudes of the Equatoguineans vis-à-vis their own linguistic repertoires; nor do so theories that favor monolingual speakers of standard (European) languages. Experiences of discrimination further foster feelings of sociocultural inferiority as part of a minority.

To conclude, I wish to emphasize from a theoretical point of view that the negative attitudes of the Equatoguineans interviewed toward their own local languages help to perpetuate not only colonial hierarchies and power relations, but also the transmission of conservative ideas and language ideologies. Bubi and Fang receive mostly negative evaluations, unless they are related exclusively to culture and local traditions, and they are always compared to (standard) European languages, which serve as superior counterpoints. In this way, Western standards and Eurocentric views on cultural and linguistic phenomena, established and spread during the colonial period, ultimately become African views as well. It will be crucial to compare the findings of the present study with the data collected in Equatorial Guinea itself in 2022. This will allow to better understand the influence of the migration context in the Madrid corpus; and to get more profound insights into the contemporary language attitudes and ideologies of Equatoguineans who live in their own country.

It is of utmost importance that work is done to effect a change of perspective in investigations of Africa-related topics or other colonized territories. African phenomena must be evaluated from an inner-African perspective, avoiding constant references to European models. It is time to leave the colonial past behind and decolonize perceptions of historically constructed dichotomies and hierarchies. This will enable researchers to focus positively on the future; after all, the recognition and consolidation of local structures, cultures, and knowledge are crucial to solving not only social problems, but political issues as well.

Acknowledgements

The support of several funding institutions and individuals was crucial for the realization of my research project on the Equatoguinean community in Madrid, and I wish to express my gratitude to all of them. The interviews with Equatoguineans in Madrid were carried out during two research stays funded by the Swiss National Science Foundation (SNSF) (project no. 173468: “Sociolinguistic integration of immigrants from Equatorial Guinea in Madrid”) and via a research scholarship (“Giner de los Ríos”) from the University of Alcalá. The transcriptions of the interviews were realized with the help of Lilli Geyer-Schuch, Sara Carreira, Laura Renna, Tabea Dürr, Emanuel Branco (University of Basel), and Adriana Orjuela López (University of Freiburg). The digitization of the interviews was carried out in collaboration with Anđelka Zečević (Faculty of Mathematics, University of Belgrade) within the project “Digital analyses of sociolinguistic data. Linguistic features in interviews with Spanish speakers from Equatorial Guinea” (project no. 190022) financed by the SNSF. The databases analyzed in this chapter were annotated together with Sara Carreira and Lilli Geyer-Schuch (University of Basel). Finally, I would like to thank Johannes Ritter (University of Basel) for his valuable support during the preparation and careful revision of this chapter; and the two anonymous reviewers for their helpful suggestions.

References

- Aixelà-Cabré, Yolanda. 2013. Of colonists, migrants and national identity. The historic difficulties of the socio-political construction of Equatorial Guinea. *Nordic Journal of African Studies* 22(1&2). 49–71.
- Aixelà-Cabré, Yolanda. 2020. Colonial Spain in Africa: Building a shared history from memories of the Spanish Protectorate and Spanish Guinea. *Culture & History Digital Journal* 9(2). e017.
- Barbosa da Silva, Diego. 2011. Política lingüística en África: Del pasado colonial al futuro global. *Estudios de Asia y África* 46(1). 65–95.
- Blommaert, Jan. 2006. Language ideology. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., 510–522. Amsterdam: Elsevier.
- Bolekia Boleká, Justo. 2001. *Lenguas y poder en África*. Madrid: Mundo Negro.
- Bolekia Boleká, Justo. 2003. *Aproximación a la historia de Guinea Ecuatorial*. Salamanca: Amarú.
- Bouchard, Marie-Eve. 2022. Postcolonial sociolinguistics: Investigating attitudes, ideologies and power in language contact settings. *Journal of Postcolonial Linguistics* 6. 1–11.

- Castillo Barril, Manuel. 1969. La influencia de las lenguas nativas en el español de La Guinea Ecuatorial. *Archivo de Estudios Africanos* 20. 45–71.
- Castillo Rodríguez, Susana. 2013. Language and the Hispanization of Equatorial Guinea. In José Del Valle (ed.), *A political history of Spanish. The making of a language*, 350–363. New York: Cambridge University Press.
- Chirilă, Elena Magdalena. 2015. *Identidad lingüística en Guinea Ecuatorial: Diglosia y actitudes lingüísticas ante el español*. Universidad de Bergen. (Master Thesis).
- De Zarco, Mariano. 1937. *Dialecto inglés - africano o Broken English de la Colonia Española del Golfo de Guinea. Epítome de la gramática seguido del vocabulario español-inglés e inglés-español*. 2nd edn. Turnhout: H. Proost / Cía.
- Del Valle, José. 2007. Glotopolítica, ideología y discurso: Categorías para el estudio del estatus simbólico del español. In José Del Valle (ed.), *La lengua, ¿patria común? Ideas e ideologías del español*, 13–29. Madrid / Frankfurt am Main: Iberoamericana / Vervuert.
- Ferguson, Charles A. 1959. Diglossia. *Word* 15(2). 325–340.
- Gallois, Cindy, Bernadette Watson & Madeleine Brabant. 2007. Attitudes to language and communication. In Marlis Hellinger & Anne Pauwels (eds.), *Handbook of language and communication: Diversity and change*, 595–618. Berlin / New York: De Gruyter Mouton.
- Gómez Vélez, Martha Isabel, Dora Cecilia Saldarriaga Grisales, María Claudia López Gil & Lina María Zapata Botero. 2017. Estudios decoloniales y poscoloniales. Posturas acerca de la modernidad/colonialidad y el eurocentrismo. *Ratio Juris* 12(24). 27–60.
- Guthrie, Malcolm. 1953. *The Bantu languages of Western Equatorial Africa*. London / New York / Toronto: Oxford University Press.
- Guthrie, Malcolm. 1971. The Western Bantu languages. In Thomas A. Sebeok (ed.), *Linguistics in Sub-Saharan Africa*, 357–366. Den Haag / Paris: Mouton.
- Kilgariff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý & Vít Suchomel. 2014. The Sketch Engine: Ten years on. *Lexicography* 1. 7–36.
- Labov, William. 2006. *The social stratification of English in New York City*. 2nd edn. Cambridge / New York: Cambridge University Press.
- Lander, Edgardo. 2000. Ciencias sociales: Saberes coloniales y eurocéntrico. In Consejo Latinoamericano de Ciencias Sociales (ed.), *La colonialidad del saber: Eurocentrismo y ciencias sociales. Perspectivas latinoamericanas*, 4–23. Buenos Aires: CLACSO, Consejo Latinoamericano de Ciencias Sociales.
- Lara Delgado, Jonhny. 2015. Pensamiento decolonial como instrumento transgresor de la globalización. *Analéctica* 1(10). n.p.

- Lipski, John M. 1992. Pidgin English usage in Equatorial Guinea (Fernando Poo). *English World-Wide. A Journal of Varieties of English* 13(1). 33–57.
- Lipski, John M. 2004. The Spanish language of Equatorial Guinea. *Arizona Journal of Hispanic Cultural Studies* 8. 115–130.
- Lipski, John M. 2012. Isn't Pidgin English just bad English? In E. M. Rickerson & Barry Hilton (eds.), *The five-minute linguist. Bite-sized essays on language and languages*, 2nd edn., 47–50. Sheffield: Equinox Publishing.
- Martín Rojo, Luisa. 2017. Language and power. In Ofelia García, Nelson Flores & Massimiliano Spotti (eds.), *The Oxford handbook of language and society*, 77–102. New York: Oxford University Press.
- Mbomío Bacheng, Joaquín. 2011. La originalidad de la literatura guineana. *Ibero-romania* 73–74. 126–143.
- Morales de Walter, Amparo. 2008. Puertorriqueños. In *Enciclopedia del español en los Estados Unidos: Anuario del Instituto Cervantes 2008*, 284–310. Madrid: Santillana / Instituto Cervantes.
- Moreno Fernández, Francisco. 2015. *Principios de sociolingüística y sociología del lenguaje*. 4th edn. Barcelona: Ariel.
- Moreno Fernández, Francisco (ed.). 2021. *Marcas y etiquetas mínimas obligatorias para materiales de PRESEEA* (Documentos PRESEEA de investigación). Alcalá de Henares: PRESEEA / Universidad de Alcalá. https://preseea.uah.es/sites/default/files/2022-02/Marcas%20y%20etiquetas%20m%C3%ADnimas%20obligatorias%20para%20materiales%20de%20PRESEEA_Moreno%20Fern%C3%A1ndez%20282021%29.pdf.
- Moustaoui Srhir, Adil. 2016. The concept of 'dialect' and 'language': A critical sociolinguistic overview. In Francisco Moscoso García & Adil Moustaoui Sghir (eds.), *VI Congreso de árabe Marroquí: Identidad y Conciencia Lingüística*, 47–60. Madrid: Ediciones Universidad Autónoma de Madrid.
- Moustaoui Srhir, Adil. 2019. Transforming the urban public space: Linguistic landscape and new linguistic practices in Moroccan Arabic. *Linguistic Landscape* 5(1). 80–102.
- Moustaoui Srhir, Adil & Francisco Moscoso García. 2019. Árabe marroquí: Dinámicas de cambios y nuevos horizontes. *Al-Andalus Magreb* 26. 1–5.
- Quijano, Aníbal. 2007. Colonialidad del poder y clasificación social. In Santiago Castro-Gómez & Ramón Grosfoguel (eds.), *El giro decolonial: Reflexiones para una diversidad epistémica más allá del capitalismo global*, 93–126. Bogotá: Siglo del Hombre Editores / Universidad Central, Instituto de Estudios Sociales Contemporáneos / Pontificia Universidad Javeriana, Instituto Pensar.
- Quilis, Antonio. 1983. Actitud de los ecuatoguineanos ante la lengua española. *Lingüística Española Actual* 5. 269–275.

- Quilis, Antonio. 1988. Nuevos datos sobre la actitud de los ecuatoguineanos ante la lengua española. *Nueva Revista de Filología Hispánica* 36. 719–731.
- Quilis, Antonio & Celia Casado-Fresnillo. 1993. La lengua española y los guineoequatorianos. Historia de unas actitudes. *Boletín de la Real Academia Española* 73. 569–587.
- Repinecz, Martin. 2019. Raza or race? Remembering slavery in Equatorial Guinean literature. *Hispanic Studies Review* 4(1). 121–135.
- Riochí Siafá, Juan. 2021. Guinea Ecuatorial, punto neurálgico de la literatura hispanoaficana. *Revista Aleph* 198. 45–58.
- Rosa, Jonathan & Christa Burdick. 2017. Language ideologies. In Ofelia García, Nelson Flores & Massimiliano Spotti (eds.), *The Oxford handbook of language and society*, 103–123. New York: Oxford University Press.
- Schlumpf, Sandra. 2018. Guineoequatorianos en Madrid: Actitudes hacia su propio español y el español madrileño. *Lengua y migración / Language and Migration* 10(2). 7–31.
- Schlumpf, Sandra. 2019. Construcción del colectivo guineoequatoriano en España a través de la prensa: *El País* y *La Vanguardia* (2010–2018). *Discurso & Sociedad* 13(2). 287–324.
- Schlumpf, Sandra. 2020a. African languages and Spanish among Equatoguineans in Madrid. *Spanish in Context* 17(1). 108–130.
- Schlumpf, Sandra. 2020b. El español hablado por los bubis y los fang de Guinea Ecuatorial. Valoraciones desde la comunidad guineoequatoriana en Madrid. *Lengua y migración / Language and Migration* 12(1). 103–125.
- Schlumpf, Sandra. 2021a. De *castizar* a *disparar* o ¿cómo hablan los guineoequatorianos el español? Terminología metalingüística y reflexiones teórico-ideológicas a partir de un corpus de entrevistas sociolingüísticas. *RILEX. Revista sobre investigaciones léxicas* 4(3). 11–51.
- Schlumpf, Sandra. 2021b. Spanisch in Afrika/Afrika in Spanien. Sprachliche Charakteristika von Spanischsprecherinnen und Spanischsprechern aus Äquatorialguinea in Madrid. *Romanistisches Jahrbuch* 72(1). 339–387.
- Shepherd, Nick. 2020. The grammar of decoloniality. In Ana Deumert, Anne Storch & Nick Shepherd (eds.), *Colonial and decolonial linguistics: Knowledge and epistemes*, 303–324. Oxford: Oxford University Press.
- Tilmatine, Mohand. 2011. El contacto español-bereber: La lengua de los informativos en Melilla. *Revista Internacional de Lingüística Iberoamericana* 9(2 (18)). 15–46.
- Trudgill, Peter. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1(2). 179–195.

- Trujillo, José Ramón. 2012. Historia y crítica de la literatura hispanoaficana. In Mbaré Ngom & Gloria Nistal (eds.), *Nueva Antología de la Literatura de Guinea Ecuatorial*, 855–907. Madrid: SIAL / Casa de África.
- Veronelli, Gabriela A. 2015. The coloniality of language: Race, expressivity, power, and the darker side of modernity. *Wagadu: A Journal of Transnational Women's & Gender Studies* 13. 108–134.
- Veronelli, Gabriela A. 2016. Sobre la colonialidad del lenguaje. *Universitas Humanística* 81. 33–58.
- Veronelli, Gabriela A. 2019. La colonialidad del lenguaje y el monolenguajear como práctica lingüística de racialización. *Polifonia* 26(44). 146–159.
- Wolff, H. Ekkehard. 2016. *Language and development in Africa: Perceptions, ideologies and challenges*. New York: Cambridge University Press.
- Woolard, Kathryn A. & Bambi B. Schieffelin. 1994. Language ideology. *Annual Review of Anthropology* 23. 55–82.
- Yakpo, Kofi. 2010. *Gramática del Pichi*. Vic (Barcelona): CEIBA / Centros Culturales Españoles de Guinea Ecuatorial.
- Yakpo, Kofi. 2013a. Pichi. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages*, vol. 1, 194–205. Oxford: Oxford University Press.
- Yakpo, Kofi. 2013b. Wayward daughter: Language contact in the emergence of Pichi (Equatorial Guinea). *Journal of African Languages and Linguistics* 34(2). 275–299.
- Yakpo, Kofi. 2016a. “The only language we speak really well”: The English creoles of Equatorial Guinea and West Africa at the intersection of language ideologies and language policies. *International Journal of the Sociology of Language* 2016(239). 211–233.
- Yakpo, Kofi. 2016b. O estatuto do pichi na Guiné Equatorial. *Revista do Instituto Internacional da Língua Portuguesa - PLATÔ* 3(6). 20–41.
- Yakpo, Kofi. 2018. ¿El nacimiento de una lengua afrohispana?: la influencia del español en el criollo inglés de Guinea Ecuatorial. In Dorothy Odattey-Wellington (ed.), *Trans-afrohispanismos. Puentes culturales críticos entre África, Latinoamérica y España*, 243–259. Boston / Leiden: Brill Rodopi.
- Yakpo, Kofi. 2019. *A grammar of Pichi*. Berlin: Language Science Press.
- Zezeza, Paul Tiyaambe. 2006. The inventions of African identities and languages: The discursive and developmental implications. In Olaoba F. Arasanyin & Michael A. Pemberton (eds.), *Selected Proceedings of the 36th Annual Conference on African Linguistics*, 14–26. Somerville, MA: Cascadia Proceedings Project.

Part IV

**Focus on microlinguistic
outcomes of language contact**

Chapter 9

Systems of pragmatic markers in contact: Processes and outcomes

Inga Hennecke

University of Tübingen

Pragmatic markers are highly polyfunctional and polysemic lexical units that generally occur in sentence-peripheral positions and do not contribute to the propositional content of an utterance. In situations of language contact, pragmatic markers are particularly susceptible to borrowing and other cross-linguistic influences because of their syntactic and semantic detachability. This paper presents a corpus-based analysis of the influence of language contact with English on the system of pragmatic markers in spoken Manitoban French, a variety of Canadian French spoken in Manitoba. To this aim, three sets of partially equivalent pragmatic markers were chosen for analysis: *comme* and *like*; *alors*, *donc*, and *so*; and *bon*, *ben*, and *well*. The analysis shows vastly different outcomes of long-term language contact on specific markers in one system. Four outcomes are discussed in this paper; namely, the emergence of new discourse-pragmatic functions, the borrowing of a marker from the other language, changes in frequency and productivity of specific markers, and the absence of specific markers in the system.

1 Introduction

Pragmatic markers remain a controversial topic in scientific discussion and there is still no consensus on their exact classification, delimitation, and definition (for a detailed overview, see Mosegaard Hansen 1998, Andersen 2001, Aijmer 2002, Aijmer & Simon-Vandenberg 2006). Pragmatic markers, which are often also referred to as discourse markers or discourse particles, are highly polyfunctional lexical units that may contain a high number of polysemic semantic meaning patterns. They also demonstrate syntactic flexibility and often occur



in sentence-peripheral positions. Pragmatic markers generally fulfill discourse-pragmatic functions and do not contribute to the propositional content of the utterance. In that sense, they do not assume a grammatical relation to the other elements of the utterance (see Hennecke 2014 for a detailed discussion). Further, some pragmatic markers, such as the English *like*, may also fulfill several types of hedging functions, particularly approximation and attenuation (see Kaltenböck et al. 2010).

The focus of the scientific debate on pragmatic markers has mostly centered around the characteristics and functions of pragmatic markers among monolingual native speakers. However, a number of studies focus on cross-linguistic issues of pragmatic markers, such as the role of pragmatic markers in situations of language contact and the implications of language contact on pragmatic markers (e.g., Mougeon & Beniak 1991; Maschler 2000; Matras 2000; Hlavac 2006; Torres & Potowski 2008). This paper aims to investigate three pairs of English and French pragmatic markers in Manitoban French, a variety of Canadian French that is spoken in the Canadian province of Manitoba. Manitoban French has been in an intensive, long-term language-contact situation with English and is therefore an ideal source for the investigation of the evolution of pragmatic markers in language contact (for more detailed information on the language contact situation in Manitoba, see Hennecke 2014).

2 Pragmatic markers in contact

Pragmatic markers are considered very susceptible to borrowing and other cross-linguistic influences because of their syntactic and semantic detachability (see Section 3.2). Still, this is not the only reason researchers are interested in bilingual pragmatic markers. It is widely considered that pragmatic markers tend to be difficult to translate from one language to another and have more than one translation equivalent. Furthermore, determining their semantic value and pragmatic functions is often a challenge, and it is not even unequivocally clear if they encode conceptual or procedural meaning. Another interesting factor comes from diachronic analysis of pragmatic markers, in that the moment of the emergence of their discourse-pragmatic functions is generally hard to determine. This is because most pragmatic markers emerged through processes of pragmaticalization from already existing lexical items (see Aijmer 1997, Dostie 2004). But it is also due to the fact that pragmatic markers generally occur in spoken language, and often only in very informal speech. Therefore, it is difficult to retrace their diachronic evolution by means of written corpus data.

While it is evident that pragmatic markers have specific characteristics in comparison to other lexical units, it is still not clear what happens to them in situations of strong language contact. Clyne (1972) investigated the borrowing of German pragmatic markers in the English discourse of German-speaking immigrants in Australia. Mougeon & Beniak (1991) examined the borrowing of English pragmatic markers in Canadian French discourse. There is a wide range of examples that account for the influence of the pragmatic markers from one language on the other language in situations of intensive and/or long-term language contact. Torres & Potowski (2008: 264) attempted to classify the possible outcomes of pragmatic markers in contact as follows:

1. The two sets of discourse markers will coexist.
2. Similar markers from each language will acquire differentiated meanings.
3. The markers from one language may replace those of the other language.

To support this classification, they cite different examples of studies on pragmatic markers in language contact. As an example of the first case, they mention Hill & Hill's (1986) work on Spanish in contact with Mexicano and Brody (1987), who analyzed Spanish in contact with different indigenous languages, for example Mexicano and Mayan (Torres & Potowski 2008: 264). As an example of the second case, they cite Solomon (1995) and her work on Spanish in contact with Yucatec (Torres & Potowski 2008: 265). The third outcome of pragmatic markers in contact was examined by Goss & Salmons (2000) in their work on Texas German. In this case, the whole German set of markers was replaced by English markers (Goss & Salmons 2000).

The peculiarities of pragmatic markers in language contact are mostly due to their general characteristics. Still, it is unclear if the three outcomes mentioned above are mutually exclusive or if long-term, intensive language contact always results in outcome number three; namely, the complete replacement of one set of pragmatic markers. The opposite option would be that sets of pragmatic markers from two languages might co-occur over an extended period of time without having too strong an influence on each other. This research question will be investigated by means of data from Manitoban French in Section 3 of this paper. Previous research on pragmatic markers in contact has mainly focused on individual markers in one contact variety or on one specific outcome of a contact situation.

The English marker *so* (see also Section 3.2) has been investigated as a potential case of borrowing in different contact situations. Mougeon & Beniak (1991)

assume that all markers first occur in the respective other language as code switches and then gradually become borrowings. Mougeon & Beniak (1991: 199) state that *so* in Ontarian French can be regarded as a core lexical borrowing. They assert that the use of *so* in Canadian French discourse may be due to intensive, long-term language contact because it is particularly prominent in Canadian French varieties that have experienced strong language contact with English (Mougeon & Beniak 1991: 201). Although they argue that the “degree of bilingualism is a poor predictor of variation in *so* usage” (Mougeon & Beniak 1991: 201), they still find that *so* is mostly used by speakers who have the most contact with English in their everyday lives. This fact leads them to the assumption that “core lexical borrowings like *so* or other sentence connectors may start out as code switches (either as single words or as part of switched sentences) which by dint of repetition become loanwords” (Mougeon & Beniak 1991: 211). That is to say, more fluent bilingual speakers introduce the English marker to French discourse and less fluent speakers repeat this linguistic behavior. Still, they consider an explanation proposed in the work of Weinreich & Haugen under which these kinds of borrowings emerge through the “acculturation of bilingual speakers who experience high levels of contact with a superordinate language” (Mougeon & Beniak 1991: 212). The phenomenon of integrating English pragmatic markers in the discourse of another language is by no means restricted to the language pair English–French. Several authors have investigated the language contact of Spanish in the US, in particular the use of *so* and its Spanish equivalent *entonces* in the Spanish discourse of bilingual speakers (Silva-Corvalán 1995, Aaron 2004, Torres 2002, Lipski 2005, Torres & Potowski 2008). The same phenomenon can even be proved for other bilingual speakers, such as Croatian–English bilinguals (Hlavac 2006). Hlavac (2006) explains the frequent occurrences of *so* in bilingual Croatian discourse by the marker’s polyfunctionality. This characteristic cannot be assigned to its Croatian equivalents (Hlavac 2006: 1896).

The present corpus analysis aims to investigate a specific set of pragmatic markers in contact in more detail to identify distinct types of language-contact phenomena. To this aim, this paper focuses on different processes of contact-induced language change in a language-contact situation. According to Heine & Kuteva (2005: 2), “contact-induced influence manifests itself in the transfer of linguistic material from one language to another”. When talking about cross-linguistic change, Heine & Kuteva assume a model language (also referred to as the source language), providing the pattern for transfer, and a replica language (also referred to as the target or borrowing language), receiving the pattern. This paper follows Heine & Kuteva (2010) in their terminology for the main types of contact-induced linguistic transfer, which are borrowing and replica-

tion (see Heine & Kuteva 2010: 87). In this sense, the cross-linguistic transfer that affects “meanings (including grammatical meanings or functions) or combinations of meanings” (Heine & Kuteva 2005: 2) will be referred to as *replication*, while cross-linguistic transfer that affects “form–meaning units or combinations of form–meaning units” (Heine & Kuteva 2005: 2) will be termed *borrowing*. Both processes will be described in more detail in the following part of this paper.

3 Pragmatic markers in contact in Manitoban French

Due to its sociolinguistic and historical evolution, the French language in Manitoba has been exposed to a strong, long-term influence of English for more than two centuries. French has long been a minority language in the province of Manitoba. At the end of the nineteenth century, all French schools were banned in Manitoba, and for most of the twentieth century, the French community was not allowed to teach their children in French. It was only in 1979 that the French community regained the same rights as the English community and French became an official language *de jure* (see Hennecke 2014 for a detailed sociohistorical description). Still, Manitoban French remains a *de facto* minority language that is only spoken by 3.2 percent of the population of Manitoba (Statistics Canada 2016).¹ Therefore, Manitoban French has been and continues to be strongly influenced by English.

The following analysis is based on a corpus of spoken Manitoban French, the FM Corpus (see Hennecke 2014 for a detailed presentation and discussion of the transcriptions and the corpus data). The corpus data consist of recordings of informal everyday conversations. They contain 35,660 tokens, divided into 15 communications from 20 speakers. The corpus data of the FM Corpus were collected in 2010 and 2012 in St. Boniface, the French quarter of Winnipeg, and consist of two-thirds of French utterances, while English utterances only make up one-third of the data. The transcriptions of the corpus data are based on the HIAT convention to transcribe spoken data (*Halbinterpretative Arbeitstranskriptionen* ‘Semi-Interpretative Working Transcriptions’, see Ehlich & Rehbein 1976).² All speakers in the corpus are aged between 17 and 30 and were born and raised in the Franco-Manitoban environment of St. Boniface or its neighboring districts. All speakers indicated French as their mother tongue and can be identified as

¹<https://www150.statcan.gc.ca/n1/pub/89-657-x/89-657-x2019014-eng.htm> (accessed on 20.11.2021)

²See <https://exmaralda.org/de/hiat> for a detailed description of the HIAT transcription convention and a transcription manual.

balanced bilingual speakers of English and French. In the following examples from the corpus data, each speaker is referred to using an anonymous speaker ID (e.g., DM, ZA, GR). Due to the prevailing sociolinguistic circumstances, all speakers were regularly exposed to English and French in everyday life from an early age. Sections 3.1 to 3.3 will present and discuss selected examples of French and English markers in the FM Corpus; namely, *comme* and *like*; *alors*, *donc*, and *so*; and *bon*, *ben*, and *well*. The three sets of markers were selected because of their occurrence in the FM data (see Section 3.4 for further discussion). All examples in Sections 3.1 to 3.3 are taken from the FM Corpus, unless indicated otherwise. Section 3.4 will then discuss the absence of certain pragmatic markers in the FM Corpus data.

3.1 The markers *comme* and *like*

In European French, the conjunction and adverb *comme* ‘like’ is a highly multifunctional lexical unit. Even in its diachronic evolution, the lexical unit *comme* has expanded semantic patterns and developed several functions that can also be found in its Portuguese and Spanish counterpart *como* and its Italian counterpart *come* (Mihatsch 2009). In Canadian French, the lexical unit *comme* shows some important peculiarities. Recent findings suggest that *comme* has developed functions that are not attested for *comme* in European French. The most salient new functions of *comme* include the extension in its use as a hedge (e.g., a quantitative approximation marker) and its use in quotation. New functions that have been detected in Canadian French include quantitative approximation, or the rounder function according to Prince et al. (1982), and the indirect discourse and autocitation functions, which will be defined as *quotative* functions in the following, and the assertion function. In current research on *like*, this function is, from a syntactic perspective, commonly referred to as sentence-final use, or, from a pragmatic perspective, as focus function (see Underhill 1988). Depending on the pragmatic function of *like* in the specific utterance, assertion can also comprehend shield functions, according to the terminology of Prince et al. (1982). In European French, *comme* cannot fulfill this set of functions. Therefore, the question arises as to which underlying process of language change can be identified for the new meanings and functions of *comme* in Canadian French.

In spoken English, *like* functions as a highly polysemous and syntactic flexible lexical unit that can take several discourse functions. According to Meehan (1991: 49), *like* has been known in its function as a conjunction since the fourteenth century, and has since developed new functions, such as its use in exemplification and its different discourse functions, such as focus and quotative.

Besides these functions, *like* can also appear as a hedge and a hesitation marker. Some of the discourse functions of *like*, such as its use as a focus and quotative marker, only emerged in more recent times and became very frequent in spoken language. This frequency in spoken language, initially restricted to young speakers of American English, rapidly expanded to other sociolinguistic groups and other varieties of English spoken outside the US. This rapid evolution of language change took place in the second half of the twentieth century (e.g., Buchstaller 2002, Buchstaller & D'Arcy 2009, Vandelanotte & Davidse 2009).

The markers *comme* and *like* both seem to be very frequent in Manitoban French. In the FM Corpus data, *comme* occurs 577 times, whereas *like* appears 255 times. When regarding the occurrences of *like* in the FM Corpus, 237 items out of 255, or 93 percent, appear in the function of a pragmatic marker. The other 7 percent include occurrences of *like* as a verb or as a comparison marker. For *comme*, 554 occurrences of the item, or 96 percent, appear in the function of a pragmatic marker.

The FM Corpus data show that *like* takes functions that are commonly attested for American and Canadian French (D'Arcy 2017). The English marker *like* is highly flexible and can appear in various positions of an utterance and mark a large scope, as in (1).

- (1) DM: *like* what if • all the services and everything was already done in French you know.

The marker *like* in the FM Corpus commonly takes various hedging functions, which are either lexical (2) or numeric (3) approximation.

- (2) ZA: • I wanna do a *like* solo album pendant l'hiver.
 ZA: • I wanna do a *like* solo album during winter.
- (3) GR: well he's got • his eeh wine cellar in the basement he's got about *like* forty bottles.

Like in the FM Corpus also functions as a focus marker, highlighting specific focal information, as in (4).

- (4) ZA: ben c'est un film d'une heure et demi hein? • *comme* it's a feature length • • • deal • it's *like* huge • •
 ZA: 'well, it's an hour and a half film, right? • *like* it's a feature length • • • deal • it's *like* huge • •'

Further, in the FM Corpus, *like* can fulfill different quotative functions that are commonly attested in different varieties of English, for example introducing quoted speech (5), quoted thought (6), or quoted attitude (7).

- (5) DM: I called his eh constituency office the other day and I said I wanted to meet with them • I was *like* “I wanna meet you” eh • they/ they phoned back and they said that it was all booked up for September I’m *like* “that’s fine because I’m on a trip.”
- (6) ZA: alors c’est eux qui s’occupent de la distribution c’est/ they do all the work for me and submit it to festivals so I’m *like* “oh good I don’t have to worry about this.”
ZA: ‘so they’re the ones who take care of the distribution / they do all the work for me and submit it to festivals so I’m *like* “oh good I don’t have to worry about this.”’
- (7) WIL: and then • all of a sudden • she’s *like* “what a great Francophone scene we have there is in Winnipeg.”

Comme in Manitoban French can take a hedging function with a large scope (8) and mark numeric approximation (9).

- (8) DM: (...) in the meantime euhm je travaillais juste à *comme* • produire le document lui-même.
DM: ‘(...) in the meantime euhm I was just working on *like* • producing the document itself.’
- (9) CAR: ça fait *comme* cinq fois qu’ (elle) • te prend avec ça ((0.8s)) un joke
CAR: ‘it’s *like* five times that she • caught you with it ((0.8s)) a joke.’

The French marker *comme* may also take focus functions in Manitoban French, as in (10).

- (10) CAR: elle est vraiment *comme* la meilleure artiste de nos jours
CAR: ‘she is really *like* the best artist of our time’

With regard to the FM Corpus, one can observe a striking use of *être comme* ‘be like’ as a quotative. As in different varieties of English, *être comme* can fulfill the functions of introducing quoted speech (11), quoted thought (12), and quoted attitude (13).

- (11) GER: so Joey *était comme* • • “how about premier novembre?” ils *sont comme* “ok” • • he (did) a writing • he *est comme* “perfect”
 GER: ‘so Joey *was like* • • “how about first November?” they *are like* “ok” • • he (did) a writing • he *is like* “perfect”’
- (12) WIL: • • ça c’est la rumeur qui/ qui passe maintenant puis *j’suis comme* “I don’t care I’m getting her out.”
 WIL: • • ‘that’s the rumour that’s going around now and *I’m like* “I don’t care I’m getting her out.”’
- (13) GR: là tout le monde *est comme* “oh my god”.
 GR: ‘here everyone *is like* “oh my god”.’

Further, different varieties of English show the use of a quotative form, *go like*, and it is possible to observe the equivalent form, *aller comme*, in the FM Corpus.

- (14) DAN: I don’t know • mais quand tu fais une fau(te) ça *va comme* “cling cling cling” • and then ça ça va venir and then on va être tout frustrés and then on va • • casser les guitars.
 DAN: ‘I don’t know • but when you do a mistake *it goes like* “cling cling cling” • and then it’ll come and then we’ll get all frustrated and then we’ll • • break the guitars.’

In conclusion, the analysis of the markers *comme* and *like* revealed that the European French equivalent *genre*, a comparably new pragmatic marker in spoken European French, is not present in the corpus data. The marker *comme* has developed new meanings and functions, such as its use as quotative, hedging, and numeric approximation, which, at first sight, appear to be replicated from the English *like*. Following Heine & Kuteva (2005, 2010), replication may “in the same way affect morphological, syntactic, and pragmatic structures, the noun phrase, and the verb phrase in the same way as the organization of clauses and clause combining” (Heine & Kuteva 2005: 261). Furthermore, Heine & Kuteva (2010: 89) state that grammatical replication contrasts with borrowing in that it does not include the transfer of phonetic material, which is a crucial point in borrowing. Hence it is important to differentiate grammatical replication from polysemy copying, also called calquing or loan translations, in which a meaning is only copied. According to Heine, “polysemy copying can be described as an abrupt rather than a gradual change, and it tends to be associated with lexical rather than grammatical replication” (2012: 126).

In conclusion, it is not possible to term the process of the extension of pragmatic functions and semantic meaning patterns of *comme* in Manitoban French

unequivocally as contact-induced language change or even replication in the sense of Heine & Kuteva (2005, 2010). Similar processes have been reported in other Romance languages (e.g., Mihatsch 2009). Therefore, language contact might also be just one specific factor among others that accelerated the process of language change in Manitoban French compared to European French (for a detailed discussion, see Hennecke 2014).

3.2 The markers *alors*, *donc*, and *so*

In current research on French pragmatic markers, *alors* and *donc* have commonly been treated together. Mosegaard Hansen explains this by the fact that both markers “originate in temporal anaphoric expressions” and that both are “frequently used in argumentational structures, where they mark a result or a conclusion” (1998: 321). In modern French, *donc* has completely lost its original temporal use but has developed new discourse-pragmatic functions. According to Mosegaard Hansen (1997: 165), *donc* has two main functions: marking a conclusion, a consequence, or a result and marking repetitions such as reformulations, paraphrases, or summaries. Furthermore, *donc* can take emphatic functions that are not restricted to imperative phrases, as shown for Old French (Mosegaard Hansen 1998: 329). For *donc* as a marker of discourse structuring, Bolly & Degand (2009) establish a differentiation between its syntactic-semantic functions and its discourse functions. As a syntactic-semantic function, they list the use of *donc* as a conclusion or consequence marker (Bolly & Degand 2009: 7). As a discourse function, they specify the use of *donc* as a repetition marker, as a marker of participative transition, and as a marker of conceptual structuring (Bolly & Degand 2009: 12). Thereby, they distinguish two kinds of repetition markers: a repetition marker of conclusive orientation that includes a recapitulation, and a repetition marker that implies reformulation and explications (Bolly & Degand 2009: 12). In contrast, *donc* as a marker of conceptual structuring marks a reorientation toward a new subject or a subject that has been mentioned earlier in the conversation (Bolly & Degand 2009: 9).

Regarding the French *alors*, three functions can commonly be distinguished – the temporal, the causal, and the discourse-structuring functions (see Degand & Fagard 2011, Le Draoulec & Bras 2007, Mosegaard Hansen 1997). The marker *alors* can function as a consequence or result marker in Modern French (Degand & Fagard 2011: 9). As a causal marker, *alors* still contributes to the propositional content of an utterance. This is not the case for the metadiscursive use of *alors*, where the marker only modifies the illocutionary force of the utterance and “can be left out without changing the semantic content” (Degand & Fagard 2011: 15).

As a metadiscursive device, *alors* structures discourse or introduces new topics or topic shifts.

Apart from the function of marking results or conclusions, Mosegaard Hansen highlights two main functions of *alors*: marking reperspectivization or reorientation (1998: 335) and marking foregrounding (1998: 348). By reperspectivization or reorientation, Mosegaard Hansen understands the uses of *alors* as a structuring device, to introduce a new topic, parentheses, or citations. Furthermore, in this function, *alors* can be used as a topic and discourse starter (Mosegaard Hansen 1998: 335). When foregrounding, *alors* marks “transitions from more backgrounded to more foregrounded material, especially, but not exclusively in narrative” (Mosegaard Hansen 1998: 348).

From a crosslinguistic perspective, the English marker *so* is generally considered the translation equivalent of the French markers *alors* and *donc*. The marker *so* is among the best-investigated pragmatic markers in the English language. It is well known that *so* is a highly multifunctional particle that can occur in different grammatical and discourse-pragmatic functions. Schiffrin (1987) detects two separate but not exclusive functions of *so*. Firstly, she points out the function of *so* as a causal marker that connects propositional content or illocutionary acts. Secondly, she focuses on the purely discursive functions of *so* as an interaction marker (Schiffrin 1987: 218). In addition, in informal speech, “*so* is a turn-transition device which marks a speaker’s readiness to relinquish a turn” (Schiffrin 1987: 218). Bolden (2009) clearly distinguishes the inferential use of *so* from its utterance-initial functioning. According to Bolden, *so* can be seen as a marker of “emergence of incipency” and “is a resource for establishing discourse coherence and [...] accomplishing understanding” (2009: 996). The utterance-initial *so* is “used in contexts where a particular course of action is oriented to by the interlocutors as having been pending or relevantly missing” and “on turn constructional units that pursue abandoned or interrupted interactional projects” (Bolden 2009: 996). In conclusion, the marker *so* generally functions as a consequence, result, or conclusion marker, to introduce a recapitulation or reformulation of something said earlier, or to mark reorientation and reperspectivization (e.g., topic starting or topic changing).

As mentioned in Section 2, *so* has already been analysed as a possible borrowing in different situations of language contact. Here, most authors claim that *so* in the discourse of bilingual speakers is a fully integrated loan or a core borrowing. It can be stated that most researchers regard the transfer of form-meaning units of pragmatic markers as a case of borrowing (e.g., Mougeon & Beniak 1991, Silva-Corvalán 1995, Torres 2002, Torres & Potowski 2008). In contrast, Lipski argues that the insertion of the English *so* into Spanish discourse is a case of

“metalinguistic bracketing” that is “Spanish discourse filtered through the English metasystem” (2005: 13). He further postulates that this is possibly due to the simultaneous activation of the two languages (Lipski 2005: 6). This means the speakers utter the English marker unconsciously and this is to some degree a sign of the acculturation that the speakers experience. This idea points in the same direction as Matras (2000), who claims a cognitive trigger for bilingual discourse markers. According to Matras, “this cognitive motivation [...] is so strong that it will at times override the social and communicative constraints on the discourse, leading to counterstrategic, accidental, or unintentional choices (i.e., slips)” (2000: 514). This idea of a metasystem or cognitive filter cannot fully be adopted in this paper, as it does not sufficiently account for the fact that not all pragmatic markers of a language seem to underlie this filter and that the various markers of a language behave very differently in language contact.

All three markers introduced above are found in the data of the FM Corpus. It is striking that the French markers *donc* and *alors* appear infrequently in the corpus data and only in purely French utterances; that is, they only occur in utterances where the matrix language is French and never proceed or follow parts of English discourse. This is striking because pragmatic markers are generally known for being easy to insert in bilingual discourse, in part due to their semantic and syntactic detachability. Further, it can be detected in the data that *donc* and *alors* are only used by a small number of speakers.

The marker *donc* only has 12 occurrences in the FM Corpus and is used by only four speakers. Still, *donc* can be found in various pragmatic functions, such as a topic changer (15), a topic starter (16), and a marker of a conclusion or consequence (17).

- (15) FLO: •• ehm •• *donc* • Inga est-ce que tu l’as commencé ton chose?
FLO: •• ‘ehm •• *so* • Inga have you started your thing?’
- (16) FLO: Word ! ••• ok ••• *donc* est-ce que vous avez besoin du temps •••
on stage • avant de commencer?
FLO: ••• ‘ok ••• *so* do you need the time ••• on stage • before you start?’
- (17) DAN: À cause y a/ dans un chanson •• je • je nomme tous les membres du
groupe. FLO : cool • ok • *donc* moi j’v/ j’... j’ai pas besoin d’l’ faire (...)
DAN: ‘Because there is/ in a song •• I • I name all the members of the band.’
FLO: ‘cool • ok • *so* I’m going to... I don’t need to (...)’

Further, *donc* in the FM Corpus occurs as a consequence or result marker, as in (18), or as a reformulation marker, as in (19).

- (18) WIL: (...) en enlevant ça de mon/ de mes/ ma liste de dépenses • • • tu m'obliges de prendre leur quinze pourcent vers des coûts • • d'autres coûts • • • et *donc* je (suis) déficitaire de ce moment là • • •

WIL: '(...) by removing that from my/ my/ my list of expenses • • • you force me to take their fifteen percent towards costs • • other costs • • • and so I (am) in deficit from then on • • •'

- (19) WIL: (...) c'est elle qui décide combien d'argent est donné • ((...)) ... ya c/ mais non • ça c'est pas efficace • • • *donc* ((1.2s)) c'est elle qui écrit la lettre (...)

WIL: '(...) it's her who decides how much money is given • ((...)) ... but no • it's not effective • • • so ((1.2s)) it's her who writes the letter (...)'

In the same way, *alors* is used infrequently in the FM data and only by a small number of speakers. Still, some of its discourse-pragmatic functions that are documented in European French are also evident in Franco-Manitoban spoken discourse. *Alors* appears as a marker of consequence or result in (20) and as a marker of repetition (explication) in (21).

- (20) JO: puis c'est comme "ya je va me changer" *alors* il se change puis là il sort puis il avait sa casquette là puis ça • j'étais comme "what the...?"

JO : 'then it's like "yah I'm going to change" so he changes then he goes out then he had his cap on then it • I was like "what the...?"'

- (21) JO: elle travaille à • • à temps partiel *alors* elle travaille les après-midis puis c'est une classe d'onzième (...)

JO: 'she works • • part-time so she works in the afternoons then it's an eleventh grade class (...)'

Furthermore, in the FM Corpus data, the marker *alors* is used for turn management, for example as a topic changer (22) and as a topic starter (23).

- (22) JO: oh il était ici avant?

NI: oui.

JO: oh ya? • • • ha.

NI: *alors* vous prenez un cours ensemble? c'est quoi? (...)

JO: 'oh he was here before?'

NI: 'yes'

JO: 'oh yah ? • • • ha'.

NI: 'so you're taking a class together? What is it? (...)'

- (23) NI: (...) t'sais comme on se sert encore de ces choses là (...) mais eh *alors* lui il doit apprendre comme comment on travail/ travaille avec le cuir • puis la fourrure

NI: '(...) you know how we still use these things (...) but *then* he has to learn how to work with leather • then fur'

In the FM data, *alors* does not appear as a discourse-structuring device, but in Example (24) it is used to bridge a moment of discourse planning.

- (24) JO: mais quand même t'sais les personnes l'appellent Macaroni puis il y avait • • des (()) comme ça • ici puis *alors*...ouais ça • • *alors* • • *alors* quand même j' pense ça/ ça eu un effet (...)

JO: 'but still, you know, people call him Macaroni and then there were • • (()) like that • here and *then*... yeah that • • *then* • • *then* still I think that had an effect (...)

Like the marker *donc* in the FM Corpus, *alors* also occurs in a high number of pragmatic functions when compared to its infrequent use.

The English marker *so* appears in purely French and purely English discourse environments in the FM Corpus and in bilingual discourse. In this context, an utterance is considered bilingual if one language occurs on the left-hand side of the marker *so* and another language on the right-hand side, as in (25).

- (25) PJ: (...) but the most of it is there *so* je pourrais envoyer ça.

PJ: '(...) but the most of it is there *so* I could send that.'

When looking at the distribution of *so* in the FM Corpus, it is striking that it appears more frequently in bilingual (20%) or French (42%) sentence environments. *So* is used in purely English discourse in only 38% of the occurrences. Therefore, for the purpose of this analysis, the focus will be on *so* in French and bilingual contexts.

When indicating resultant parts of utterances, *so* in the FM Corpus can mark results, as in (26), and consequences, as in (27).

- (26) PJ: Ils ont fait une autre comme négative *so* it's just like totally the wrong pictures and they fucked with them.

PJ: 'They did another as negative *so* it's just like totally the wrong pictures and they fucked with them.'

- (27) DR: puis Damian va être ici aussi *so* he's canning the date with you tomorrow...
 DR: 'Then Damian will be here too *so* he's canning the date with you tomorrow...'

When marking a conclusion, *so* often includes the pragmatic functioning of introducing an explication, as in (28), or a reasoning (of something said earlier).

- (28) FLO: aah vous avez l'âge à mon petit frère (puis) ma petite soeur ((1.s)) je me sens vieille.
 GER: c'est des jumeaux?
 FLO: non • • mais l'une (est née) en quatre-vingt-onze puis l'un en quatre-vingt-treize • • *so* you are in the middle *so*
 FLO: 'ahh you are the age of my little brother (then) my little sister ((1.s)) I feel old'.
 GER: 'are they twins?'
 FLO: 'no • • but one (was born) in ninety-one then one in ninety-three • • *so* you are in the middle *so*'

The marker *so* as a repetition marker can introduce a further explication (29) or a reformulation (30) of something said earlier.

- (29) GR: ça parle de • • • comment que • l/ les français on voulait/ • les francophones on voulait nos droits puis là il y avait un backlash politique • sévère • • *so* il y a des anglophones • on pense • qui ont • • brûlé le bâtiment de la Société Franco-Manitobaine
 GR: 'it talks about • • • how • the French wanted • the Francophones wanted our rights and then there was a political backlash • severe • • *so* there are Anglophones • we think • who • burned the Société Franco-Manitobaine building'
- (30) DM: Ça j'ai écrit en anglais mais je voulais vraiment que ça soit en français aussi • euhm • • *so* eu-h • c'est c'est bien mais là • puisque ça traite de la culture dakota • euh • t'sais les amérindiens, right? c'est traduit en dakota aussi, *so* c'est trilingue in the end
 DM: 'I wrote this in English but I really wanted it to be in French as well • euhm • • *so* eu-h • it's good but • since it deals with Dakota culture • euh • you know Native Americans, right? It's translated into Dakota as well, *so* it's trilingual in the end'

Further, *so* also can fulfill typical discourse-management functions, such as topic starting (31) and topic changing (32).

- (31) CAR: hallo.
ME: hallo.
FLO: ya so moi je vais juste vous parler un petit peu parce que • il va avoir • du temps... ya il ya avoir du • temps pendant le show après le show comme ben/ après votre set • •
CAR: 'hallo'
ME: 'hallo'
FLO: 'yah so I'm just gonna talk to you a little bit because • there's gonna be • time ... yah there's gonna be • time during the show after the show like after your set • •'
- (32) WIL: shut up. shut up • so ya c'est un projet avec le CJP. C'est un projet originally du CJP.
WIL: 'shut up. shut up • so yah this is a project with the CJP .◡ this is an original project of the CJP.'

Unlike *alors* and *donc* in Franco-Manitoban discourse, *so* can occur in utterance-final positions in the FM Corpus data without difficulty (33).

- (33) DM: well they're pretty lucky they had/ they were pretty close to him so eh yeah.

In conclusion, the corpus-based analysis of the markers *alors*, *donc*, and *so* show that the markers *alors* and *donc* appear on a low-frequency basis, while the marker *so* occurs very frequently, especially in bilingual contexts and monolingual French discourse. Still, *alors* and *donc* have not lost any of their semantic meanings or pragmatic functions that are attested in spoken European French. Despite its increase in frequency and its use in bilingual and French discourse, *so* has not developed new functions or new meaning patterns in Manitoban French. The results from the corpus analysis and previous research on the marker *so* in other contact varieties indicate that *so* is indeed a case of borrowing from English (e.g., Mougeon & Beniak 1991, Torres & Potowski 2008).

3.3 The markers *ben*, *bon*, and *well*

The pragmatic markers *bon* and *ben* are derived from the adjective *bon* 'good' and from the adjective and adverb *bien* 'good' (Mosegaard Hansen 1998: 222). Waltereit (2007: 91) traces the reduced form of *bien*, *ben*, as a pragmatic marker back

to the eighteenth century and *bon* as a pragmatic marker even further, back to the sixteenth century (Waltereit 2007: 92). According to Mosegaard Hansen (1998: 225), the marker *bon* has two main functions: its interjective use and its proper discourse-marking use. The former includes the utterance-initial *bon*, which is mostly retroactive and indicates acceptance. In contrast, the latter use includes *bon* in non-utterance-initial positions. Here, *bon* can appear either in turn-final or turn-medial positions or inside a sentential structure (Mosegaard Hansen 1998: 234). Beeching ascribes the following functions and meanings to *bon*: “positive evaluation, acceptance, *mot de la fin*, provisional acceptance (stage-marking) and concession” (2011: 102). Here, provisional acceptance may mark conflicts of opinions among speakers. Furthermore, she recognizes the function of *bon* as a face-threat mitigator, a hesitation and repair marker, and a pause filler.

According to Mosegaard Hansen (1998: 247), in contrast to *bon*, *ben* marks the unacceptability and irrelevance of a discourse phenomenon. It can mark inaccuracy, lack of importance, or the obvious and superfluous (Mosegaard Hansen 1998: 247). Furthermore, she states that “*ben* always functions on a level of utterance content” (Mosegaard Hansen 1998: 234). Waters (2009) puts the focus on the discourse-structuring functions of *ben*. For her, the main functions of *ben* are as an initial turn-opener and “at the boundary between two intonation units” (Waters 2009: 15). In these two positions, *ben* then fulfills diverse pragmatic functions, all of which comment on the preceding utterance of the previous speaker or the current speaker themselves (Waters 2009: 15).

The marker *well* is among the best-investigated markers of the English language, and a large number of studies examine this marker from a wide range of perspectives (e.g., Lakoff 1973, Schiffrin 1987, Jucker 1997, Aijmer & Simon-Vandenberg 2003, Beeching 2011). This is mainly due to its frequency in spoken English and to the pragmatization pathways it has undergone (see Beeching 2011).

Most researchers agree on the fact that *well* can, among other things, embody some sort of positive value judgement (Aijmer & Simon-Vandenberg 2003), conformity to a norm (Bolinger 1989), or acceptance (Carlson 1984). Some authors additionally point out that the pragmatic marker *well* has a wide range of meanings that vary from partial agreement to complete disagreement (Cuenca 2008). If the speaker aims to express (partial) disagreement, the pragmatic *well* does indeed function as a face-threat mitigator in that it can introduce a dispreferred response or express demur (Beeching 2011). Here, the speaker can mark concession, flag incoherence (Beeching 2011), or indicate a discrepancy between propositional attitudes of the speaker and the hearer (Smith & Jucker 2000). In

this case, the speaker tries to reestablish common ground that was lacking before (Smith & Jucker 2000). *Well* as a face-threat mitigator is mostly employed in an utterance-initial position. This syntactic position is also commonly used to flag a conclusion or a partial conclusion. It is striking that *well* is mostly employed in utterance-initial positions, while it occurs rarely in utterance-medial or utterance-final positions. As an item in utterance-final positions, *well* is commonly used as a bracketing and hesitation device. On a discourse-structuring level, *well* may function as a repair and hesitation device, a boundary marker, or a pause filler.

The FM Corpus data show that the pragmatic marker *bon* is only used twice in the whole corpus data and only by one speaker. In these two occurrences, *bon* fulfills discourse-structuring functions, as in (34).

- (34) WIL: finalement la semaine passée j'appelle parce que j'ai envoyé • • • un courriel à ce temps là ((1.s)) basically disant "*bon* • • • Promo Musique vous allez pas couvrir ces coûts là vous allez couvrir ces coûts là au lieu • " (...)
WIL: 'finally last week, I call because I sent • • • an email at that time ((1.s)) basically saying "*well* • • • Music promotion you're not going to cover those costs there you're going to cover those costs there instead • " (...)'

In contrast, in the FM Corpus, *ben* marks a wide range of polysemous senses that vary from complete agreement (35) to partial agreement (36), partial disagreement (37), and complete disagreement (38).

- (35) GER: c'est comme "ooh • • c'était ça le maximum? • • • ha • mettez une liste avec tous les maximums"
WIL: *ben* oui • • c'est pas difficile • like what
GER: 'it's like "ooh • • was that the maximum? • • • ha • put a list with all the maximums"'
WIL: '*well* yes • • it's not difficult • like what'
- (36) ZA: ya • • t/ ya t'as pas le choix là? mais eeh...I wonder if you can peel off the sponsors ((laughing)) in fact just like/ • I like just...
PJ: *ben* il faut (quand même) connaître (()) but
ZA: 'yah • • yah you have no choice? But eeh... I wonder if you can peel off the sponsors ((laughing)) in fact just like/ • I like just...'
PJ: '*well* you have to know anyway (()) but'
- (37) CAR: c'est comme vingt secondes.
FLO: ah ok.

DAN: *ben* trente secondes.

CAR: 'it's like twenty seconds.'

FLO: 'ah ok.'

DAN: '*well* thirty seconds.'

- (38) CAR: ya peut-être une minute.

DAN: non • on smash pas les guitars pendant (())

CAR: *ben* non parce que faut qu'on • faut qu'on smash • • and then toi tu va chercher ta guitare •

CAR: 'yah maybe a minute ago.'

DAN: 'no • we don't smash the guitars during (())'

CAR: '*well* no because we have to • we have to smash • • and then you go and get your guitar •'

Further, *ben* can mark that something is obvious. In (39), the speaker signals that the previous utterance was superfluous or not necessary for the current conversation.

- (39) GR: in what sense? ((laughing))

DM: *ben* • • • in the fullest sense right?

GR: ya. ya. ce serait awesome.

GR: 'in what sense?' ((laughing))

DM: '*well* • • • in the fullest sense right?'

GR: 'yah. yah. That would be awesome.'

When structuring discourse, *ben* can be used not only to introduce new or pending topics, but also to mark the beginning of a subtopic or a bracket (40), introduce reported speech (41), or flag a conclusion (42).

- (40) NI: on arait • on avait une atten/ une entente avec/ • *ben* en plus ils chargent comme 30 dollars de l'heure so

NI: 'we would • we had a / an agreement with/ • *well* besides that they charge like 30 dollars an hour so'

- (41) WIL: puis là dans un courriel elle dit "*ben* chose certaine • les CBL là seront eh seront couverts là ces coûts là seront couverts • • (...)"

WIL: 'and then in an email she says "*well*, one thing for sure • the CBL there will be eh will be covered there these costs there will be covered • • (...)"'

- (42) DM: ya il était comme “je suis fou de Winnipeg” e h ya. c’est awesome!
c’est comme • il est déjà franco-manitobain là you know (())) *ben* c’est ça.
DM: ‘yah he was like “I’m crazy about Winnipeg” e h yah. It’s awesome!
it’s like • he’s already Franco-Manitoban there you know (())) *well* that’s
it.’

The pragmatic marker *well* is not a particularly frequent item in Franco-Manitoban spoken language. Still, a range of varying senses and functions that are attributed to the marker in spoken colloquial English can also be found in the FM Corpus data. *Well* in the FM data can mark complete agreement only when it is combined with another item of positive evaluation, as in (43).

- (43) GR: c’est ça • • you make them look good.
WIL: ya exactement • • *well* • oui • • • puis I wanna use that argument and
I know it’s true I just need the backup. • • •
GR: ‘that’s it • • you make them look good.’
WIL: ‘yah exactly • • *well* • yes • • • then I wanna use that argument and I
know it’s true I just need the backup. • • •’

In Example (44), *well* expresses partial agreement with the framework of the preceding speaker.

- (44) ME: so you wanna go to • • to Europe?
NI: eh • y/ yeah ((1s)) yeah • • • *well* like • • I’ve been to France ((1.3s)) and
I’ve been to London. I’d go back ‘cause I didn’t spend very much time there
• •

Concerning its pragmatic functions, *well* in the FM Corpus can mark concession (45) or can be used as a face-threat mitigator (46).

- (45) ME: he died last year.
DM: yeah. • • *well* they’re pretty lucky they had/ they were pretty close to
him so eh. yeah. yeah. yeah.
- (46) ZA: is any of it good? • • Who are these people? I’ve never heard of these’
WIL: eehm...
ZA: hm *well* ((1.2)) do these guys actually get to have careers? comme •
est-ce qu’ils font de l’argent?
ZA: ‘is any of it good? • • Who are these people? I’ve never heard of these’
WIL: ‘eehm...’
ZA: ‘hm *well* ((1.2)) do these guys actually get to have careers? Like • do
they make money?’

It is striking that in most occurrences in the FM data, *well* acts as a discourse-structuring device, such as for bracketing (47) and at the beginning of reported speech (48).

- (47) GR: they were gonna enact a law that made the province bilingual ((1.5s)) but (()) la crise linguistique *well* there's a huge backlash from the public within the party. yeah I guess.
- (48) SA: je commençais à enseigner and like • (après) comme trois ans j'avais comme • • • comme presque cent étudiants like/ like to myself and I was like "ah *well* I guess comme • je devrais peut-être (()) l'enseignement".
SA: 'I was starting to teach and like • (after) like three years I had like • • • like almost a hundred students like/ like to myself and I was like "ah *well* I guess like • maybe I should (()) be teaching".'

In conclusion, the corpus-based analysis of the markers *bon*, *ben*, and *well* demonstrated that the markers *bon* and *well* only occur infrequently and mostly in monolingual contexts in the FM Corpus data. In contrast, the marker *ben* appears particularly frequently in the data, but has not undergone any other changes in its semantic meaning patterns or its discourse-pragmatic functions. Furthermore, it is striking that *ben* generally occurs in monolingual contexts and cannot be considered a case of borrowing in English discourse. It can be speculated that the marker *ben* is preferred over other partially equivalent markers such as *bon*, *(en)fin*, or *bref* because of its strong semantic overlap with the English *well*. Still, it remains unclear why the markers *ben* and *well* do not coexist to the same degree in the FM data. More large-scale corpus data is needed to investigate this phenomenon more closely.

3.4 The absence of pragmatic markers in Manitoban French

Research on pragmatic markers has shown that co-occurrence of discourse features is very common, not only in French but also in English (Pichler & Levey 2010). For instance, speakers of European French have a wide range of co-occurring markers at their disposal and these markers are used frequently in spoken language. Surprisingly, very frequent markers from European French, such as *en-fin/fin* 'so, well', *bref* 'well', *tu vois* 'you see', *genre* 'like', and *quoi* 'what', do not occur at all in the Manitoban French corpus data. Furthermore, even frequent markers that have emerged in Quebec French, such as *(ça) fait que* 'well' and *coudon* 'so', do not appear at all in the FM Corpus.

In addition, it is striking that all markers analyzed in Sections 3.1 to 3.3 seem to have undergone changes in frequency and/or productivity. While *comme*, compared to European French, has experienced a huge increase in frequency that goes hand-in-hand with a broadening of productivity, other markers such as *bon* have seen a considerable decrease in frequency, together with a possible, yet not unequivocally provable, decrease in productivity. At this point, it is important to mention that an increase in frequency of a certain phenomenon does not necessarily correlate with an increase in productivity of the same phenomenon and vice versa (see Poplack 2001, Poplack & Levey 2010). Using a variationist sociolinguistic approach, Poplack (2001) proved that while the use of the subjunctive in Quebec French is decreasing, the use of the subjunctive with the verbs *valoir* 'to be worth' and *falloir* 'must' is increasing (Poplack & Levey 2010). This means that only a small number of verbs are used with the subjunctive in spoken Quebec French, but this small number of verbs is used considerably more often with the subjunctive than before (Poplack & Levey 2010).

In the use of pragmatic markers in Manitoban French, we see a somewhat similar but not comparable evolution. While a small number of markers, such as *like*, *comme*, *so*, and *ben*, are used very frequently, there is no comparable co-occurrence of markers. There may be two explanations for the lack of variation in the discourse-marking system in Manitoban French. Sankoff et al. (1997) observed in their corpus data from Anglophone L2 French speakers in Montreal that the speakers used significantly fewer pragmatic markers when speaking in their L2 than in their L1. In this case, the use of pragmatic markers increased in parallel to the L2 language skills (Sankoff et al. 1997: 213). Therefore, native language skills may be one plausible reason for the size of the discourse-marking system in Manitoban French. Furthermore, Sankoff et al. (1997: 214) identified a correlation between the use of certain pragmatic markers and sociolinguistic factors such as a speaker's childhood environment, gender, and social class. An in-depth analysis of sociolinguistic factors relating to the speakers in the FM Corpus is not possible within the framework of this study. Hence, the description of the corpus data in Section 3 must suffice to provide an insight into the speakers' sociolinguistic environment and their language skills (see also Hennecke 2014). All the speakers in the FM Corpus consider French as their L1, despite the strong influence of English in all situations of their everyday lives. All the speakers in the corpus received their education up to their high school diploma exclusively in French. Furthermore, all the speakers live in St. Boniface, the French quarter of Winnipeg, and actively participate in social and cultural activities in their community. These facts do not provide evidence that a lack of language skills may trigger a lack of variation in the discourse-marking system of Manitoban

French. A certain influence of specific sociolinguistic factors cannot be ruled out, but these factors do not seem to be the only reason for the evolution of pragmatic markers in Manitoban French. Further, the evolution of spoken Manitoban French and its sociohistorical development may have influenced the system of pragmatic markers. Here, the strong language contact with English and the constant situation of bilingual discourse in the everyday lives of the speakers may play a role in the absence of certain markers. More research on this topic needs to be done to find unequivocal explanations for this specific phenomenon.

4 Conclusion

In previous studies on pragmatic markers, it has been stated that these items may undergo different processes in language contact (e.g., Mougeon & Beniak 1991, Torres & Potowski 2008). These studies claim that pragmatic markers are well suited to borrowing in language contact but that similar markers from two languages may also acquire different meanings or be replaced by one item from one language. It has even been suggested that two sets of pragmatic markers may coexist or that all markers from one language may replace all markers from the other language (e.g., Brody 1987, Goss & Salmons 2000). This study aimed to investigate the processes and outcomes of language change in a long-term situation of language contact by means of a self-compiled corpus of bilingual Franco-Manitoban conversations.

The analysis of pragmatic markers in the FM Corpus shows four different outcomes of pragmatic markers in contact. Firstly, the marker *comme* takes functions from the English equivalent *like* and sees a rise in frequency. Secondly, the English marker *so* is borrowed in French discourse, while its equivalent French markers *alors* and *donc* are used infrequently but keep their functions in French discourse. Thirdly, the French marker *ben* and its equivalent *well* occur frequently and mostly in monolingual discourse. In contrast, the marker *bon* is almost inexistent in the corpus data. Finally, some markers that are frequent in varieties of European or Québec French do not occur at all in Manitoban French (e.g., *bref*, *quoi*, *(en)fin*).

The analysis also reveals five different outcomes of language contact on the system of pragmatic markers in Manitoban French:

1. Contact-induced language change
 - a) Emergence of new semantic meaning patterns
 - b) Emergence of new discourse-pragmatic functions

2. Borrowing of a marker from the other language
3. Frequency change
 - a) For the benefit of a marker from the same language
 - b) For the benefit of a marker from the other language
4. Coexistence of two markers
 - a) Coexistence in their original linguistic system
 - b) Coexistence in the same linguistic system
5. Absence of pragmatic markers

This analysis showed that the classification provided by Torres & Potowski (2008) is too limited and restricted, in that language contact can have complex and diverse impacts on systems of pragmatic markers in contact. Further, the analysis demonstrated that the different types of language change are not mutually exclusive in one system of pragmatic markers in contact. In conclusion, this corpus analysis revealed that pragmatic markers in a contact situation might indeed undergo different processes of language change. The analysis highlighted three processes; namely, the contact-induced change of the marker *comme*, the borrowing of the marker *so*, and a decline in frequency, particularly in the case of the markers *donc* and *bon*.

These findings support previous research results on pragmatic markers in language contact (e.g., Torres & Potowski 2008, Hlavac 2006). Furthermore, this study highlights the necessity of detailed cross-language analyses of pragmatic markers to determine their contact-induced changes in their pragmatic functioning and meaning patterns more precisely.

References

- Aaron, Jessi Elana. 2004. "So respetamos un tradición del uno al otro": *So* and *entonces* in New Mexican bilingual discourse. *Spanish in Context* 1(2). 161–179.
- Aijmer, Karin. 1997. *I think* – An English modal particle. In Toril Swan & Olaf J. Westvik (eds.), *Modality in Germanic languages. Historical and comparative perspectives*, 1–48. Berlin: De Gruyter Mouton.
- Aijmer, Karin. 2002. *English discourse particles: Evidence from a corpus*. Amsterdam: John Benjamins.

- Aijmer, Karin & Anne-Marie Simon-Vandenberg. 2003. The discourse marker *well* and its equivalents in Swedish and Dutch. *Linguistics* 41. 1123–1161.
- Aijmer, Karin & Anne-Marie Simon-Vandenberg (eds.). 2006. *Pragmatic markers in contrast*. Oxford: Elsevier.
- Andersen, Gisle. 2001. *Pragmatic markers and sociolinguistic variation: A relevance-theoretic approach to the language of adolescents*. Amsterdam: John Benjamins.
- Beeching, Kate. 2011. The translation equivalence of *bon*, *enfin*, *well* and *I mean*. *Revue française de linguistique appliquée* XVI(2). 91–105.
- Bolden, Galina B. 2009. Implementing incipient actions: The discourse marker ‘so’ in English conversation. *Journal of Pragmatics* 41. 974–998.
- Bolinger, Dwight. 1989. *Intonation and its uses: Melody in grammar and discourse*. London: Edward Arnold.
- Bolly, Catherine & Liesbeth Degand. 2009. Quelle(s) fonction(s) pour *donc* en français oral ? Du connecteur conséquentiel au marqueur de structuration du discours. *Lingvisticæ Investigationes* 32(1). 1–32.
- Brody, Jill. 1987. Particles borrowed from Spanish as discourse markers in Mayan language. *Anthropological Linguistics* 29. 507–521.
- Buchstaller, Isabelle. 2002. He goes and I’m like: The new quotatives re-visited. *Internet Proceedings of the University of Edinburgh Postgraduate Conference*. <http://www.lel.ed.ac.uk/~pgc/archive/2002/proc02/buchstaller02.pdf> (21 November, 2020).
- Buchstaller, Isabelle & Alexandra D’Arcy. 2009. Localized globalization: A multi-local, multivariate investigation of quotative *be like*. *Journal of Sociolinguistics* 13. 291–331.
- Carlson, Lauri. 1984. ‘Well’ in dialogue games: A discourse analysis of the intersection ‘well’ in idealized conversation. Amsterdam & Philadelphia: John Benjamins.
- Clyne, Michael. 1972. Perception of code-switching by bilinguals: An experiment. *ITL - International Journal of Applied Linguistics* 16. 45–48.
- Cuenca, Marie-Josep. 2008. Pragmatic markers in contrast: The case of *well*. *Journal of Pragmatics* 40. 1373–1391.
- D’Arcy, Alexandra. 2017. Discourse-pragmatic variation in context: Eight hundred years of LIKE. *ICAME Journal* 42(1). 219–224.
- Degand, Liesbeth & Benjamin Fagard. 2011. *Alors* between discourse and grammar: The role of syntactic position. *Function of Linguagae* 18(1). 29–56.
- Dostie, Gaétane. 2004. *Pragmaticalisation et marqueurs discursifs: Analyse sémantique et traitement lexicographique*. Brüssel: De Boeck & Larcier.

- Ehlich, Konrad & Jochen Rehbein. 1976. Halbinterpretative Arbeitstranskriptionen (HIAT). *Linguistische Berichte* 45. 21–41.
- Goss, Emily L. & Joseph C. Salmons. 2000. The evolution of a bilingual discourse marking system: Modal particles and English markers in German-American dialects. *International Journal of Bilingualism* 4(4). 469–484.
- Heine, Bernd. 2012. On polysemy copying and grammaticalization in language contact. In Claudine Chamoreau & Isabelle Léglise (eds.), *Dynamics of contact-induced language change*, 125–166. Berlin: De Gruyter Mouton.
- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Heine, Bernd & Tania Kuteva. 2010. Contact and grammaticalization. In Raymond Hickey (ed.), *The handbook of language contact*, 86–105. West Sussex: Blackwell.
- Hennecke, Inga. 2014. *Pragmatic markers in Manitoban French - A corpuslinguistic and psycholinguistic investigation of language change*. Ruhr-Universität Bochum. (Doctoral dissertation).
- Hill, Jane H. & Kenneth C. Hill. 1986. *Speaking Mexicano: Dynamics of syncretic language in Central Mexico*. Tucson: University of Arizona Press.
- Hlavac, Jim. 2006. Bilingual discourse markers: Evidence from Croatian–English code-switching. *Journal of Pragmatics* 38. 1870–1900.
- Jucker, Andreas. 1997. The discourse marker *well* in the history of English. *English Language and Linguistics* 1(1). 91–110.
- Kaltenböck, Gunter, Wiltrud Mihatsch & Stefan Schneider (eds.). 2010. *New approaches to hedging*. Bingley: Emerald.
- Lakoff, George. 1973. Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic* 4(2). 458–508.
- Le Draoulec, Anne & Myriam Bras. 2007. *Alors* as a possible temporal connective in discourse. *Cahiers Chronos* 17. 81–94.
- Lipski, John M. 2005. Code-switching or borrowing? No sé so no puedo decir, you know. In Lotfi Sayahi & Maurice Westmoreland (eds.), *Selected Proceedings of the Second Workshop on Spanish Sociolinguistics*, 1–15. Somerville, MA: Cascadia Proceedings Project.
- Maschler, Yael. 2000. Discourse markers in bilingual conversation. *Special issue of International Journal of Bilingualism* 4. 437–561.
- Matras, Yaron. 2000. Fusion and the cognitive basis for bilingual discourse markers. *International Journal of Bilingualism* 4(4). 505–528.
- Meehan, Teresa. 1991. It's like 'What's happening in the evolution of like?': A theory of grammaticalization. *Kansas Working Papers in Linguistics* 16. 37–51.

- Mihatsch, Wiltrud. 2009. The approximators French *comme*, Italian *come*, Portuguese *como* and Spanish *como* from a grammaticalization perspective. In Corinne Rossari, Claudia Ricci & Adriana Spiridon (eds.), *Grammaticalization and pragmatics. Facts, approaches, theoretical issues*, 65–91. Leiden: Brill.
- Mosegaard Hansen, Maj-Britt. 1997. 'Alors' and 'donc' in spoken French: A re-analysis. *Journal of Pragmatics* 28(2). 153–187.
- Mosegaard Hansen, Maj-Britt. 1998. *The function of discourse particles: A study with special reference to spoken standard French*. Amsterdam: John Benjamins.
- Mougeon, Raymond & Edouard Beniak. 1991. *Linguistic consequences of language contact and restriction: The case of French in Ontario, Canada*. Oxford: Oxford University Press.
- Pichler, Heike & Stephen Levey. 2010. Variability in the co-occurrence of discourse features. *Language Studies Working Papers* 2. 17–27.
- Poplack, Shana. 2001. Variability, frequency and productivity in the irrealis domain of French. In Joan L. Bybee & Paul J. Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, 405–428. Amsterdam: John Benjamins.
- Poplack, Shana & Stephen Levey. 2010. Contact-induced grammatical change. In Peter Auer & Jürgen Erich Schmidt (eds.), *Volume 1 Theories and methods: An international handbook of linguistic variation*. 391–419. Berlin: Mouton de Gruyter.
- Prince, Ellen, Joel Frader & Charles Bosk. 1982. On hedging in physician-physician discourse. In Robert Dipietro (ed.), *Linguistics and the professions*, 83–97. New York: Ablex Publishing Corporation.
- Sankoff, Gillian, Paul Thibault, Naomi Nagy, Hélène Blondeau, Marie-Odile Fonollosa & Lucie Gagnon. 1997. Variation in the use of discourse markers in a language contact situation. *Language Variation and Change* 9(2). 191–217.
- Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge: Cambridge University Press.
- Silva-Corvalán, Carmen. 1995. Spanish in four continents: Studies in language contact and bilingualism. *International Journal of Applied Linguistics* 10(1). 141–145.
- Smith, Sara & Andreas Jucker. 2000. *Actually* and other markers of an apparent discrepancy between propositional attitudes of conversational partners. In Gisle Andersen & Thorstein Fretheim (eds.), *Pragmatic markers and propositional attitude*, 207–237. Amsterdam: John Benjamins.
- Solomon, Julie. 1995. Local and global functions of a borrowed/native pair of discourse markers in a yucatec maya narrative. In *Proceedings of the twenty-first annual meeting of the berkeley linguistics society*, 287–298.

- Torres, Lourdes. 2002. Bilingual discourse markers in Puerto Rican Spanish. *Language in Society* 31(1). 65–83.
- Torres, Lourdes & Kim Potowski. 2008. A comparative study of bilingual discourse markers in Chicago Mexican, Puerto Rican, and MexiRican Spanish. *International Journal of Bilingualism* 12(4). 263–279.
- Underhill, Robert. 1988. *Like* is, like, focus. *American Speech* 63(3). 234–246.
- Vandelanotte, Lieven & Kristin Davidse. 2009. The emergence and structure of *be like*. *Cognitive Linguistics* 20(4). 777–807.
- Waltereit, Richard. 2007. À propos de la genèse diachronique des combinaisons de marqueurs: L'exemple de *bon ben* et *enfin bref*. In Gaétane Dostie & Claus D. Pusch (eds.), *Les marqueurs discursifs*, 94–109. Paris: Larousse/Armand Colin.
- Waters, Sophia. 2009. The semantics of French discourse particles 'quoi' and 'ben'. In *Selected Papers from the 2009 Conference of the Australian Linguistic Society*, 1–25.

Chapter 10

Phonic adaptation of Spanish Anglicisms in Mexico and Spain: A corpus data analysis

Linda Bäumler

University of Vienna

In this chapter, I discuss the phonic adaptation of Anglicisms by Spanish natives from Mexico and Spain. Since phonemes of the source language (in this case, English) which do not exist in the recipient language (in this case, Spanish) are of special interest in loanword phonology, I based my study on the realization of the English voiced alveopalatal affricate /dʒ/. Previous research describes mainly two different ways of pronouncing this foreign phoneme: the realization of grapheme-phoneme correspondence as in *manager* ['manaxer] on the one hand, and the adaptation of the foreign phoneme as [j] as in ['manajer] on the other. However, the comparative oral corpus data analyzed in this study also reveals many cases of phoneme importation, corresponding to pronouncing [dʒ]. Moreover, the data suggest that regarding the perception process, the most striking feature of the foreign affricate is [+VOICE], since more speakers realize a voiced palatal sound than an unvoiced postalveolar affricate when imitating the foreign sound. In addition, statistical analysis using a generalized linear mixed-effects model revealed that, at least in this data set, a speaker's country of origin does not affect the pronunciation of the loanword. This suggests that in a globalized world, geographic proximity to the US might not be a relevant factor influencing the realization of the phoneme. However, the level of exposure and affinity a speaker has to the English language and American culture is the most relevant factor actually influencing the pronunciation of Anglicisms in this study.



1 Introduction

1.1 Loanword phonology in Spanish-English language contact

For loanword phonology, those phonemes of the source language that do not exist in the phonological system of the recipient language are of special interest. Since Spanish incorporates an increasing amount of English loanwords due to globalization (Cabanillas & Martínez 2012: 95; Oncins-Martínez 2009: 116; Gómez Capuz 2001: 16; Schweickard 1991), this provides us with vast material to examine how Spanish natives pronounce these unknown phonemes in the respective Anglicisms. In general, speakers tend to either adapt the foreign phoneme to the Spanish phoneme system or to imitate the English model.

In the case of imitation, the literature describes several different possibilities: First, the unknown English sound can be replaced by a native Spanish sound that is close to the foreign sound (Kang 2011: 8; Gómez Capuz 2001: 27, 52; Pratt 1980: 155). An example of this would be the realization of *brunch* as [brantʃ] instead of [bɾantʃ]. Second, sounds of the source language (English) that exist in the phonological system of the recipient language (Spanish), but only in a specific phonological environment, are used in new phonological environments. Haugen (1950: 216) refers to this phenomenon as *phonemic redistribution*. An example of this is the realization of the consonant [tʃ] in word-final position, as in the Anglicism *sandwich*. Note that in Standard Spanish, the affricate is not used in word-final position (Gómez Capuz 2001: 49). Third, the foreign sound is imported to the borrowing language, a process that Haugen (1950: 216) refers to as *phonemic importation* (e.g., realization of [ʌ] in *brunch*). For Spanish, Gómez Capuz (2001: 52) states that even though the influence of English is constantly rising, phonemes are at present not imported and speakers would prefer the first two ways of imitation. However, a prominent example of an English phoneme that got imported is the fricative [ʃ], as in *show*. On the other hand, if speakers tend not to imitate the foreign English sound, they might realize Spanish grapheme-phoneme correspondences, and therefore realize *brunch* as [bruntʃ] rather than [bɾantʃ], as in English, for instance.

In the literature, different factors that influence the pronunciation of loanwords are discussed. Various authors argue that the age of the loanword affects its pronunciation (Rodríguez González 2017: 127; Haspelmath 2009: 42; Gómez Capuz 2001: 16; Haugen 1950: 216). Imitation of the foreign sound is more likely to occur when speakers are pronouncing younger Anglicisms than when pronouncing older ones. In addition, more frequent loanwords are more likely to undergo phonic adaptation (Rodríguez González 2018: 354). Moreover, the way

in which loanwords enter the language has been mentioned as a relevant factor in this process. Pratt (1980: 16) distinguishes between *eye loans*, that is, loanwords that entered the recipient language through the written channel, and *ear loans*, which were introduced orally into the recipient language. In the former case, the graphic and therefore visual code plays a major role, while in the latter, perception influences the integration of the loanword. Furthermore, it has been suggested that speakers' competence in English and their level of education affect the pronunciation of Anglicisms (Rodríguez González 2017: 102, 123, 127; Gómez Capuz 2001: 42). Speakers with good language skills and/or a high level of education are claimed to be more likely to imitate the foreign English sound. Moreover, the high prestige of the English language might prompt the imitation of the latter (Pustka 2021: 508; Gómez Capuz 2001: 42; Meisenburg 1992: 51). In addition, due to geographic proximity to the US, a higher impact of English on Hispanoamerican varieties of Spanish than on those of Spain has been suggested (Pustka 2021: 504; Rodríguez González 2017: 127; Oncins-Martínez 2009: 116).

To identify and analyze phonic adaptation of loanwords and the respective variables that influence the process, huge oral corpora¹ are needed. Nevertheless, Spanish Anglicisms have mostly been studied by means of written corpora (Rodríguez González 2017: 100; Gómez Capuz 2001: 6). Therefore, the pronunciation of Anglicisms has only occasionally been the subject of studies. Pratt (1980) works among other things with television recordings, but discusses the phonic realization of Anglicisms only briefly. Lorenzo (1997) examines the integration of vowels but analyzes mainly written data. Gómez Capuz (2001) analyzes the phonic assimilation of Anglicisms by means of the oral corpus Val.Es.Co (Briz & Grupo Val.Es.Co. 1995) and provides the reader with Anglicisms used in spontaneous speech. Although this gives insights into many different observable realizations, the corpus does not allow for systematic comparison between many speakers. LaCharité & Paradis (2005) analyze English and French loanwords in different oral corpora, among which is also Mexican Spanish. Data were gathered from written documents and spontaneous speech, and the pronunciations of the loanwords were then verified with a minimum of three L1 speakers (LaCharité & Paradis 2005: 228). Even though their corpus contains a huge amount of loanwords, it lacks realizations of many different speakers that could be compared. Rodríguez González (2017) analyzes the adaptation of consonants and vowels that are unknown to the Spanish phoneme system and bases his analysis on his personal observations of the spoken language during different decades, especially on the occasion of the preparation of dictionaries of Anglicisms (Rodríguez

¹For the definition of the term *corpus* used by the author, see Section 2.1.1.

González 2017: 100). At the time of writing, to the best of my knowledge, no large-scale corpus-phonological data set exists for Spanish Anglicisms. However, this is indispensable to investigate the pronunciation of Anglicisms systematically. Such corpus data not only allow for detection of different realizations and comparative studies, but also for revealing the different social and linguistic factors that influence the pronunciation of the loanwords.

1.2 The English phoneme /d͡ʒ/

When comparing the phonological systems of English and Spanish, many consonants that do not exist in Spanish but do in English stand out. Among these is the English voiced alveopalatal affricate /d͡ʒ/ that appears in words such as *jersey* and *jazz*. Standard Spanish² does not possess an alveopalatal affricate that would be voiced. Instead, it has the voiceless variant that appears for example in *chico* [ˈt͡ʃiko] ‘small boy’ and shares the same place and manner of articulation, but unlike /d͡ʒ/ is not voiced.

The literature attests the adaptation of the English voiced affricate as [j] by Spanish natives in Spain (Rodríguez González 2017: 107; Gómez Capuz 2001: 7).³ Gómez Capuz (2001: 29) assumes that when speakers perceive the loanword, the sonority of the phoneme plays a major role, and Spanish natives therefore tend to realize [j] instead of the voiceless affricate [t͡ʃ], which except for [+VOICE], shares the same features as [d͡ʒ]. This is sometimes also reflected in the spelling of the respective loanwords: *yins* (< jeans), *yip* (< jeep) or *yonqui* (< junkie) (Rodríguez González 2017: 107).

Generally, Rodríguez González (2017: 107) states that mostly speakers who are not familiar with English realize the Spanish grapheme-phoneme correspondence and therefore pronounce [x] instead of [j], as in *jeans* [xeans] (versus [jins]), *manager* [maˈnaxer] (versus [ˈmanajer]), or *challenger* [t͡ʃalenxer] (versus [t͡ʃalenjer]). Gómez Capuz (2001: 21) claims that the higher the age of the loanword, the more likely [x] and therefore grapheme-phoneme correspondence is realized.

²Note that the voiced alveopalatal affricate is known in certain Spanish varieties – in Paraguay, some areas of Ecuador, and Argentina, in the realization of <ll>, <y> or <(h)i> (Real Academia Española & Asociación de Academias de la Lengua Española 2011: 226).

³Note that Gómez Capuz (2001: 29) refers to a fricative and transcribes the allophone by means of the *Alfabeto de la Revista de Filología Española* with the symbol [y]. Rodríguez González (2017: 107) also transcribes [y]. However, I decided in this work to transcribe [j] and refer to an approximant realization, since previous research revealed that in Spanish, realization of [j] is far more frequent than [j̞] (Martínez Celdrán 2015).

In some cases, the realization is also found as [g], for example [ˈbangi] for *bungee* (Rodríguez González 2017: 108). Rodríguez González (2017: 107) documents cases where the English phoneme is realized as the voiceless variant [tʃ] in intermediate or word-final position (as in *pidgin*, *backstage*, *porridge*). Additionally, Cassano (1973: 211–213) documents one case each of [dʒ] in New Mexico and Argentina in the initial position in *ginger-beer*. In general, he claims that the development of the voiced alveopalatal affricate is independent of foreign influence and characteristic of several Hispanoamerican varieties of Spanish. He supposes that English loanwords might play a reinforcing role in the integration of the voiced alveopalatal affricate.

This study aims to investigate the different realizations of the phoneme in question and to analyze the factors that influence whether speakers imitate the English phoneme. For this study, it was of particular interest to know whether geographical proximity is still a relevant factor in the realization of Anglicisms, as suggested previously (Pustka 2021: 504; Rodríguez González 2017: 127; Oncins-Martínez 2009: 116). However, I question if through the interplay of modern communication tools (e.g., the Internet and social media) and American cultural imperialism (Gray 2007, Hamm & Smandych 2005) – the worldwide diffusion of culture, trademarks, values, media, and language from the US – the effect of greater geographic proximity might no longer play a crucial role.

2 Methodology

In the following section, I will describe the applied methodology: the corpus that served as database and the analysis I carried out, namely the audiovisual coding, the calculation of interrater reliability, and conducting the statistical analysis.

2.1 Corpus

2.1.1 Participants

For this study, the corpus designed for the dissertation project of the author (Bäumler 2023) served as the database. In the following, I will refer to the corpus as the *English Influence on Spanish (EIS)* corpus. Note that corpus phonology traditionally includes naturalistic and experimental data (Eychenne 2021: 324, Chaudron 2007: 736). The definition of *corpus* applied in this study is therefore not limited to spontaneous speech, since experimental data, such as reading a word list, are needed to enable comparison between speakers. The EIS corpus was designed to investigate possible factors influencing the adaptation of English loanwords

in Spanish. Therefore, it contains speech samples of 71 natives from two Spanish-speaking countries, namely Mexico and Spain. These two countries were chosen in order to test whether geographical proximity is still a relevant factor in the realization of Anglicisms (cf. Section 1.2). Informants were recruited primarily in the capital of each country, namely Madrid and Mexico City, but also in rural areas in the surroundings of the capitals, namely in Pedrezuela, Spain, that counts 5,892 habitants (INE 2019) and Chiconcuac, Mexico, which has 19,656 habitants (INEGI 2009). These two rural areas were chosen as they seem relatively comparable with regard to their social composition and their infrastructure as they are equally distant from the respective capital (both can be reached in 45 minutes by car). Due to the rather short distance between the urban and the rural areas, it can be assumed that the rural speakers use the same variety as the speakers in the respective capitals.

For reasons of research economy and to guarantee the speakers' confidence in the investigator (Pustka et al. 2018: 13; Milroy & Gordon 2003: 32), recruitment took place using a snowball technique, ensuring an equal distribution between females and males and across three different generations. This led to 25 speakers in each of the cities and to 10 (Pedrezuela) and 11 (Chiconcuac) speakers in the respective rural areas. The youngest speaker was 21 years old; the oldest was 92. The mean age was 48 years. The boxplot in Figure 1 shows the distribution of age.

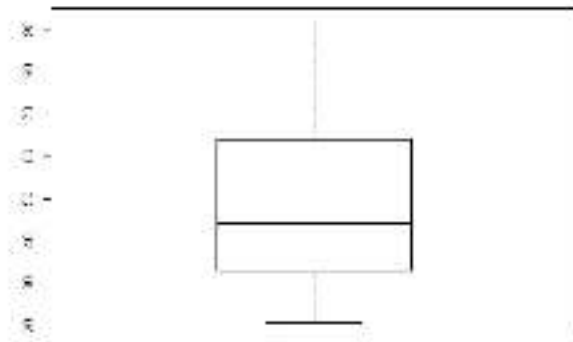


Figure 1: Distribution of age of the informants.

2.1.2 Task

In the framework of the project, speakers were asked to complete various tasks, including reading a word list. Since Anglicisms are only found rarely in spon-

taneous speech,⁴ this makes it possible to compare the pronunciation between many speakers. The word list consists of 102 items, among which there are 77 Anglicisms, the other 25 items being fillers and therefore not Anglicisms. Among the 77 Anglicisms are 29 Anglicisms that are morphologically and/or orthographically integrated or semantic loans (e.g., *administración*) and therefore also function as fillers, since they distract the informants. The word list provides seven Anglicisms that contain the voiced alveopalatal affricate in the source language (English), namely *jersey*, *ginger-ale*,⁵ *jet-lag*, *manager*, *jazz*, *gentleman*, and *digital*. Informants were asked to read the list aloud without familiarizing themselves with the words in advance. Recordings were carried out with the digital audio recorder ZOOM H4n Handy at a sampling rate of 44.1 kHz and a 16-bit depth. To further improve audio quality, the condenser microphone AKG C 520 was used.

2.1.3 Internal variables

Besides the speech material, the EIS corpus provides us with vast metadata. As previous research considers word frequency and age of the loanword as possible influencing factors in adaptation of loanwords, these variables were included in the analysis. Frequency values of each loanword were extracted from the corpora *American Spanish Web 2011* (esamTenTen11) and *European Spanish Web 2011* (eseuTenTen11), both accessible via the software Sketch Engine (Kilgarriff et al. 2014). Frequency values were divided by the total number of tokens included in the respective corpora to enable comparison between the two. In the analysis, frequency values from the American corpus were included only for speakers from Mexico, whereas values from the European corpus were only included for speakers from Spain in order to deal with differences in frequency between Mexico and Spain. Table 1 gives an overview of the frequency values for each word from the two corpora: *digital* is the most frequent, followed by *jazz* in both corpora. *Jet-lag* and *ginger-ale*, on the other hand, represent the less-frequent ones.

The age of the loanwords – also assumed as a relevant factor when it comes to the adaptation of loanwords – was defined as the time the loanword has existed in the target language (Spanish). To approximate this value, I extracted the year the loanword first appeared in a Spanish dictionary in the *Nuevo tesoro lexicográfico de la lengua Española* (NTLLE) (Real Academia Española 2019), which unifies an extensive selection of different dictionaries. I am aware of the limitations of

⁴In the interviews which were done in the framework of the project, only 141 occurrences of Anglicisms (78 types) were found in total among all speakers.

⁵Note that in this study, the analysis was limited to the phoneme in the first syllable to ensure that for each word, one phoneme was analyzed.

Table 1: Absolute and relative frequency of each loanword according to the American Spanish Web and European Spanish Web corpus.

	Freq. in European corpus	%	Freq. in American corpus	%
<i>digital</i>	263,013	130.10	652,311	87.26
<i>jazz</i>	37,182	18.39	96,871	12.30
<i>manager</i>	26,020	12.87	69,935	9.35
<i>jersey</i>	15,356	7.60	27,222	3.64
<i>gentleman</i>	858	0.42	1,622	0.22
<i>jet-lag</i>	196	0.10	230	0.03
<i>ginger-ale</i>	4	<0.01	12	<0.01

this method, since loanwords have to be more or less known by the community already before entering a dictionary. Nevertheless, this method was evaluated as the only possibility to obtain the data. Table 2 gives an overview of the year the different loanwords appeared in a dictionary in the NTLLE for the first time.

Table 2: Information on the first appearance of each loanword.

	Date of first appearance	Calculated age (years)	Dictionary of first appearance
<i>gentleman</i>	1895	124	<i>Zerolo</i>
<i>jersey</i>	1917	102	<i>Alemany y Bolufer</i>
<i>digital</i>	1983	36	<i>DRAE</i>
<i>jet-lag</i> ⁶	1984	35	<i>DRAE</i>
<i>manager</i>	1984	35	<i>DRAE</i>
<i>ginger-ale</i>	2014	5	<i>DRAE</i>
<i>jazz</i>	2014	5	<i>DRAE</i>

2.1.4 External variables

In addition, a large amount of extra-linguistic metadata on the informants was collected in the framework of the EIS project. In this analysis, I included the common sociolinguistic variables *age* and *sex*. To account for geographic differences between Mexico and Spain, the provenance of the speakers was also included. This makes it possible to test the hypothesis of previous research, which suggests a higher impact of English on Spanish in Hispanoamerica than in Spain (cf.

⁶First appearance as *jet*.

Section 1.1). Moreover, I included the informants' environments in the analysis to test the hypothesis that residents of rural areas are more commonly associated with conservative speech forms than those in cities (Sandøy 2014: 229; Chambers & Trudgill 1998: 46; Caballero Fernández-Rufete 1994: 134). However, this study aims to test whether this hypothesis is still true in a globalized world. Certainly, results in this regard have to be treated with caution, as Mexican immigrants from rural areas (Fussell 2004: 937) might have returned from the US and may use Anglicisms to a particular extent. Since contact with migrants who used to live in the US and therefore might incorporate non-adapted Anglicisms in their speech may also influence the realization of Anglicisms of speakers who have not lived in the US, the (binary) variable of speakers' contact with former migrants was also included in the study.

As only four speakers in the studied population speak an additional first language besides Spanish (in Mexico, one speaker indicates Totonaco and one Nahuatl as first languages; in Spain, one speaker indicates Valencian and one French as first languages), this variable was not included in the statistical analysis. However, the number of foreign languages spoken by the speakers was included.

Moreover, I included the variable *months spent abroad*, assuming that this variable might also influence the pronunciation of Anglicisms, since people who have lived in other (not only English-speaking) countries might have had more contact with foreign languages and cultures and might as a result show a more open attitude toward loanwords.

In addition, a variable called *Language Exposure and Affinity Score (LEAS)* that describes the affiliation of the speakers with the English language and American culture was included. To establish this score, speakers indicated on a five-point Likert scale how often they are in contact with the English language and rated their attitude toward the language and American⁷ culture. They were asked how often they read in English, watch films in English, make use of the Internet in English, and are in contact with the English language in their daily lives. They also rated their attitude toward the English language and American culture and indicated how many times they had visited an English-speaking country.⁸ In consequence, up to four points were added to the score per question. Furthermore, I included the variables *education* and *English level* in the analysis. Note that the latter – just as the *LEAS* – is based on self-evaluation, since speakers

⁷See Section 1.2 for an explanation as to why I chose to concentrate on American culture.

⁸Since this last information was not measured by means of a Likert scale, it was included in the score as follows: no points were added if the informant had never been to an English-speaking country, two points if they had been to one English-speaking country, and four points if they had been to two or more English-speaking countries.

rated their competence as *zero*, *basic*, *intermediate*, or *advanced*. This approach naturally represents a limitation, but was considered the most effective way to obtain the data.

2.2 Analysis

In the following, I will describe the different steps of data analysis, namely audio-visual coding, the calculation of interrater reliability, and the statistical analysis.

2.2.1 Audiovisual coding

To enable the analysis, data were first processed in PRAAT (Boersma & Weenink 2021) before being segmented and orthographically transcribed. Subsequently, three linguists with phonetic background (the author of the article, one MA student, and another person who had a master's degree in linguistics) transcribed all 568 Anglicisms phonetically. Two of the transcribers speak German as their first language and English and Spanish as foreign languages. The other speaks English as their first language. All of them studied Spanish and/or English linguistics and therefore display extended knowledge in either Spanish or English linguistics or both.

The three coders transcribed the data individually before the transcriptions were compared to grant objectivity. In doing so, each coder transcribed the data audiovisually, aligning the phonetic transcription with the acoustic signal.

2.2.2 Interrater reliability

To ensure the reliability of the analysis, the transcription of the phoneme in question by the three coders was compared and a Fleiss' kappa (Fleiss et al. 2003) was calculated. For calculation, the R package DescTools (Signorell 2021) was used. Fleiss' kappa was elaborated for the calculation of interrater reliability of more than two coders. Landis & Koch (1977) proposed the interpretation of kappa values as shown in Table 3.

Landis & Koch (1977) mention in their paper that these divisions are clearly arbitrary and should only be seen as useful benchmarks. The arbitrary nature of such benchmarks has been criticized, with observers pointing out that the effects of prevalence and bias on kappa must be considered when judging its magnitude (Sim & Wright 2005: 264). It has to be taken into account that kappa can be influenced by the number of categories in the measurement scale: The more categories available, the greater the potential for disagreement among raters, resulting in a lower kappa with many categories than with few. With these limitations

Table 3: Interpretation of kappa value (Landis & Koch 1977: 165).

Kappa statistic	Interpretation
<0.00	Poor agreement
0.00–0.20	Slight agreement
0.21–0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–1.00	Almost perfect agreement

in mind, the interpretation provided by Landis & Koch (1977) will only be used as orientation in this study. As phonetic transcription opens up many different possible categories, it can be assumed that calculated kappa values in phonetic studies could be even higher if categories were further reduced.

2.2.3 Statistical analysis

After calculating the interrater reliability, the transcriptions in which two or all three transcribers matched were selected. In cases where the transcriptions of all three transcribers gave different results, the author relistened to the phoneme and decided on one solution.

Data were then processed in R (R Core Team 2021). Analysis of the data took place in two steps. First, the transcribed data were analyzed descriptively to get an overview of the used phoneme adaptations and their respective frequencies. In a second step, statistical modelling was applied to infer the effects of the variables on the realization of the phoneme in question. For this, the R package lme4 (Bates et al. 2015) for fitting linear mixed-effects models was used. For statistical modelling, dichotomous coding of the transcribed data took place.

Therefore, I coded all realizations of grapheme-phoneme correspondences as *Spanish*, while all imitations of the English model were coded as *English*. The first group therefore included realizations as [ˈmanaxer] or [dixiˈtal]. In the latter group, I included all cases where no grapheme-phoneme correspondences were realized and therefore imitation of the English model can be assumed. Therefore, cases such as [ˈmæniðʒə] that were pronounced source language-like (*phoneme importation*), and cases like [ˈmanajer], which represent instances where the speakers imitate the English model but replace it with a phoneme that is native to their L1 (*phoneme adaptation*), were included in this group. Figure 2 gives an overview of the dichotomous coding.

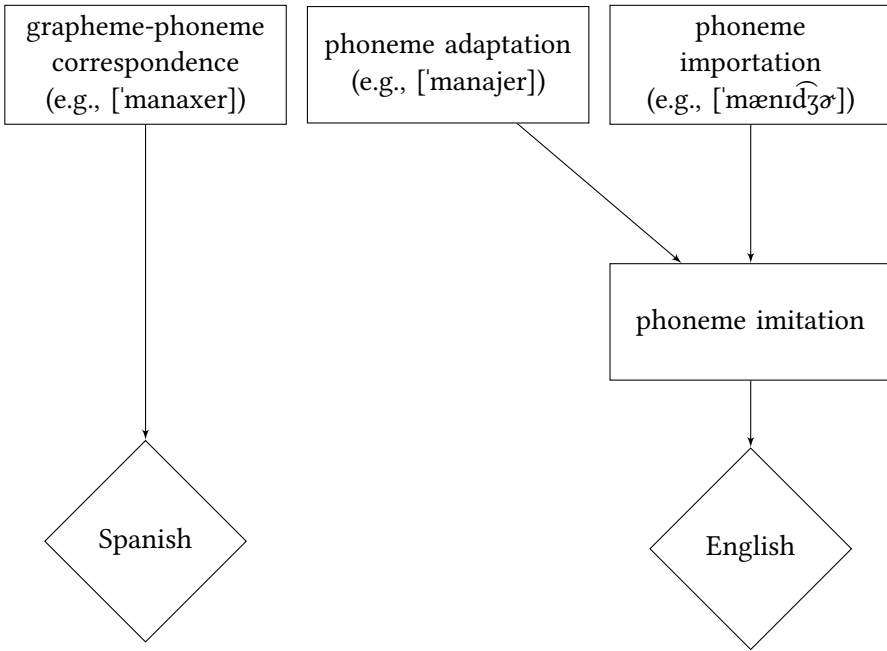


Figure 2: Dichotomous coding of the variable.

Prior to the calculation of the generalized mixed model, a directed acyclic graph (DAG) was developed (Figure 3) using the R package dagitty (Textor et al. 2016). In this graph, all assumed causal relationships between the dependent variable and independent variables (cf. Sections 2.1.3 and 2.1.4) and between each of the independent variables are illustrated, resulting in a very dense network of potential causal implications. As a next step, I aimed to identify among these relationships only those who actually (i.e., in terms of statistical significance) and directly (i.e., not confounding by another independent variable) infer the pronunciation. For this purpose, the generalized linear mixed model was built, employing a bottom-up and top-down approach.

I did not include the phoneme realizations of *jersey* in the statistical modelling, since it can be assumed that Mexicans and Spaniards do not refer to the same item with this lexeme (see Section 4 for further explanation). Before setting up the model, the variable *word frequency* was log-transformed to account for the differences between the different word frequencies. In addition, I scaled this variable and the other continuous variables to ensure that all continuous variables are on the same scale. For the categorical variables *area*, *country of origin*, and *sex*, the reference level was set as follows: city (*area*), Spain (*country of origin*),

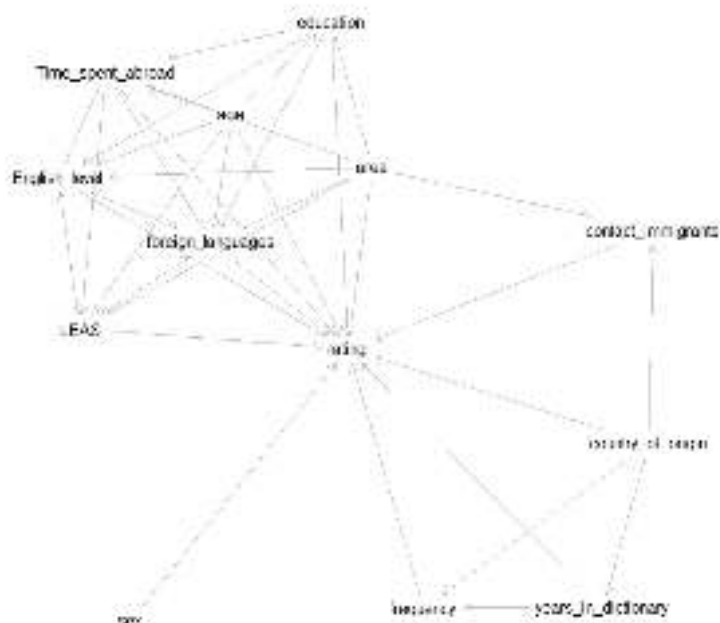


Figure 3: Directed acyclic graph of potential predictors of English phoneme realization.

female (*sex*). Further, I added the random effects *speaker*, *word*, *following vowel*, and *position* to the model to account for group-level variations. The random effect *speaker* therefore accounts for variations in pronunciation between the informants, whereas *word* accounts for any word-level differences in pronunciation. The random effect *following vowel*, on the other hand, accounts for differences in pronunciation due to the quality of the following vowel, which can be either a low, medium, or high vowel (/a/, /e/, or /i/, respectively). With the inclusion of this random effect, I therefore aim to expose whether the phonetic environment influences the realization of the phoneme. Finally, the random effect *position* accounts for variations in pronunciation between the different positions in which we find the phoneme in question, namely word initially and word internally (in both cases in onset position).

3 Results

3.1 Fleiss' kappa

The calculation of Fleiss' kappa revealed an interrater reliability of 0.71. According to Landis & Koch (1977), this represents "substantial agreement." Since phonetic transcriptions naturally show many different categories, the Fleiss' kappa values indicate a sufficient interrater reliability.

3.2 Realization of the phoneme in question

For the phoneme in question, all in all, nine different realizations among the speakers were found. Table 4 gives an overview of the different realizations of <ge>, <gi>, and <j> in all words among all speakers and their respective frequencies.

Table 4: Realizations of the phoneme in question in all words among all speakers.

Realization	Frequency in %
$\widehat{d}ʒ$	32.2
$\widehat{tʃ}$	5.8
ʃ	7.0
j	16.3
g	2.0
ç	9.3
x	16.3
χ	4.8
h	0.4
other	3.6
error	2.2

The analysis shows that in most cases (32 percent), speakers import the English phoneme. In the respective corpus, importation of $\widehat{d}ʒ$ is therefore not rare. In total, speakers realize grapheme-phoneme correspondences in 30.8 percent of cases. Note that the different allophones [x] (16.3 percent), [ç] (9.3 percent), and [χ] (4.8 percent) constitute this group. While the latter, which represents a posterior pronunciation of [x], is typical of the center and north of the Iberian Peninsular (Hualde 2014: 149), the voiceless palatal fricative [ç] appears in the speech

of Mexicans in front of the fronted vowels [i] and [u], as in *digital*. Even though the realization of the glottal [h] does not represent a standard Spanish grapheme-phoneme correspondence, it was included in this group, since it can more likely be interpreted as allophonic variation than as imitation of the English model.

In between the importation of the foreign phoneme and the realization of the Spanish grapheme-phoneme correspondence, different allophones that aim to imitate the English phoneme can be observed. Regarding the place of articulation, they range from postalveolar ([tʃ], 5.8 percent, place of articulation shared with the English phoneme) to palatal pronunciations ([ʃ], 7 percent and [j], 16.3 percent) and velar realizations ([g], 2 percent). Further, regarding the manner of articulation, speakers realize affricate ([tʃ]), plosive ([ʃ] and [g]), and approximant ([j]) pronunciation. Regarding the voicing, 25.3% of the speakers realize a voiced phone, whereas 5.8% realize an unvoiced alternative.

The analysis of each word individually shows that the realization varies considerably between words. Table 5 gives an overview of the frequencies in percent of the realizations for each word.

Table 5: Realizations of the phoneme in question in the different words among all speakers in percent.

Realization	<i>ginger-ale</i>	<i>jet-lag</i>	<i>gentleman</i>	<i>jazz</i>	<i>manager</i>	<i>jersey</i>	<i>digital</i>
ḋʒ	60.6	54.9	36.6	36.6	19.7	16.9	0
tʃ	5.6	9.9	5.6	5.6	8.5	5.6	0
ʃ	0	8.5	7.0	19.7	1.4	12.7	0
j	5.6	8.5	12.7	25.4	52.1	9.9	0
g	4.2	0	7.0	0	2.8	0	0
ç	2.8	2.8	1.4	0	0	0	57.7
x	12.7	4.2	15.5	8.5	8.5	36.6	28.2
χ	1.4	4.2	1.4	2.8	0	9.9	14.1
h	0	0	0	0	0	2.8	0
other	4.2	4.2	8.5	0	7.0	1.4	0
error	2.8	2.8	4.2	1.4	0	4.2	0

The results show that *digital* displays the most cases of realization of grapheme-phoneme correspondences, followed by *jersey*, with 100 percent and 49.3 percent, respectively. In the first case, no imitation of the English phoneme was observed. In all other words, cases of phoneme imitation outweigh those of grapheme-

phoneme correspondences. That is, speakers are more likely to imitate the English model than to realize grapheme-phoneme correspondences. Within this group, where realizations of phonic imitation outweigh realizations of grapheme-phoneme correspondences, differences regarding the frequencies of phonic importation can be observed. *Ginger-ale* and *jet-lag* show the most cases of phonic importation, with 60.6 percent and 54.9 percent, respectively. *Gentleman* and *jazz* both show phoneme importation in 36.6 percent of cases. When pronouncing *manager*, on the other hand, only 19.7 percent of the speakers import the English phoneme. Nonetheless, the latter shows a strikingly elevated number of cases of the palatal approximant, with 52.1 percent of the speakers adapting the foreign sound to the respective realization.

3.3 Effects influencing the imitation of the English phoneme

To infer the effect of the selected variables on the phonic realization, a general linear mixed model was used, as described in Section 2.2.3. To summarize, for this model, the realization of Spanish grapheme-phoneme correspondence (or the imitation of the English realization) was set as the dependent variable. The results of this analysis are shown in the forest plot (Figure 4). Plots were created using the package *ggplot2* (Wickham 2016).

In this plot, for every independent variable, the effect on the realization is depicted by a horizontal bar, corresponding to the 95 percent (thin bar) and the 50 percent (thicker bar) credibility interval of the predicted effect. The x-axis represents the probability of the type of loanword realization; whereas positive values correspond to a higher probability of an imitation of the English phoneme, negative values correspond to a higher probability of realization of grapheme-phoneme correspondence. The mean effect of each predictor is marked by a dot. When interpreting these plots, it is important to consider that if the 95 percent credibility interval overlaps the zero-line (dashed vertical line), the model does not yield a distinct direction of the effect for this particular predictor, and therefore should be seen as statistically insignificant.

The model reveals a significant effect of the *LEAS* on the realization of loanwords. The higher the exposure and affinity to the English language and American culture (assessed by *LEAS*), the more likely it was the English phoneme was imitated, and the less likely the realization was grapheme-phoneme correspondent. Moreover, the plot shows an effect of word frequency on the loanword: the more frequent the loanword, the higher the probability that speakers pronounce the phoneme in question grapheme-phoneme correspondent. Furthermore, the type of area also impacted the realization, as being situated in cities or urban

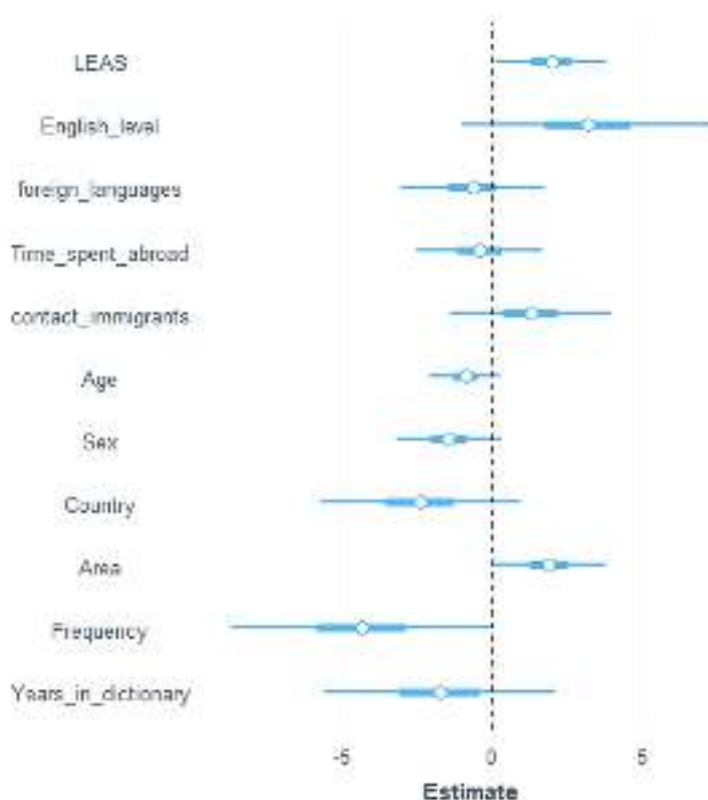


Figure 4: Forest plot of predictors of English phoneme realization.

areas was associated with a higher probability of an imitation of the English model. Yet further analysis showed that these two predictors interact with *LEAS*, thus distorting their effects. For this reason, it must be assumed that the effects observed in those measured are in fact not there.

Similarly, the analysis revealed that *age*, *foreign languages*, and *education* are confounds of *LEAS*. The effects found for these variables when considered separately in the model vanished in the process of the bottom-up analysis, taking into account interactions and confounding between the different independent variables. All these initially observed effects vanished with the integration of *LEAS* into the model, therefore indicating that the abovementioned variables may also influence the realization, but only indirectly, since they are collinear with *LEAS* and data are better explained by the latter. For instance, when analyzing *age* separately, it may seem as if a direct influence on the realization of loanwords exists;

however, this influence is actually explained by differences in *LEAS*, which decreases with increasing age. Respectively, we can assume a causal relationship between *education* and the realization insofar as speakers with a higher level of education may have higher scores in English media consumption. The variable *education* was not included in the final model since it interacts with *LEAS* and influences its values in the final model.

The final model does not reveal any direct influence of *English level* on pronunciation. However, bottom-up analysis showed that when only including *LEAS* and *English level* in the model, the effects both variables show in univariate analysis persist, although the effect of *LEAS* is higher than the effect of *English level*. It is therefore possible to assume a bidirectional causal relationship between the two variables, not only as speakers with higher language skills may have higher scores in English media consumption, but also since English media consumption may also improve language skills.

Furthermore, in the data analyzed, there is no indication of an effect of the country of origin on the preferred type of loanword realization – no significant difference between Spaniards and Mexicans was found. In addition, contact with former migrants who used to live in the US and the age of the loanwords did not influence the pronunciation of the Anglicisms.

The analysis of the random effects yielded no remarkable differences between the speakers. Standard deviation between the speakers is 1.69. The analysis of the random effect *word*, on the other hand, revealed greater differences in the realization between words. The forest plot in Figure 5 shows the group-level intercepts per word. The x-axis indicates the realization in relation to the population mean, while negative values indicate the realization of grapheme-phoneme correspondences and higher values the imitation of the English sound. However, these posterior intercepts have to be interpreted with caution, since they are drawn to the mean by the *pooling* effect during regression modelling, which prevents the model from overfitting. Figure 5 shows a certain tendency toward the probability of grapheme-phoneme correspondent realization being higher for *jet-lag* and *digital*, whereas the English phoneme in *manager* is more likely to be imitated. However, the depicted confidence intervals underline the uncertainty of the results and do not allow for specific interpretation.

Moreover, the analysis of the random effect *following vowel* revealed that in cases where the phoneme in question is followed by /i/, speakers are more likely to realize grapheme-phoneme correspondences than when followed by /a/ or /e/, as shown in Figure 6.

The analysis of the random effects showed that *following vowel* influences *word*. As such, it is not simply a case of the the different words showing dif-

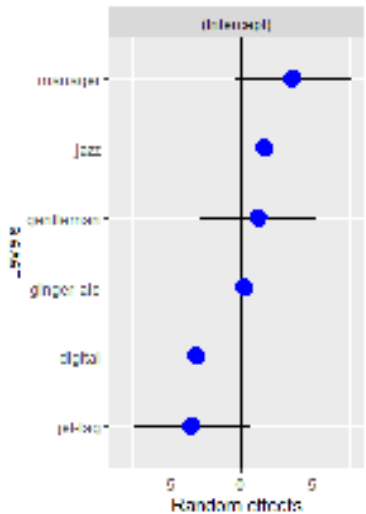


Figure 5: Forest plot of group-level intercepts per word.

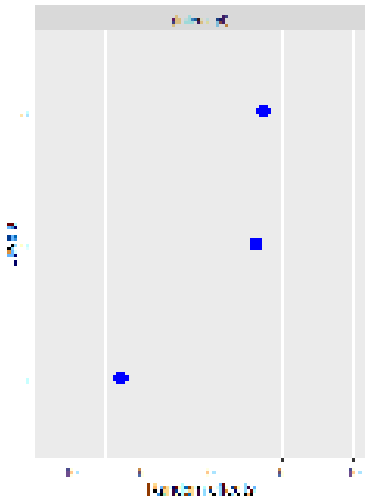


Figure 6: Forest plot of group-level intercepts per following vowel.

ferent realizations; the following vowels are also an important factor. Since the analysis of the random effect *position*, meaning if the phoneme in question is located word initially or internally, did not yield any remarkable differences, I did not include it in the final model.

4 Discussion

In this section, I will discuss the results described above; I will begin with the different phones, before discussing the results of the statistical modelling.

First of all, the analysis showed many cases of phoneme importation that have barely been reported in previous literature (cf. Section 1.2). Since native Spanish speakers can nowadays be constantly exposed to the model sounds via English media, it can be assumed that this promotes the importation of foreign phonemes. Additionally, since some Spanish varieties, such as those found in Paraguay and some areas of Ecuador and Argentina (Real Academia Española & Asociación de Academias de la Lengua Española 2011: 226), display the voiced alveopalatal affricate [dʒ] in the realization of <ll>, <y>, and <(h)i>, research on these varieties is needed to analyze whether the realization of the respective phoneme in English loanwords is even higher in the respective geographical areas.

Moreover, in cases where speakers imitate the foreign sound, but do not realize the voiced postalveolar affricate [dʒ], the analysis revealed that speakers realize phones of different places of articulation. Despite the realization of [tʃ], in all other cases, speakers shift the articulation of the phone further to the back. Since only 5.8 percent of the speakers imitate the English phoneme with the voiceless postalveolar affricate [tʃ], this leads to the assumption that the place of articulation is not decisive in imitation. The palatal pronunciation, which is realized by 23.3 percent and therefore by the majority of the speakers who imitate the English model but do not pronounce [dʒ], is nonetheless the closest to the postalveolar articulation point. Regarding the manner of articulation, results showed that only 5.8 percent of the speakers realize an affricate (other than the English variant), which leads to the assumption that the manner of articulation is not striking when it comes to the perception and subsequently to the realization of the phoneme. Concerning the voicing of the phone, the finding that 25.3 percent of the speakers realize a voiced phone (other than [dʒ]) in contrast to 5.8 percent who realize a unvoiced one suggests that the voicing plays a major role in the perception and realization of the English model. It obviously plays such a big role that speakers would rather realize the voiced palatal approximant [j], which is posterior in means of place of articulation, than the unvoiced variant of

the English affricate, which not only shares the same place of articulation, but also the same manner of articulation.

These results suggest that the voicing of the English affricate is the most striking feature when Spanish natives perceive the foreign sound. This has also been shown by Gómez Capuz (2001: 29). It can be assumed that it is more important in the imitation process than the place of articulation, even though the latter plays a major role in comparison to the manner of articulation, which seems rather unimportant in the selection of the closest sound in the native system.

Furthermore, the results show varying frequencies of the presented phones in the different words. This finding underlines that “each word has its own history.”⁹ Regarding the Anglicism *digital*, the first appearance used in a technical way dates from 1983, but the word was already used in Spanish in the nineteenth century to name a plant (> lat. DIGITALIS). Before the loanword entered the Spanish language via English, it was therefore already established and is not perceived as an Anglicism. It can consequently rather be seen as a semantic loan. Hence, we can assume that the new meaning coming from the Anglicism was linked to the already existing Spanish pronunciation. Concerning the Anglicism *jersey*, a considerable difference between Mexicans and Spaniards can be found; while 88 percent of the Spaniards realize grapheme-phoneme correspondences, only 12 percent of Mexicans do so. This geographical difference can be explained as the Anglicism is common in Spain to name the respective garment. It first appeared in a dictionary with this meaning in 1917, while it was already known in the Spanish lexicon for a much longer period (at least since 1705) as the denomination of the isle in UK waters. This suggests that the new meaning was linked to the already existing Spanish pronunciation of the English phoneme. Since Mexicans are more likely to use *suéter* for the respective garment, they do not link the same object to the Anglicism as Spaniards. One can assume that Mexicans more likely refer to the respective garment used in American football. Remember that *jersey* was not included in the statistical analysis, since speakers in each country use the term to refer to different items.

As shown above, *ginger-ale* and *jet-lag* showed the most cases of phonic importation, with 60.6 percent and 54.9 percent respectively. The analysis of the respective entries in the NTLLE shows that these Anglicisms range among the “younger” ones, with *ginger-ale* being included in 2014 and *jet-lag* in 1984. Further, it was shown that *gentleman* and *jazz* both showed phoneme importation

⁹According to Pustka (2007: 40), the quote “chaque mot a son histoire” can be assigned to Gilliéron: “Cette phrase ne se trouve pas explicitement dans l’œuvre de Gilliéron, mais lui est attribuée par de nombreux auteurs” (Pustka 2007: 40).

in 36.6 percent of cases. The analysis of the NTLLE showed that *gentleman* was first included in a dictionary in 1895, while the latter was not included until 2014. We therefore find an older and a very young loanword in this group. The analysis of *manager* showed lower rates of phonemic importation, while revealing high rates of adaptation to the palatal approximant. It can be assumed that the intervocalic position of the phone favors the articulation of the approximant instead of the affricate, which is more likely to be produced word initially. The first dictionary entry of *manager* dates from 1984, the same year as for *jet-lag*, which shows high rates of phonemic importation. At first glance, these findings already suggest that the date of entrance alone cannot be a reliable factor for predicting the adaptation of the foreign phoneme. Moreover, it has to be kept in mind that the year of first incorporation in a dictionary can only approximate the time, the loanword exists in the recipient language: First, loanwords are often already known for a mostly indeterminable time period in the speech community before they get incorporated in the dictionary. Second, the decision of incorporation of new items can also be influenced by purist tendencies of the language academies.

Instead, it can be assumed that the way of entrance (*eye loans* versus *ear loans*) and transparent orthography play a non-negligible role in the adaptation of English loanwords in Spanish. Even though some researchers claim that orthography only plays a limited role in loanword phonology (LaCharité & Paradis 2005: 241), others underline the importance of the graphic system (Pustka 2021: 505; Vendelin & Peperkamp 2006; Peperkamp & Dupoux 2003: 369): Since Spanish has a very transparent orthography, it can favor the realization of Spanish grapheme-phoneme correspondences. This analysis clearly shows many cases of grapheme-phoneme correspondences. However, the limitation of the chosen method, namely the reading of a word list, has to be considered. An influence of the visual code cannot be excluded. Therefore, future studies should also investigate the phonic realization of Spanish Anglicisms in more naturalistic data, such as interviews. However, comparability between the data is not as high as it is in the case of reading of a list, since Anglicisms occur only rarely in spontaneous speech (cf. Section 2.1.2). Moreover, picture naming tasks or discourse completion tasks, where at least the graphic code is not present, might be a fruitful possibility to analyze phonic adaptation.

Concerning the effects of the different variables on the realization of the phoneme, the generalized linear mixed model revealed a significant effect of *LEAS*. Further, the apparent effect of *area* and *frequency* might in fact only be the result of an interaction with *LEAS*. The effects of the variables *age*, *education*, *English level*, *foreign languages*, *area*, and *frequency* found when integrated separately into the model are partially supported by previous literature. Rodríguez González (2017:

127) and Gómez Capuz (2001: 42) suggest that speakers with a higher level of competence in English tend to imitate the English model. Rodríguez González (2017: 102, 123) concludes the same for speakers with a higher level of education, while Rodríguez González (2018: 354) states that more frequent loanwords are more likely to undergo phonic adaptation.

However, as this study examines this phenomenon by analyzing comparative data of various speakers in connection with sociogeographical variables, it allows for statistical modelling of multiple predictors. The statistical model suggests that the abovementioned predictors, also discussed in previous research, do not directly explain the change of preference regarding the realization of the English phoneme, but that these predictors are rather explained by the collinear predictor of English exposure and affinity (i.e., *LEAS*). In other words, the actual decisive predictor that drives the preference of imitation of the English model is the exposure and affinity to the English language and American culture. However, the limitation of the chosen method – the reading of a word list – has to be considered: speakers might adopt the pronunciation of the Anglicisms in different context situations.

Moreover, the statistical analysis did not reveal significant differences between speakers in Mexico and Spain. The assumption expressed in previous literature (Pustka 2021: 504; Rodríguez González 2017: 127; Oncins-Martínez 2009: 116) that English has a greater influence in Hispanoamerica than in Spain could not be confirmed by these data. Therefore, at least the analyzed data suggest that in a globalized world, geographic proximity to the US might not be as relevant, since speakers can expose themselves to the English language via easily accessible media. Besides written input, spoken English is available via streaming services as Netflix or video blogs shared via social media – just to mention some. However, more research is needed to compare the realization in different Spanish varieties. Moreover, the analysis did not reveal a significant influence of speakers' contact with people who have previously lived in the US and therefore may use a higher number of Anglicisms. In Mexico, in particular, one might assume that such contact situations could influence the pronunciation of speakers who have not lived in the US. Since the data analyzed in this study did not reveal such a correlation, it is reasonable to assume that contact with English in general, which does not necessarily have to happen through returning migrants, is a more crucial predictor. In addition, the analysis did not reveal any significant influence of the internal variable *years in dictionary* on the realization of the Anglicisms. The assumption made in previous literature (Gómez Capuz 2001: 21) that older Anglicisms are less likely to be pronounced source language-like, cannot be confirmed with these data.

The finding of the descriptive analysis described above – that the pronunciation of the phoneme in question varies significantly between words – can be confirmed by the statistical analysis. However, the analysis of the random effects shows that it is not only the history of the word itself that influences the realization of the phoneme. The realization also depends on the quality of the following vowel (low/medium/high), indicating that when followed by the high vowel /i/, the phoneme is more likely to be realized grapheme-phoneme correspondent. Future research could also incorporate Anglicisms where /d͡ʒ/ is followed by /u/ and /o/ to test these occurrences. In this sample, the position of the phoneme (word initially versus internally) did not influence the realization when it comes to the difference between imitation and realization of grapheme-phoneme correspondences. However, within the group of imitation, it can be assumed that word initially, more speakers opt to import the phoneme than in intervocalic position (as described above). This might be due to the fact that the consonant is more salient in word-initial position. Nevertheless, more data in intervocalic position are needed.

The results of the statistical modelling show that when multiple potential predictors are included in complex statistical models, effects that would eventually remain hidden by other types of analysis can be presented. Therefore, and due to the fact that complex variables such as *LEAS* have not been used in previous research, to the best of my knowledge, future research on the realization of Spanish Anglicisms should be based on comparative corpora that include sufficient sociodemographic data to allow for statistical modelling.

5 Conclusion

Concerning the distribution of the different realizations, the analysis of the corpus data revealed that in 32.2 percent of the analyzed cases, speakers imported the English phoneme /d͡ʒ/ when realizing the respective Anglicisms. The importation of the voiced alveopalatal affricate /d͡ʒ/, which is not native to the Spanish phoneme system, is therefore not rare. Besides phoneme importation, in other cases of phoneme imitation, speakers rather realize a voiced palatal sound ([j] or [ɟ]) than the unvoiced alveopalatal affricate ([t͡ʃ]). The voicing might therefore play a major role in the perception of the English phoneme. The realization of the grapheme-phoneme correspondence also plays an important role in the realization of the English phoneme, which may be due to the transparent nature of Spanish orthography.

Moreover, concerning the sociogeographic variables influencing the realization of the English phoneme, it has been shown that primarily speakers' expo-

sure and affinity to the English language and American culture was decisive; the higher one speakers' *LEAS*, the more likely imitation of the English model was observed. In contrast, at least the data studied here suggest that in a globalized world, geographic proximity may – at least regarding the regions compared in this study – no longer be a relevant factor.

Abbreviations

EIS	English Influence on Spanish
DAG	directed acyclic graph
LEAS	Language Exposure and Affinity Score
NTLLE	Nuevo tesoro lexicográfico de la lengua Española

Acknowledgements

I want to kindly thank Frederik Hartmann (University of North Texas) for providing advice and expertise for statistical analysis and comments on the manuscript, and Elissa Pustka (University of Vienna), the editors and two blind reviewers for a critical reading of this paper. All remaining errors are my own. Furthermore, I want to thank Luke Green and Martin Baumgartner for conducting the additional phonetic transcriptions and of course all the speakers who patiently participated in the study. The research stays in Mexico and Spain were funded by grants from the University of Vienna and the Rosita Schjerve-Rindler Fonds.

References

- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.
- Bäumler, Linda. 2023. *Anglicisms in Mexico and Spain in times of globalization: Use, pronunciation and perception*. University of Vienna. (Doctoral dissertation).
- Boersma, Paul & David Weenink. 2021. *Praat: Doing phonetics by computer [Computer program]*. www.praat.org (18 December, 2021).
- Briz, Antonio & Grupo Val.Es.Co. 1995. *La conversación coloquial. Materiales para su estudio*. Valencia: Universidad de Valencia.
- Caballero Fernández-Rufete, Salvador. 1994. Anglicismos: Un trabajo de campo. *Revista de filología inglesa* (18). 143–160.
- Cabanillas, Isabel De La Cruz & Cristina Tejedor Martinez. 2012. Email or correo electrónico? Anglicisms in Spanish. *Revista Española de Lingüística Aplicada* 25(1). 95–118.

- Cassano, Paul Vincent. 1973. The influence of American English on the phonology of American Spanish. *Orbis* 22(1). 201–214.
- Chambers, J. K. & Peter Trudgill. 1998. *Dialectology*. 2nd edn. Cambridge: Cambridge University Press.
- Chaudron, Craig. 2007. Data collection in SLA research. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition*, 762–828. Malden, MA: Blackwell.
- Eychenne, Julien. 2021. Corpus phonology. In Christoph Gabriel, Randall Gess & Trudel Meisenburg (eds.), *Manual of romance phonetics and phonology*, 319–341. Berlin/Boston: De Gruyter.
- Fleiss, Joseph L., Bruce Levin & Myunghee Cho Paik. 2003. *Statistical methods for rates and proportions*. 3rd edn. Hoboken, NJ: Wiley.
- Fussell, Elizabeth. 2004. Sources of Mexico's migration stream: Rural, urban, and border migrants to the United States. *Social Forces* 82(3). 937–967.
- Gómez Capuz, Juan. 2001. Estrategias de integración fónica de los anglicismos en un corpus de español hablado: Asimilación, compromiso y efectos estructurales. *Estudios de Lingüística* 15. 5–86.
- Gray, Jonathan. 2007. Imagining America: *The Simpsons* go global. *Popular Communication* 5(2). 129–148.
- Hamm, Bernd & Russell Smandych (eds.). 2005. *Cultural imperialism. Essays on the political economy of cultural domination*. Toronto: University of Toronto Press.
- Haspelmath, Martin. 2009. Lexical borrowing: Concepts and issues. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 35–54. Berlin/New York: De Gruyter Mouton.
- Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language* 26(2). 210–231.
- Hualde, José Ignacio. 2014. *Los sonidos del español*. Cambridge: Cambridge University Press.
- INE, Instituto Nacional de Estadística. 2019. *Cifras oficiales de población resultantes de la revisión del Padrón municipal a 1 de enero*. <https://www.ine.es/jaxiT3/Datos.htm?t=2881#!tabs-tabla> (17 November, 2021).
- INEGI, Instituto Nacional de Estadística Geografía e Informática. 2009. *Prontuario de información geográfica municipal de los Estados Unidos Mexicanos*. https://www.inegi.org.mx/contenidos/app/mexicocifras/datos_geograficos/15/15030.pdf (17 November, 2021).
- Kang, Yoonjung. 2011. Loanword phonology. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 2258–2282. Malden, MA: Wiley-Blackwell.

- Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý & Vít Suchomel. 2014. The Sketch Engine: Ten years on. *Lexicography* 1. 7–36.
- LaCharité, Darlene & Carole Paradis. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36(2). 223–258.
- Landis, J. Richard & Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33(1). 159–174.
- Lorenzo, Emilio. 1997. Tratamiento del vocalismo inglés en español. *Estudios ingleses de la Universidad Complutense* 5. 81–90.
- Martínez Celdrán, Eugenio. 2015. Naturaleza fonética de la consonante ‘ye’ en español. *Normas* 5. 117–131.
- Meisenburg, Trudel. 1992. Graphische und phonische Integration von Fremdwörtern am Beispiel des Spanischen. *Zeitschrift für Sprachwissenschaft* 11(1). 47–67.
- Milroy, Lesley & Matthew Gordon. 2003. *Sociolinguistics: Method and interpretation*. Chichester: Wiley-Blackwell.
- Oncins-Martínez, José. 2009. Towards a corpus-based analysis of Anglicism in Spanish: A case study. *International Journal of English Studies* 1. 115–132.
- Peperkamp, Sharon & Emmanuel Dupoux. 2003. Reinterpreting loanword adaptations: The role of perception. In *Proceedings of the 15th international congress of phonetic sciences*, 367–370.
- Pratt, Chris. 1980. *El anglicismo en el español peninsular contemporáneo*. Madrid: Gredos.
- Pustka, Elissa. 2007. *Phonologie et variétés en contact. Aveyronnais et Guadeloupéens à Paris*. Tübingen: Narr.
- Pustka, Elissa. 2021. Loanword phonology in Romance. In Christoph Gabriel, Randall Gess & Trudel Meisenburg (eds.), *Manual of romance phonetics and phonology*, 503–528. Berlin/Boston: De Gruyter.
- Pustka, Elissa, Christoph Gabriel, Trudel Meisenburg, Monja Burkard & Kristina Dziallas. 2018. (Inter-)Fonología del Español Contemporáneo (I)FEC: Metodología de un programa de investigación para la fonología de corpus. *Loquens* 5(1). 1–16.
- R Core Team. 2021. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Real Academia Española. 2019. *Nuevo tesoro lexicográfico (NTLLE)*. <https://apps.rae.es/ntlle/SrvltGUISalirNtlle>.

- Real Academia Española & Asociación de Academias de la Lengua Española. 2011. *Nueva gramática de la lengua española: Fonética y fonología*. Madrid: Espasa.
- Rodríguez González, Félix. 2017. Variaciones fonológicas en el uso del anglicismo: Panorama y revisión crítica. *Revista Española de Lingüística* 47(2). 99–134.
- Rodríguez González, Félix. 2018. Aspectos ortográficos del anglicismo. *Lebende Sprachen* 63(2). 350–373.
- Sandøy, Helge. 2014. Linguistic globalization: Experiences from the Nordic laboratory. In Eline Zenner & Gitte Kristiansen (eds.), *New perspectives on lexical borrowing: Onomasiological, methodological and phraseological innovations*, 225–249. Berlin/Boston: De Gruyter Mouton.
- Schweickard, Wolfgang. 1991. Anglizismen im Spanischen. *Terminologie et Traduction* 1. 75–86.
- Signorell, Andri et mult. al. 2021. *DescTools: Tools for descriptive statistics*. R package version 0.99.44. <https://cran.r-project.org/package=DescTools>.
- Sim, Julius & Chris C. Wright. 2005. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy* 85(3). 257–269.
- Textor, Johannes, Benito van der Zander, Mark S Gilthorpe, Maciej Liśkiewicz & George TH Ellison. 2016. Robust causal inference using directed acyclic graphs: The R package ‘dagitty’. *International Journal of Epidemiology* 45(6). 1887–1894.
- Vendelin, Inga & Sharon Peperkamp. 2006. The influence of orthography on loan-word adaptations. *Lingua* 116(7). 996–1007.
- Wickham, Hadley. 2016. *Ggplot2: Elegant graphics for data analysis*. New York: Springer. <https://ggplot2.tidyverse.org>.

Chapter 11

Imitating the Arabic model: The case of valency-increasing operations in the Old Spanish translation of the Arabic *Kalīla wa-Dimna*

Hans-Jörg Döhla

University of Tübingen

Learned language contact, explicitly the translation of Arabic works into Old Spanish, offers quite a range of possibilities for the detailed study of morphosyntactical and semantic structures in the target language of Old Spanish. This chapter investigates causative constructions in Old Spanish using the medieval translation (ca. 1251 AD) of the Arabic *Kalīla wa-Dimna* as its primary source. In this context, three main formal strategies can be identified: (1) zero morphology or conversion; (2) denominal derivations of the patterns *a-N-ar* and *en-N-ar*; and (3) analytic constructions by means of the causative verbs *fazer* ‘to make’, *mandar* ‘to order’, and *enbiar* ‘to send’. In the case of (1), it can be shown that the valency increase of Old Spanish intransitive verbs is motivated by the Arabic model that displays a stem II or IV verb that has factitive and causative meaning. As far as formal strategy (2) is concerned, the motivation behind its augmented use in Old Spanish can be traced not only to the Arabic semantic model, but also to the formal similarity of the *a-* prefix, found both in Arabic and Old Spanish. Finally, strategy (3) also follows the analytical constructions found in the Arabic model, especially regarding the original semantics of the auxiliary causative verbs. The analysis of the Old Spanish causative constructions confirms that translation is an act of negotiation and compromise and is influenced deliberately where the given repertoire of the target language exhibits a general predisposition that permits such contact-induced structures.



1 Introduction

The main focus of this chapter is the contrastive historical analysis of the Old Spanish translation, *Calila e Dimna*, and its famous and widespread Arabic model, *Kalila wa-Dimna* (see Section 2), with particular attention paid to Old Spanish factitive and causative constructions (see Section 5). In this study, translation as such is regarded as a particular form of language contact, where structures and semantic concepts may be imitated in the target language by the translators following patterns found in the source model (Döhla 2008, Haßler 2001).¹ At the same time, it can be observed that structures of the target language text diverge from those of the model language in order to comply with the linguistically acceptable repertoire of the target language, especially if the target language, like written Old Spanish, is still a young language lacking certain syntactic and semantic fine-tuning (Bossong 1979, Galmés de Fuentes 1996, Huffman 1973). As far as the Arabic–Old Spanish translations of the thirteenth century are concerned, the studies by Bossong (1979, 2008) and Döhla (2008, 2009) demonstrate how thoroughly the translators operated while translating the bulk of texts from Arabic, predominantly using structures from their available Romance repertoire and modifying them creatively by derivational and analytic strategies.

The main research question of this study concerns the morphological strategies used by the translators in order to create factitive and causative meanings and the possible influence of the formal side of the derivational and analytic constructions exhibited in the Arabic model. The study itself is intended to be purely descriptive and non-statistical (except for Table 4, taken from Sanaphre Villanueva 2010), since only the full corpus of translated texts of the thirteenth century (see Faulhaber 2004) would fully justify statistical analyses. Unfortunately, the full corpus, Old Spanish and Arabic, has not yet been edited for large-scale comparative studies.

Before getting to the descriptive analyses in Section 5, Sections 3 and 4 are dedicated to the differences between Arabic and (Old) Spanish morphology and to general aspects of factitive and causative constructions.

¹Translations are popular windows for matter borrowings, that is the adoption of entire constructions of the source language, including the phonological shape and its corresponding meaning (Sakel 2007: 15), to enter the target language. There are also Arabisms found in *Calila e Dimna* (Döhla 2009), although these are few in number (44) considering the overall amount of 2,535 lexemes (types).

2 Contact between Spanish and Arabic and *Calila e*

Dimna

When compared to other areas where Romance languages are spoken, the Iberian Peninsula of the Middle Ages stands out due to the coexistence of two different faith systems (Christianity and Islam), several different realms (the northern Christian shires and emerging kingdoms versus the Muslim sphere of control in the southern regions), and, most importantly, two different language families (Indo-European and Semitic). As can be imagined and as is well known, out of this coexistence a number of language-contact scenarios emerged. In the first place, the Christians living under Muslim rule have to be mentioned. According to Latin sources of the Christian North, they called themselves *Mozarabs*,² which points to the fact that they were ‘Arabicized’ (not ‘Islamized’), since the term derives from the Classical Arabic *mustaʿrab* ‘Arabicized’ (Corriente 2008a: 383). Besides this inclination toward the Arabic language (Bossong 2007: 69), their inherited language was Ibero-Romance, which was also used by Muslims and Jews in al-Andalus.³ Apart from a high number of early Arabisms, especially the bilingual *ḥaraḡāt* (Spanish *jarchas*) (Corriente 1997, 2008b), the final verses of the Andalusi poem pattern called *muwaššah*, which bear witness to the creative and aesthetic use of Arabic-Romandalusi (see Note 3) code-switching, must be mentioned. While this kind of language contact took place on a popular, albeit court, level in the case of the *ḥaraḡāt*, the thirteenth century marked a turning point in language use, when the most powerful of the advancing Christian kingdoms, Castile, promoted its vernacular Castilian language as the official language of the court. Thus, the first document written in Old Spanish is the *Tratado de Cabreros* of 1206 (Wright 2000), a treaty between Alfonso VIII of Castile and Alfonso IX of

²In fact, the term *Muzaraues* (Pérez 2010: 497) can be found in a Latin document from the northern region of León as far back as 1024. However, the adoption of Arabisms in the Latin of the Christians in the north can be dated back to the ninth century, with *citara* ‘curtain’ (812), *barrius* ‘quarter’ (887), *adorra* ‘kind of shirt’ (887), *zuramen* ‘Moorish tunic’ (887), and *adria* ‘special kind of share’ (896) (Pérez 2010: 172, 95, 13, 805, 815). This early appearance of Arabisms in Latin documents of the northern Christian realms can be linked sociohistorically to the migration of Mozarabs from the south to the northern no-man’s-land, which was repopulated actively by the Christians. On the other hand, even though the term *Mozarab* comes from Arabic, there are no traces of the use of the Arabic term among Muslim writers to refer to Christians in al-Andalus (Simonet 1983: VIII).

³This Ibero-Romance language has been called *mozárabe/Mozarabic* by some scholars (Beale-Rivaya 2016, Galmés de Fuentes 1983, Hitchcock 2008: XVII) and, more appropriately, *romandalusí* by Corriente (2008b). The term proposed by Corriente suggests a Romance origin of this language used not only by the Christian, i.e., Mozarabic, population of al-Andalus but by all other groups, too (Corriente 2008b: 98).

León, 24 years before the unification of Castile and León under Fernando III in 1230. This first use of Old Castilian as the official language of the court paved the way for an early elaboration process (Haugen 1983: 273),⁴ even before the first vernacular grammar was written, which was published by Antonio de Nebrija in 1492. The elaboration process was twofold; one originated from the desire to express mostly Christian and Classical topics in the vernacular in prose and rhyme, while another made Old Castilian the target language of the translation of scientific and wisdom literature from Arabic. These Arabic documents had been falling into the hands of the Christians, advancing their *reconquista* from the northern parts of the peninsula toward the south, from the eleventh century onward, thus already leading to Arabic–Latin translations in Toledo in the twelfth century. However, as far as the patron of the translations of the thirteenth century, Alfonso X the Wise, is concerned, a clear preference for the vernacular can be detected, as can be read in the foreword of the *Lapidario* (1245), a treatise on the magic properties of (gem) stones.

[E]t dizien le Yhuda Mosca el Menor, que era mucho entendido en la arte de astronomia et sabie et entendie bien el arauigo et el latin. Et de que por este iudio, su fisico, ouo entendido el bien et la grand pro que en el iazie, mando gelo trasladar de arauigo en lenguaie castellano por que los omnes lo entendiesses mejor. (Rodríguez M. Montalvo 1981: 19)

[A]nd he was called Yehuda Mosca Junior, who was very adept in the art of astronomy and he knew and understood Arabic and Latin well. And since, with the help of this Jew, he [the infant Alfonso] had understood the greatness and the great advantage which was lying in it [the *Lapidario*], he ordered him to translate it from Arabic into the Castilian language, so that people would understand it better. (Own translation)

Thus, under the patronage of Alfonso X the Wise (approximately 1250–1282), there was a wave of translations⁵ into Old Castilian, comprising many topics,

⁴In the terminology of Haugen (1983: 273), “elaboration” is “the continued implementation of a norm to meet the function of a modern world. [...] A modern language of high culture needs a terminology for all the intellectual and humanistic disciplines, including the cultural under-world that runs from low to popular”. The norm-driven implementation of Castilian in Spain was not followed up until the eighteenth century.

⁵In general, the translation process was usually undertaken by a team of translators, with one translator reading the text aloud or translating it into Andalusí Arabic. Then, another member of the team would translate the text heard into Old Castilian, which was then written down (Hilty 1954: 3, Hilty 2005: XXIX, Rodríguez M. Montalvo 1981: 19).

such as astronomy/astrology, zoology, veterinary medicine, agriculture, mineralogy, and oriental wisdom. This special type of language contact (Döhla 2008, Haßler 2001) initiated a process of acculturation which, according to Bossong (1979: 6), is the “Prozeß der Universalisierung und/oder Komplektisierung einer R[esponse]-Sprache unter dem Einfluß und durch Anregung einer S[timulus]-Sprache” ‘the process of universalization and/or complexitization of the response language under the influence of a stimulus language’ (own translation). In this context, universalization refers to the elaboration of the lexicon, whereas complexitization alludes to the development of syntactic expressivity. Vestiges of both processes are clearly detectable in the Arabic–Old Castilian translations of the thirteenth century. However, even though there are studies regarding the lexicon (Bossong 1978, 1979, Döhla 2009) and contact-induced syntactic structures (in chronological order: Tallgren 1934, Dietrich 1937, Hottinger 1958, Huffman 1973), this whole area of investigation still displays several desiderata, which can best be summed up in a general comparative grammar of Arabic and Old Spanish based on the translations of the thirteenth century.

The linguistic material discussed in the following sections is taken from the Alfonsine Old Castilian version of the famous Arabic mirror of princes called *Kalīla wa-Dimna*. Its origin lies in the Old Indian *Pañcatantra* ‘Five stories’, a collection of fables and allegories compiled sometime between the third and fourth centuries. It was translated into Middle Persian and expanded at the same time in the sixth century, and from there it was rendered into Arabic in 750 AD by Ibn al-Muqaffaʿ. This Arabic translation has been transmitted in many copies, which can be subclassified into different lines of manuscript traditions, one of them being the Iberian branch with a particular order and number of chapters. Without going into further detail, at this point, it is sufficient to note that the translation, which was presumably ordered by the infant Alfonso in 1251 (Alfonso the X from 1252 onward), has survived to the present day in two manuscripts. One of them (manuscript A; kept at the library of the Real Monasterio de San Lorenzo de El Escorial, sig. h-III-9), dates from the first third of the fifteenth century, and quite accurately reflects the model of Arabic manuscripts that are considered part of the Iberian branch.⁶ The other (manuscript B; kept at the library of the Real Monasterio de San Lorenzo de El Escorial, sig. x-III-4), dates from 1467 and exhibits influences from other Arabic manuscript traditions and, possibly, from the Hebrew translation of *Kalīla wa-Dimna* elaborated by Rabbi Joël in Spain in

⁶For example ms. Arabe 3478, supplément 1795, Bibliothèque Nationale, Paris, cited here as Ibn al-Muqaffaʿ (n.d.); ms. 4095/3, Süleymaniye Kütüphanesi, Istanbul, cited here in the edition of ʿAzzām ʿAbdu & Bek (1941).

the twelfth and thirteenth centuries as well (Döhla 2009: 57, 78). This means that manuscript B, despite its similarities with manuscript A, is not a mere copy of manuscript A. As far as the examples of the Old Castilian *Calila e Dina* are concerned, they are taken from Döhla (2009). The Arabic model sentences are taken from those manuscripts specified in Note 6.

3 Arabic and Spanish contrastive morphology

Before presenting a brief typology of formal valency-increasing strategies as well as addressing the concrete valency-increasing operations found in the Old Spanish version of *Kalila wa-Dimna* in more detail and contrasting them with the respective source constructions of the Arabic model, it seems reasonable to dedicate a brief paragraph to the differences between Arabic and (Old) Spanish concerning their basic morphological strategies.

In general, the main difference between Arabic, a Semitic language of the Afroasiatic language family, and (Old) Spanish, a Romance language of the Indo-European language family, lies in the use of transfixation as the major derivational device in Arabic. In this context, “[a] **transfix** may be defined as a discontinuous affix that disrupts the base to which it is attached” (Broselow 2000: 552). Other authors refer to this morphological type as “templatic morphology” (Lieber 2009: 81-82), “root and pattern morphology”⁷ (Holes 2004: 99), or as Spanish “interdigitalización” ‘interdigitation’ (Bossong 2008: 21).

In practice, as opposed to the exclusively linear, sequential, or concatenative morphology found in Romance languages, Arabic morphological derivation departs from a lexical root consisting of a fixed sequence of three consonants⁸ in most cases. From there, “Arabic is like a mathematical game. You take the root of a word, [...], and you start playing” (Drißner 2015: 4), creating different patterns according to fixed structures by adding the vowels /a, i, u/, and/or /a:, i:, u:/, by omitting a vowel between root consonants and/or by lengthening the middle

⁷Broselow (2000) distinguishes two general patterns of transfixation; one that she calls “segmental transfixation”, where the opposition between two derivations lies in the exchange of vowels, such as in the Arabic *ḥazina* ‘to be sad’ versus *ḥazana* ‘to make sad’, and a second that corresponds to the root and pattern morphology.

⁸Wehr (1952) lists 2,967 roots with three consonants and 362 with four consonants (Drißner 2015: 7). The roots with three consonants also contain those roots where one consonant is represented by the glottal stop (*hamza*) or by the half-consonants *wāw* and *yāʾ*, or where the second and the third consonants are identical, thus appearing as one geminated consonant. However, the same derivational principles are applied to these particular types of roots, even though they might experience some phonetical alternations and adaptations.

consonant (see Table 1). At the same time, it should be mentioned that Arabic, just as Spanish does, makes use of other types of affixes, such as prefixes and suffixes, for inflectional and derivational morphology, which can also be combined with basic lexemes, that is with stems already modified by transfixation. This way, verbal, nominal and adjectival derivations are created, all of them belonging more or less to the same semantic field. For example, the consonant sequence *k-t-b* construes lexemes that have something to do with the domain of ‘writing’, thus making the nominal derivations displayed in Table 1 possible, among others.

Table 1: Some basic derived nouns of the root *k-t-b* in Classical Arabic

lexeme	pattern	meaning
<i>kātib</i>	$C_1\bar{a}C_2iC_3$	‘writer, scribe, secretary’
<i>kuttāb</i>	$C_1uC_2C_2\bar{a}C_3$	‘writer, scribe, secretary (PL)’ [but also sg ‘Koran school’]
<i>maktab</i>	$maC_1C_2aC_3$	‘office, desk’ (<i>nomen loci</i> , literally ‘place where you write’)
<i>makātib</i>	$maC_1\bar{a}C_2iC_3$	‘office, desk (PL)’
<i>kitāb</i>	$C_1iC_2\bar{a}C_3$	‘book’
<i>kutub</i>	$C_1uC_2uC_3$	‘book (PL)’
<i>miktāb</i>	$mi-C_1C_2\bar{a}C_3$	‘stylus’ (<i>nomen instrumenti</i> , literally ‘device you write with’)

As can be noted, the so-called ‘broken plural’ forms *kuttāb*, *makātib*, and *kutub* are not formed by simply adding a plural suffix, but by reshaping the form of the construction, applying different transfix patterns.

At the same time, the trilateral root interlocks with a pattern to produce a number of so-called verbal forms that usually indicate semantically related meanings. Table 2 displays eight derived forms of the root *k-t-b*.⁹

Eight of the possible ten derived, and commonly used, verb stems (Alward et al. 2019, Fischer 1972: §§163–173, Holes 2004: 100–105) are to be found in the *kataba* entry in Wehr’s dictionary (1976: 812; see Table 2). Stems II, III, and IV are created – departing from *kataba* – by gemination of the middle consonant (II), by lengthening the vowel /a/ between C_1 and C_2 (III), and by extending the root by

⁹For the basic verbal form, like *kataba* (stem I), Arabic grammarians have always used the simplest of all verbal forms, which corresponds to 3SG.M.PFV of stem I. However, whenever I provide the meaning of a basic verbal form, I use the English infinitive; thus, instead of the literal ‘he wrote’, I use ‘to write’.

Table 2: The derived verbal stems of the root *k-t-b* in Classical Arabic

stem	lexeme	pattern	meaning
I	<i>kataba</i>	$C_1aC_2aC_3a$	‘to write’
II	<i>kattaba</i>	$C_1aC_2C_2aC_3a$	‘to make write’
III	<i>kātaba</i>	$C_1āC_2aC_3a$	‘to exchange letters, to correspond’
IV	<i>ʔaktaba</i>	$ʔaC_1C_2aC_3a$	‘to make write, to dictate’
VI	<i>takātaba</i>	$taC_1āC_2aC_3a$	‘to write to each other, to exchange letters’
VII	<i>inkataba</i>	$inC_1aC_2aC_3a$	‘to subscribe’
VIII	<i>iktataba</i>	$iC_1<t>aC_2aC_3a$	‘to copy, to be registered, to subscribe, to contribute’
X	<i>istaktaba</i>	$ista-C_1C_2aC_3a$	‘to ask to write, to dictate, to make write, to have a copy made’

means of the prefix /ʔa-/ while omitting any vowel between C_1 and C_2 . Stem VI is produced by adding the prefix *ta-* to stem III, whereas stem VIII is formed by the insertion of the infix /-t-/ after C_1 . Finally, stem VII is created by adding the prefix *n-* to stem I, preceded by *i-*, in order to avoid an onset with two consonants, whereas in stem X, the root pattern *-ktaba* receives the prefix *st-*, preceded by the same prosthetic vowel as in stem VII. As can be seen in Table 2, each stem is connected to the semantic field of ‘writing’. In particular, stems II and IV are used to construe factitive and causative meanings in Arabic. The main semantic values of the other stems are as follows: stem III: conative; stem V: reflexive, effective; stem VI: reflexive, reciprocal; stem VII: reflexive, passive; stem VIII: reflexive; stem X: reflexive (Corriente 2002: 153, Fischer 1972: §§163–172, Holes 2004: 100–105). I will come back to the concrete application of these root-transfix patterns, especially of those representing causative values (stems II and IV), in the following sections.

This formal flexibility of the consonantal roots, in combination with different nuances in meaning, is used productively and creatively by the *bulaḡāʔ* ‘eloquent (PL)’, such as Ibn al-Muqaffaʔ, who translated *Kalīla wa-Dimna* from Middle Persian into Arabic. The following two examples illustrate the creative potential of Arabic derivational morphology, which is difficult to imitate using the concatenative morphology of Romance and Germanic languages.

The first example demonstrates the application of a common linguistically aesthetic technique found in Classical Arabic literature and, thus, in *Kalīla wa-*

Dimna: double constructional parallelism, where both the form, that is the derivational pattern, and the meaning of two adjacent nouns are synomorph and synonym at the same time, as shown in Example (1).¹⁰

- (1) (Arabic; Semitic, Afroasiatic)
 bahāʔ-u-h̄^ū wa-ḡamāl-u-h̄^ū
 beauty-NOM-POSS.3SG and-beauty-NOM-POSS.3SG
 ‘his beauty and his beauty’¹¹

The Old Spanish equivalent in *Calila e Dimna* is *su beldad e su fermosura* (Döhla 2009: AI.31e), both synonyms, but not synomorphs, since both nouns do not follow the same derivational pattern as the Arabic nouns do ($C_1aC_2āC_3$). *Beldad* is formed with the nominal suffix *-dad*, whereas *fermosura* exhibits the nominal suffix *-ura*. The locating of these Old Spanish synonyms in translation from Arabic may serve to shed some light on synonymy in Old Spanish.

The second example is even more striking and even less imitable by Old Spanish derivational means. In this case, the verb stem IV is formed and combined with a meaning that does not appear as such in any of the many Arabic dictionaries. The Arabic *nahr* means ‘river’, derived from stem I of the verb *nahara*, ‘to flow copiously, to stream forth’. Taking the root *n-h-r* as the morphological and ‘river’ as the semantic point of departure, Ibn al-Muqaffaʔ derives a verb following the pattern of stem IV, *ʔanhara*, as can be seen in Example (2).

- (2) fa-ya-kūn-u maʔal-ī fī ǧālīka miṭla
 so-3SG.M-be.IPFV-IND example-POSS.1SG in DEM.DIST similar.to
 ʔl-kalb-i ʔllaǧī **yu-nhir-u**¹²
 DEF-dog-GEN REL.3SG 3SG.M-come.by.river.IPFV-IND
 ‘So my example would be similar to the dog who was coming by a river’.
 (Cheikho 1905: 38)

¹⁰In the following Arabic examples, the transcriptions make use of the special characters recommended by the *Deutsche Morgenländische Gesellschaft*. However, my transcriptions include the following three special principles: any long vowel or any consonant that is not represented in the written *rasm* is in superscript; any /a/ represented by *yāʔ* in writing receives a grave accent (*ā*); the omission of the vowel represented by the letter *ʔalif*, usually in definite articles and some derived verbal stems (VII, VIII, X), in combined reading of the individual elements of a sentence, is represented by the single quotation mark ‘.

¹¹Of course, in order to reflect the construction of synonyms, we could translate this as ‘his magnificence and his beauty’.

¹²The conjugation of the perfective verbal paradigm makes use of suffixes, whereas the imperfective paradigm uses prefixes (and occasional suffixes as well). Both paradigms also differ in the assignment of vowels to the stem patterns.

This semantic innovation, created by Ibn al-Muqaffaʿ, was analyzed correctly and translated adequately by the Medieval Spanish translator, using the complex analytic construction *que yva por un rrio* ‘who was going by a river’.

After comparing both Old Spanish manuscripts with the Arabic manuscripts of the Iberian branch (Döhla 2009), it can be stated that the translator(s) of the Arabic *Kalīla wa-Dimna* were skilled in both Arabic and the Old Spanish repertoire available to them, only leaving a few ‘dark’ passages in Old Spanish and only displaying a low number of misreadings (or mishearings).

With this in mind, in Section 4, I present a brief overview of the actual construction analyzed in the subsequent paragraphs.

4 Factitives and causatives as valency-increasing operations morphology

Valency-decreasing (by means of anticausative, passive, antipassive, incorporative, reflexive, reciprocal, and medial constructions) and valency-increasing operations (Comrie & Polinsky 1993, Dixon & Aikhenvald 2000, Haspelmath & Müller-Bardey 2004, Payne 1997: 175–222) are very common in the languages of the world. Among the latter constructions, causatives in particular have been attracting the attention of linguists for decades (see Comrie 1989, Lehmann 2016, Shibatani 1976, Shibatani & Pardeshi 2002, Song 1996, 2001).

In this paragraph, as a first step, I depart from an onomasiological point of view as far as the two relevant valency-increasing operations – factitive and causative – are concerned. This seems to be the logical consequence resulting from the previous section, where the flexibility and creativity of the Arabic root and pattern morphology was contrasted with the linear, semantically less flexible morphology of Old Spanish.

Thus, the term *factitive* is used in this study to refer to those constructions where the meaning side represents a complex event, where the caused embedded action usually denotes a qualitative state (*to be someone*), a change of state (*to become something/somebody*), or an intransitive¹³ action (*to fall*), meaning the embedded action or declarative statement is monovalent. In all three cases, the causer, the additional actant who causes the embedded action to happen, is responsible for the qualitative state (*to make someone be something/somebody*), the change of state (*to make someone become something/somebody*), or the intransitive action (*to make fall, to fell*). Thus, the valency-increasing operation results

¹³Intransitive verbs can be stative (non-active) or active, meaning the only actant may take the role of the undergoer or the agent.

from the addition of the causer as the agent (i.e., syntactic subject) of the factitive construction. In this way, the former subject, be its semantic role that of the agent or the undergoer, is demoted to object status. A case in point is the application of apophony, or ablaut, in Indo-European languages such as English, German, or Latin, where vowel fronting is used to create a factitive meaning, as in the English *to fall* > *to fell*, the German *fallen* > *fällen*, and the Latin *cadere* > *caedere*. The same segmental transfixation can be found in some Arabic verbs like *ḥazina* ‘to be sad’ versus *ḥazana* ‘to make sad, to sadden’ (both are verbs with the pattern of stem I) (Wehr 1976: 174). But Arabic also makes use of the derived stems II and IV to express the same meaning of ‘to make sad’: *ḥazzana* (II) and *ʔaḥzana* (IV).

The term *causative* is used in this study to refer to a similar complex event as in the case of factitives, but this time, the embedded action is represented by a verb, which is transitive or divalent. Hence, the embedded action that already comprises an agent and a patient is extended by the influence of another agent, the causer, who acts on the embedded agent in a way that s/he performs the action with the embedded patient. The embedded agent is degraded to the role of causee (see Figure 1), which in several languages is coded as a direct object, so that, syntactically, there are two direct objects representing the extralinguistic referents of the embedded action, as in *he made his son write a song* or *he made him write it*.

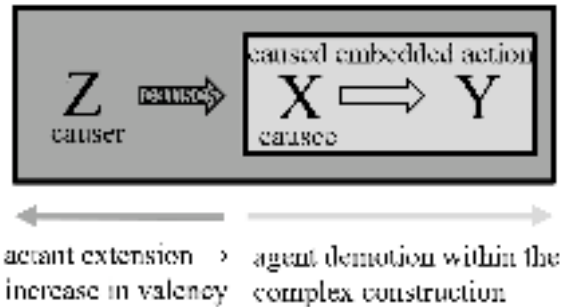


Figure 1: The complex event expressed by causative constructions comprising three actants: the causer (Z) of the embedded action, which is performed by the causee (X), who is the agent of the embedded action that is executed upon the patient (Y).

The same double object constructions can be observed in Arabic taken from the monolingual dictionary *Al-Munğid*, illustrating the use of *ʔaktaba* and its meaning (see Example 3).

pointed out, the research question behind this analysis is to see whether the Arabic model may have influenced the translator(s)' choice as far as the formal side of the Old Spanish factitives and causatives is concerned.

5 Factitive and causative constructions in the Old Spanish translation of *Kalīla wa-Dimna*

5.1 Zero morphology and conversion

The first set of examples of valency-increase consists of verbs that, in principle, are intransitive, but are used with two actants in *Calila e Dimna*. These transitive uses of otherwise intransitive verbs mostly fly under the radar of Hispanic lexicologists, because the verb does not display any morphological change that would indicate a difference in use and meaning. Moreover, the fact that the Old Spanish of the thirteenth century in general, and that used in the translated texts from Arabic in particular,¹⁴ is a VS(O) language (Bossong 2006, López García 2000, Neumann-Holzschuh 1997) makes the detection of transitive uses a difficult undertaking, because any superficial analysis would read the NP after the verb as the subject of the sentence, simply following the entrenched idea already existing in the lexicologist's mind. Only a meticulous syntactic and semantic analysis could find the cases in point, as in the following examples from *Calila e Dimna*.

The first example concerns the Old Spanish verb *encaresçer*, which has the basic meaning 'to become expensive'. It is a deadjectival derivation from *caro* 'expensive' by means of the circumfix *en-ADJ-esçer*, the suffixal part of which was already used in Latin to create an inchoative verb from an adjectival root (see below) and in Vulgar Latin to generate a factitive meaning (Väänänen 1978: §316). This basic meaning of *encaresçer* is also attested in *Calila e Dimna*, where we can find the sentence presented in Example (4).

- (4) Et acaesçio que encaresçio la
 and happen.3SG.PST.PFV COMPL become.expensive.3SG.PST.PFV DEF.F
 miel e la manteca
 honey and DEF.F fat
 'And it happened that honey and fat became (more) expensive'. (Döhla
 2009: A.VI.9)

¹⁴Classical Arabic, like Biblical Hebrew, follows the Proto-Semitic VS(O) order. However, the Modern Arabic dialects and Modern Hebrew prefer an SV(O) order.

The word order in the subordinate clause introduced by the complementizer *que* is VS. The fact that there is a mismatch as far as the agreement between the sentence-initial verb (in SG) and the subject-NP (in PL, since there are two nouns) is concerned, is certainly due to the Arabic model, which displays the following text:

- (5) wa-wāfaq-a ḍālika ḡal-ā
 and-become.PFV-3SG.M DEM.DIST become.expensive.PFV-3SG.M
 'l-saman-u wa-'l-ḡasal-u¹⁵
 DEF-fat.M-NOM and-DEF-honey.M-NOM
 'and it happened (that) fat and honey became (more) expensive' (Ibn al-Muqaffaʿ n.d.: fo. 139r)

In Example (5), it is possible to observe the same construction and agreement in the Arabic pattern as found in the Spanish translation.

In contrast to this example of the basic use of *encarezcer* in Old Spanish, the following attested use (see Example 6) in *Calila e Dimna* certainly diverges from the latter by its increase in valency and by a degree of difficulty regarding the assignment of the correct meaning in the given context. The particular chapter where the sentence in (4) can be found (chapter VI in Döhla 2009), deals with the benefits of friendship. Thus, a raven wants to befriend a mouse, who is hiding in a hole in the ground, frightened to come out because of the apparent danger, that is the presence of the raven on a nearby tree. Nevertheless, the raven keeps on talking to the mouse, trying to convince it to become its friend. In the course of the dialogue, the raven utters the following sentence:

- (6) a. e non me encarezca-s la cosa
 and NEG 1SG.OBJ make.expensive.PRS.SBJV-2SG DEF thing
 (Döhla 2009: AIII.31a)
 b. e non encarezca-s el amor¹⁶
 and NEG make.expensive.PRS.SBJV-2SG DEF love
 (Döhla 2009: BIII.31a)

¹⁵In the actual manuscript, *saman-* and *ḡasal-* are vocalized with *-a* (ACC), which does not seem to be correct. It should be *-u*, which marks the nominative case.

¹⁶Interestingly, the word *friendship* does not appear in any Arabic manuscript. However, one of the aesthetic techniques of the translator(s) was to play around with the sequence of phonemes between the source and the target language. This way, *amor* was chosen as the equivalent of *ḡumūr*, the plural of *ḡamr* 'matter', just because of phonological similarities concerning the consonants. Another case in point is the translation of *ṣabr* 'patience' with *sufremento* because of the *suf-* onset. *Sufremento* actually means 'suffering'.

With the basic meaning of *encarezcer* in mind and considering the transitive construction, the literal meaning of the two sentences would be ‘and do not make the matter expensive for me’ and ‘and do not make the friendship expensive’, respectively. The verb clearly exhibits the 2SG -s (*encarezca-s*) and a direct object and, thus, is used divalently. However, when looking at the Arabic model in Example (7), the actual meaning of *encarezcer* becomes clear.

- (7) wa-lā tu-ṣaʿfīb ʔal-ʔumūr-a
 and-NEG 2SG.M-make.difficult.JUS DEF-matter;PL-ACC
 ‘and do not make the matter difficult’ (ʔAzzām ʔAbdu & Bek 1941: fo. 48v)

The Classical Arabic verbal form *tuṣaʿfīb* is in the jussive mood; it is basically the indicative imperfective form of the verb with elision of the final vowel. The derivational pattern follows stem II, that is with factitive/causative meaning. Thus, *no encarezcas* means ‘do not make difficult’, which is not surprising at all considering the fact that the adjective *caro*, apart from ‘expensive’, also had the meaning ‘difficult’ in Old Spanish (Kasten & Cody 2001: 138). In this way, the translator(s) extended the meaning of the Old Spanish *encarezcer* according to the Arabic model:

ṣaʿfuba (stem I) ‘to be difficult’ / ṣaʿb ‘difficult’ → ṣaʿfaba (stem II) ‘to make difficult’

caro ‘difficult’ → *encarezcer* ‘to make difficult’

It is interesting to observe that none of the vocabularies dealing exclusively with the lexicon of *Calila e Dimna* (Holmes 1936, Pérez 1943, Stinson 1967) and none of the major Old Spanish dictionaries (Alonso 1996, Kasten & Cody 2001) indicate the correct meaning of *encarezcer* in the context of the above-mentioned text passage.

But there are other examples of the same kind where the Old Spanish translation follows the Arabic model, increasing the valency of an otherwise intransitive Old Spanish verb.

- (8) a. e festina al tardinero
 and make.hurry.3SG.PRS.IND OBJ;DEF slow.person
 ‘and makes the slow person hurry’ (Döhla 2009: I.174t)
 b. wa-yu-sarriʿ-u ʔl-baʿīʔ-a
 and-3SG.M-make.hurry.IPFV-IND DEF-slow.person-ACC
 ‘and makes the slow person hurry’ (Ibn al-Muqaffaʿ n.d.: fo. 49r)

In (8a), the verb *festinar*, which is normally used intransitively, just like the classical Latin *festināre*, in the sense of ‘to hurry, to hasten’, is extended to a di-valent scope in order to translate the factitive meaning of *sarrafa* ‘to urge someone, to hurry someone’. This is a stem II verb pattern, derived from the basic stem *sarufa* ‘to be fast, to hurry’.

All cases of valency conversions found in the Old Spanish translation of *Kalila wa-Dimna* share the property of zero morphology, thus not overtly indicating the change in meaning and syntactic scope. In the two examples discussed here, only *encarescer* follows the Arabic model as much as it can, given the morphological possibilities of Old Spanish. After editing the two Old Spanish manuscripts and comparing them meticulously with the Arabic model (Döhla 2009), it can be said that the attitude of the translator(s) can be determined as one that makes every effort to express as much as possible of the Arabic model by means of the Old Spanish lexicon. At the same time, the translator(s) are adventurous and creative and do not hesitate to introduce lexico-semantic contextual innovations based on their solid knowledge of the Old Castilian language.

5.2 Denominal derivations

The second set of verbs that are used to create a factitive meaning consist of denominal and deverbal derivations. In this case, a nominal root, for example, is taken with its basic meaning as the point of departure for verbalization by adding the prefix *a-* or *en-*, cutting off the final vowel of the noun and adding whatever conjugational pattern (following the *a*-conjugation) is required by the given context. Alternatively, a verb is prefixed by *a-*, thus creating the same factitive/causative meaning. Apart from being a productive method of creating factitive verbal constructions with the help of morphological means in *Calila e Dimna*, both strategies are also present in other Romance languages, such as French, Italian, and Portuguese, and Judeo-Spanish. However, as I will show in the following examples, the Old Spanish translator(s) of *Kalila wa-Dimna* made use of this construction to reproduce the Arabic stem II or stem IV derivations that have a factitive/causative meaning. Cases in point are the following three passages from *Calila e Dimna* with the respective Old Spanish translation and the Arabic model.

- (9) a. et esfuërça al cobarde, e **encobarda/acouarda**
 and strengthen.3SG.PRS OBJ;DEF coward and make.coward.3SG.PRS
 al esforçado
 OBJ;DEF brave
 ‘and it [fate, fortune] strengthens the coward and makes a coward out
 of the brave’ (Döhla 2009: A/BI.174t)

- b. wa-yu-šağǧīf-u ʾl-ġabān-a
and-3SG.M-encourage.IPFV-IND DEF-coward-ACC
wa-yu-ḡabbīn-u ʾl-šağāf-a
and-3SG.M-make.coward.IPFV-IND DEF-brave-ACC
'and it encourages the coward and makes a coward out of the brave'
(Ibn al-Muqaffaʿ n.d.: fo. 49r)
- (10) a. e **apriuado** lo mas que a todos
and make.confidant.3SG.PST.PFV OBJ.3SG.M more than OBJ all.PL
sus vasallos
POSS.3PL vassal.PL
'and he [the king] made him confidant, closer than all his other
vassals' (Döhla 2009: AXII.16)
- b. et **puso-le** en mayor pryuança que a
and put.3SG.PST.PFV-OBJ.3SG.M LOC more confidence than OBJ
ninguno de sus vasallos
none GEN POSS.3PL vassal.PL
'and he granted him more confidence than any other of his vassals'
(Döhla 2009: BXII.16)
- c. wa-ʔaḥašš-a-h^ū dūna
and-make.private.PFV-3SG.M-OBJ.3SG.M more.than
ʔaṣḥāb-i-h^ī
companion.PL-GEN-POSS.3SG
'and he made him confidant, more than his companions'
(ʿAzzām ʿAbdu & Bek 1941: 94v)¹⁷
- (11) a. El rrey non **apriuua** a los omnes por
DEF king NEG make.confidant OBJ DEF.M.PL men because.of
la priuança de sus padres, ...
DEF.F.SG confidence GEN POSS.3PL forefather.PL
'The king does not make men confidants because of the confidence of
their forefathers'. (Döhla 2009: AI.41a)
- b. ʔinna ʾl-sultān-a lā **yu-dnī**
truly DEF-sultan-ACC NEG 3SG.M-make.confidant.IPFV;IND

¹⁷The edition of ʿAzzām ʿAbdu & Bek (1941) contains *aḥtaṣṣa*, stem VIII. In Ibn al-Muqaffaʿ (n.d.), the actual verb is missing.

'l-riğāl-a li-qurb-i
 DEF-men-ACC because.of-proximity-GEN
 ʔabāʔ-i-him
 forefather.PL-GEN-POSS.3PL.M
 'The sultan does not make men confidants because of the proximity
 of their forefathers'. (Ibn al-Muqaffaʔ n.d.: fo. 29v)

In Example (9a), the translator(s) created a new word that can only be found in Müller (1987: I, 443), with the single medieval reference of *Calila e Dimna*. All other pieces of textual evidence are from later centuries, from the end of the fifteenth century onward (*acobardar* earlier than *encobardar*, the latter being very rare). Here, the translator(s) followed a similar derivational pathway as in Arabic, always making use of the morphological repertoire available to them in Old Spanish:

cobarde 'coward' → *ser cobarde* 'to be a coward' → *encobardar/acovardar*
 'to make (someone a) coward'

ğabān 'coward' → *ğabuna* 'to be a coward' → *ğabbana* 'to make (someone a) coward'.

However, the translator(s) were not able to reproduce the aesthetic morphological rhyme parallelism applied in the following Arabic formula with C_1 - C_2 - C_3 and D_1 - D_2 - D_3 representing two different consonantal roots:

$wa-yu C_1 a C_2 C_2 i C_3 u$ 'l- $D_1 a D_2 ā D_3 a$ $wa-yu D_1 a D_2 D_2 i D_3 u$ 'l- $C_1 a C_2 ā C_3 a$

Nevertheless, they were still eager to get as close as possible to the etymological pattern found in the Arabic model, thus using *esforçar* – *esforçado* and *cobarde* – *encobardar/acovardar*, where both pairs are etymologically related through derivation just as *šagğāfa* – *šagğāʔ* and *ğabbana* – *ğabān*.

In the next two passages presented above, there is another denominal (10) and a deverbal derivation (11). In both cases, the resulting verbal form exhibits the *a-* prefix, *apriuar* and *apriuar*. Whereas **privadar* is not documented according to *Real Academia Española* (n.d.) and Pharies (2002), *privar* can be found in other medieval texts with the meaning 'to dispose of' (Kasten & Cody 2001: 569), which has been inherited from its Latin etymon *privāre*. The Old Spanish *privar* in the meaning 'to make someone a confidant' also appears once in *Calila e Dimna* (Döhla 2009: AI.140b). Unfortunately, none of the Arabic models exhibits

the same sentence (‘and you made him a confidant’), which could be an indication as to the nature of the Arabic verb serving as a model.

Nevertheless, it can be stated that *aprivadar* and *apriver* follow the morphological pattern of Arabic stem IV derivations (*ʔaḥaṣṣa* and *ʔadnà*, respectively) that have an overt *ʔa-* prefix. Of course, this *a-* prefix is certainly not a matter borrowing from Arabic.¹⁸ It is definitely taken from Romance matter by making use of an already existing construction, but it is used productively to emphasize the factitive meaning of the verb. Thus, the translator(s) chose linguistic structures as Romance as possible by using original Old Spanish matter, but they also displayed a certain “voluntad de dejarse influir” (Galmés de Fuentes 1996: 230) ‘willingness to be influenced’, to adapt to Arabic structures where their intuition of acceptable Romance grammar would permit. In the case of the translator(s) of *Kalīla wa-Dimna*, for example, they did not accept another valency-increasing strategy (that does not generate a factitive/causative meaning), namely the *figura etymologica*. This is a special case of paronomasia, where the monovalent nature of an intransitive verb is changed to divalent by adding an etymologically related noun as a direct object, often in order to generate an intensive adverbial meaning, as in the following Arabic example from *Kalīla wa-Dimna*.

- (12) fa-baynamā humā ka-ḏālik ʔid ḥār-a
 and.then-while 2.DU like-DEM.DIST and.then moo.PFV-3SG.M
 ʔl-ṭawr-u ḥuwār-aⁿ šadīd-aⁿ
 DEF-bull-NOM mooing-ACC.INDF intense-ACC.INDF
 ‘and then, while they were like this, then the bull mooed intensely’
 (literally ‘... the bull mooed an intense mooing’) (Ibn al-Muqaffaʿ n.d.: fo. 30r/v)

In Example (12), the intransitive verb *ḥāra* ‘to moo’ is extended by adding the verbal noun (Ar. *maṣḍar*) of the same verb *ḥāra* as a direct object, both verb and object thus containing the same three consonants *ḥ-w-r*. This type of object is called “inner” or “absolute” object (Fischer 1972: §§376–377). As I mentioned before, there is no single instance of the *figura etymologica* construction found in the two manuscripts of the Old Spanish translation of *Kalīla wa-Dimna*. The respective passages in manuscripts A and B are the following:

¹⁸Pharies (2002: 189) states that as far as the development of parasyntetic verbal constructions is concerned, “se desarrollan en romance las llamadas estructuras parasintéticas, como *en ... ecer* y a ... *ar*, cuyo element prefijal es semánticamente vacío” (‘the so-called parasyntetic constructions were developed in Romance, just like *en ... ecer* and *a ... ar*, the prefixes of which are semantically empty’ (own translation)).

- (13) Et estando amos asy, bramo Çeñçeba / el buey muy
 and be.ACT.PTCP both like.this moo.3SG.PST.PFV PN DEF bull very
 fuerte...
 strong
 ‘and while both were like this, Çeñçeba / the bull mooed intensely’
 (Döhla 2009: A/BI.45a)

Both manuscripts display the adjective *fuerte* ‘strong’ in the function of an adverb. However, Galmés de Fuentes (1996: 201) cites the sentence “E estando amos así, bramió Çeñçeba muy fuerte bramido”, a construction that is supposed to reproduce literally the *figura etymologica*. Unfortunately, this sentence is a pure invention by Alemany y Bolufer (1915: 76), from where it was taken. Yet the Old Spanish *figurae etymologicae* can be found in scientific translations from Arabic during the second half of the thirteenth century and, even more frequently, in the Aljamiado-Morisco literature of the sixteenth and seventeenth centuries: *lloró lloramiento muy grande* ‘and he cried heavily’ (*Libro de las batallas*, 16th c.; Galmés de Fuentes 1996: 203).

5.3 Analytic causative constructions by means of *fazer*, *mandar*, and *enbiar*

The last set of verbal constructions with increased valency concerns analytic ones with the following structure in the Old Spanish version of *Calila e Dimna*: *fazer* ‘make’ + INF (Alfonso Vega 2006, Aranda Ortiz n.d., Sanaphre Villanueva 2010).¹⁹ This pattern follows one that can also be found in other Romance languages. Cases in point can be found in the following passage.

- (14) a. Et esto le faze descuydar de sy e
 and DEM.PROX OBJ.3SG.M make.3SG.PRS neglect.INF GEN himself and
 de su fazienda et faze-lo olvidar aquello
 GEN POSS.3SG affair and make.3SG.PRS-OBJ.3SG forget.INF DEM.DIST
 en que esta, et faze-le dexar la carrera
 LOC what be.3SG.PRS and make.3SG.PRS-OBJ.3SG leave.INF DEF path
 por que se ha de saluar
 through which REFL have.3SG.PRS PREP save
 ‘and this [the little sweetness of this world] makes him [the human
 being] neglect himself and his affair, and it makes him forget what

¹⁹There are also complex lexico-syntactic constructions with the pattern *fazer* COMPL CLAUSE.

the expected pattern. However, Old Spanish causatives may also be expressed by other analytic verbal constructions, among others, by *enbiar* and *mandar*, which are, together with *fazer*, the most common causative analytical verbs. A pertinent example can be found in the following passage from *Calila e Dimna*, where all three analytical constructions are present.

- (15) a. Mucho **me** **as** **fecho** **aboreçer** la **pryuança**
 much OBJ.1SG have.AUX.2SG.PRS make.PTCP hate.INF DEF proximity
 de Sençeba, e **yo enbiar le he** **dezir**
 GEN PN and 1SG send.INF OBJ.3SG.M have.AUX.1SG.PRS say.INF
 loque tengo en coraçon, et **mandar le**
 what have.1SG.PRS LOC heart and order.INF OBJ.3SG.M
he que se vaya do
 have.AUX.1SG.PRS COMPL REFL go.3SG.PRS.SBJV where
 quisiere.
 want.3SG.FUT.SBJV

‘You have made me hate a lot the proximity of Sençeba, and I will let him know what I have in mind, and I will order him to go wherever he wants to’. (Döhla 2009: AI.157a)

- b. la-qad **tarak-ta-nī** **kārih-aⁿ**
 truly-already let.PFV-2SG.M-OBJ.1SG hate.ACT.PTCP-ACC.INDF
 li-muğāwarat-i šanzaba wa-ʔana **mursil-uⁿ**
 PREP-proximity-GEN PN and-1SG send.out.ACT.PTCP-NOM.INDF
 ʔilay-hⁱ **wa-dākir-aⁿ** **lah^ū** **mā**
 to-OBJ.3SG.M and-tell.ACT.PTCP-ACC.INDF OBJ.3SG.M what
 waqaʔ-a fī nafs-ī
 happend.IPFV-IND-POSS.3SG.M LOC soul-POSS.1SG
 wa-ʔāmar-u-h^ū ʔl-laḥāq-a ḥaytu
 and-1SG;order.IPFV-IND-POSS.3SG.M DEF-entering-ACC where
 ʔaḥabb-a
 want.PFV-3SG.M

‘Truly, you already made me hate the proximity of Šanzaba, and I will send out to him and tell him what I have in mind, and I will order him to go wherever he wants to’. (Ibn al-Muqaffaʔ n.d.: fo. 45r)

Interestingly, the decision made by the translator(s) in choosing an analytical verbal construction is somewhat predetermined by the Arabic model, in that each Old Spanish causative auxiliary verb can be traced back to exactly the same verbal semantics in the Arabic model:

taraka 'to let' → *fazer* 'to make, to let'

mursil ‘sending out’ (← stem IV verb *ʔarsala* ‘to send out’) → *enbiar* ‘to send’

ʔamara ‘to order’ → *mandar* ‘to order’

Besides this, the causative verb is also followed by an infinitive form in Arabic, namely the active participle (gerund) or the *maṣḍar* (verbal noun), *kāriḥ/dākir* and *laḥāq*. The use of the infinitive can also be observed in the Old Spanish translation, except for the last sentence with *mandar*, which exhibits a subordinate clause introduced by the complementizer *que*, a structure that was on the rise in the thirteenth century (Sanaphre Villanueva 2010). However, the respective sentence in manuscript B (the passage in Example 15a is from manuscript A) also displays the infinitive after *mandar*.

- (15) c. et mandar-le-he salyr de my tierra
and order.OBJ.3SG.M-have.AUX.1SG.PRS leave.INF ABL POSS.1SG land
'and I will order him to leave my land'. (Döhla 2009: BI.157a)

First of all, it has to be stated that, to my knowledge, there are no studies concerning analytical causative constructions in Arabic. Those that deal with causatives are usually restricted to derivational verb patterns (Saad 1974, Ford 2009). However, the fact that they appear in *Kalila wa-Dimna* is certainly worthy of a separate study in the future.

As far as the Old Spanish analytic causative constructions by means of *fazer*, *mandar*, and *enbiar* are concerned, it must be mentioned that all three verbal constructions were already in use before the translation of *Kalīla wa-Dimna* into Old Spanish in 1251, since all three of them appear in *El Cantar de Mio Çid*, the Spanish national epic, often considered to be the oldest documented Spanish literary work, despite the discrepancies regarding the dating of the only existing manuscript and its chronological relationship to the actual original composition (Montaner Frutos 2011: 276–290). However, it has to be emphasized that historical research regarding the innovative emergence of causative constructions out of the Latin heritage, a language without morphological (Lehmann 2016) but with increasingly productive analytical causatives in Late Latin (Hoffmann 2018: 109), still leaves a lot to be researched and certainly deserves a dedicated in-depth study, considering the oldest vestiges of Proto-Romance and Castilian, which can be dated with accuracy.

The situation is slightly better concerning the Middle Ages, but it is still more or less limited to the studies of Alfonso Vega (2006), Davies (1995, 2000), and Sanaphre Villanueva (2010), with only the latter providing reliable statistical data for *fazer*, *mandar*, and *enbiar*, which I summarize in Table 4.

Table 4: Frequencies of *mandar*, *fazer*, and *enbiar* analytic causative constructions from the thirteenth to the sixteenth century. The absolute numbers are taken from Sanaphre Villanueva (2010: 88 and 148).

	13th c.	14th c.	15th c.	16th c.
corpus size	257,222	257,355	257,631	258,200
	words	words	words	words
<i>mandar</i>				
total instances	521	363	495	457
causative instances	389 (75%)	273 (75%)	331 (67%)	255 (55%)
<i>mandar</i> INF	292	147	223	149
<i>mandar que</i> CLAUSE	97	126	105	106
<i>mandar a</i> INF	–	–	3	–
<i>fazer</i>				
total instances	2,809	3,053	1,449	1,792
causative instances	357 (13%)	244 (8%)	221 (15%)	400 (22%)
<i>fazer</i> INF	344	239	210	371
<i>fazer que</i> CLAUSE	13	5	11	29
<i>enbiar</i>				
total instances	237	277	182	140
causative instances	61 (26%)	84 (30%)	39 (20%)	42 (30%)
<i>enbiar</i> INF	35	59	2	19
<i>enbiar que</i> CLAUSE	11	4	1	6
<i>enbiar a</i> INF	15	21	34	17
<i>enbiar a que</i> CLAUSE	–	–	2	–

As can be seen in Table 4, analytic causative constructions with *mandar* and *fazer* are far more numerous than those with *enbiar*. The fact that *mandar* exhibits higher absolute frequencies than *fazer* from the thirteenth to the fifteenth century could be linked to the respective text genres and discourse traditions that were used in the corpus, like chronicles and medieval epic poems, where constructions with *mandar* (and even *enbiar*) are expected to be quite numerous. Interestingly, Sanaphre Villanueva (2010: 4) did not consider any legal works that make frequent use of causative constructions. In fact, the impact of linguistic structures of the judicial discourse tradition on the development of the Spanish language has, so far, not been analyzed thoroughly. In the same way, the linguistic properties and traditions of the royal scriptorium of Alfonso X and its influence on contemporary and later forms of Old Spanish have never been studied systematically. This royal scriptorium not only comprises the translated texts, the linguistic structures of which are often motivated by the Arabic model, but also the bulk of legal and historical texts of the thirteenth century.

Of course, the analytic causative structures found in the Old Spanish language of *Calila e Dimna* are taken from the Romance repertoire the translator(s) were able to make use of. However, the Arabic model served as source and multiplier, so that the use of *enbiar*, *fazer*, and *mandar* increased in numbers in the given texts, themselves serving as models and sources for other Old Spanish works of the fourteenth and fifteenth centuries.²¹ Nevertheless, an overall contrastive study of causative constructions in *Calila e Dimna* in particular and in other translated texts in general is still a desideratum.

6 Conclusion

As I have shown throughout this preliminary study, this special type of learned language contact scenario, namely the translation of Arabic works into Old Spanish, offers quite a range of possibilities for the detailed study of morphosyntactical and semantic structures in the target language Old Spanish. The analyses confirm that translation is an act of negotiation (Eco 2004) and compromise (Haßler 2001: 169) and is influenced deliberately where the given repertoire of the target language permits such contact-induced structures. On other occasions (e.g., Döhla et al. 2022), I have spoken of the preexistence of a general predisposition for some structures to be accepted, which leads to an increase in frequency of

²¹Such as *Calila e Dimna* for *El conde Lucanor* by Don Juan Manuel (14th century). Both works are taken into consideration by Sanaphre Villanueva (2010).

these patterns in language-contact situations or in translated texts. The analyses of the examples in Section 3 and Section 5 have demonstrated the creative potential of the translator(s) as well as their limits.

At the same time, I have indicated throughout this chapter where there are still desiderata for future research. In this respect, this preliminary study serves as a starting point for the establishment of a large-scale project with the goal of an Arabic–Old Spanish historical contrastive grammar.

Abbreviations

JUS	jussive	PN	proper name
ACT	active	PREP	preposition

References

- Aleman y Bolufer, José. 1915. *La antigua versión castellana de Calila y Dimna*. Madrid: Suc. de Hernando.
- Alfonso Vega, Milagros. 2006. Verbos causativos. In Concepción Company Company (ed.), *Sintaxis histórica de la lengua española. Primera parte: La frase verbal*. Vol. 2, 971–1052. México: FCE/UNAM.
- Alonso, Martín. 1996. *Diccionario Medieval Español*. Salamanca: Universidad Pontificia de Salamanca.
- Alward, Ali S., Muna Y. Al-Qeili & Abdulhameed A. Ashuja'a. 2019. Translating the ten Arabic verb patterns into English: A morpho-semantic study. *Journal of Social Studies* 25(4). 113–135.
- Aranda Ortiz, Antonio. N.d. *La expresión de la causatividad en español actual*. Zaragoza: Libros Pórtico.
- Beale-Rivaya, Yasmine. 2016. At the crossroads of languages: The linguistics [sic] choices along border communities of the Reconquista in the eleventh and twelfth centuries. In Albrecht Classen (ed.), *Multilingualism in the Middle Ages and Early Modern Age: Communication and miscommunication in the premodern world*, 127–143. Berlin/Boston: Walter de Gruyter.
- Bossong, Georg. 1978. *Los Canones de Albateni*. Tübingen: Niemeyer.
- Bossong, Georg. 1979. *Probleme der Übersetzung wissenschaftlicher Werke aus dem Arabischen in das Altspanische zur Zeit Alfons des Weisen*. Tübingen: Niemeyer.
- Bossong, Georg. 2006. La sintaxis de las Glosas Emilianenses en una perspectiva tipológica. In José Luis Girón Alconchel & José Jesús de Bustos Tovar (eds.), *Actas del VI Congreso Internacional de Historia de la Lengua Española (Madrid, septiembre 2003)*, 529–543. Madrid: Arco Libros.

- Bossong, Georg. 2007. *Das maurische Spanien*. München: Beck.
- Bossong, Georg. 2008. Creatividad lingüística en las traducciones alfonsíes del árabe. *ALCANATE, Revista de estudios alfonsíes* VI. 17–38.
- Broselow, Ellen. 2000. Transfixation. In Geert Booij, Christian Lehmann & Joachim Mugdan (eds.), *Morphologie. Ein internationales Handbuch zur Flexion und Wortbildung/Morphology. An international handbook on inflection and word-formation*, vol. 1, 552–557. Berlin/New York: Walter de Gruyter.
- Cheikho, Louis (ed.). 1905. *La version arabe de Kalilah et Dimnah d'après le plus ancien manuscrit daté*. Beyrouth: Impr. Catholique.
- Comrie, Bernard. 1989. *Language universals and linguistic typology*. 2nd edn. Chicago: The University of Chicago Press.
- Comrie, Bernard & Maria Polinsky. 1993. *Causatives and transitivity*. Amsterdam/Philadelphia: John Benjamins.
- Corriente, Federico. 1997. *Poesía dialectal árabe y romance en Alandalús*. Madrid: Gredos.
- Corriente, Federico. 2002. *Gramática árabe*. Barcelona: Herder.
- Corriente, Federico. 2008a. *Dictionary of Arabic and allied loanwords*. Leiden/Boston: Brill.
- Corriente, Federico. 2008b. *Romania arábica: Tres cuestiones básicas: arabismos, "mozárabes" y "jarchas"*. Madrid: Trotta.
- Davies, Mark. 1995. The evolution of the Spanish causative construction. *Hispanic Review* 63(1). 57–77.
- Davies, Mark. 2000. Syntactic diffusion in Spanish and Portuguese infinitival complements. In Steven Dworkin & Dieter Wanner (eds.), *New approaches to old problems: Issues in Romance historical linguistics*, 109–127. Amsterdam/Philadelphia: John Benjamins.
- Dietrich, Günter. 1937. *Beiträge zur arabisch-spanischen Übersetzungskunst im 13. Jh.* Kirchhain: S/N.
- Dixon, R.M.W. & Alexandra Y. Aikhenvald. 2000. *Changing valency: Case studies in transitivity*. Cambridge: Cambridge University Press.
- Döhla, Hans-Jörg. 2008. La traducción como contacto de lenguas: El caso de las traducciones árabe-castellanas del siglo XIII. In Hans-Jörg Döhla, Raquel Montero Muñoz & Francisco Báez de Aguilar González (eds.), *Lenguas en diálogo. El iberorromance y su diversidad lingüística y literaria. Ensayos en homenaje a Georg Bossong*, 87–110. Frankfurt/Madrid: Vervuert/Iberoamericana.
- Döhla, Hans-Jörg. 2009. *El libro de Calila e Dimna (1251). Nueva edición y estudio de los dos manuscritos castellanos*. Zaragoza: IEIOP.

- Döhla, Hans-Jörg, Anja Hennemann & Benjamin Meisnitzer. 2022. Contact-induced grammaticalization – Spanish influence on Paraguayan Guaraní. In *Linguistic hybridity. Contact-induced and cognitively motivated grammaticalization and lexicalization processes in Romance languages*. Heidelberg: Universitätsverlag Winter.
- Drißner, Gerald. 2015. *Arabic for nerds*. Berlin: pochemuchka.
- Eco, Umberto. 2004. *Mouse or rat? Translation as negotiation*. London: Phoenix.
- Faulhaber, Charles B. 2004. Semitica iberica: Translations from Hebrew and Arabic into the Medieval Romance vernaculars of the Iberian peninsula. *Bulletin of Spanish Studies* LXXXI(7). 873–896.
- Fischer, Wolfdietrich. 1972. *Grammatik des klassischen Arabisch*. Wiesbaden: Harrassowitz.
- Ford, David C. 2009. The three forms of the Arabic causative. *Occasional papers in applied linguistics* 2. 1–10.
- Galmés de Fuentes, Álvaro. 1983. *Dialectología mozárabe*. Madrid: Gredos.
- Galmés de Fuentes, Álvaro. 1996. *Influencia sintáctica y estilística del árabe en la prosa medieval castellana*. 2nd edn. Madrid: Gredos.
- Haspelmath, Martin & Thomas Müller-Bardey. 2004. 107. Valency change. In Geert Booij, Christian Lehmann & Joachim Mugdan (eds.), *Morphologie. Ein internationales Handbuch zur Flexion und Wortbildung/Morphology. An international handbook on inflection and word-formation*, vol. 2, 1130–1145. Berlin/New York: Walter de Gruyter.
- Haßler, Gerda. 2001. Übersetzung als Sprachkontakt. In Gerda Haßler (ed.), *Sprachkontakt und Sprachvergleich*, 153–171. Münster: Nodus.
- Haugen, Einar. 1983. The implementation of corpus planning: Theory and practice. In Juan Cobarrubias (ed.), *Progress in language planning: International perspectives*, 269–289. Berlin et al.: Mouton.
- Hilty, Gerold. 1954. *El libro conplido en los iudizios de las estrellas*. Madrid: Real Academia Española.
- Hilty, Gerold. 2005. *Aly Aben Ragel: El libro Conplido en los Iudizios de las Estrellas. Partes 6 a 8*. Zaragoza: Instituto de Estudios Islámicos y del Oriente Próximo.
- Hitchcock, Richard. 2008. *Mozarabs in Medieval and Early Modern Spain: Identities and influences*. Aldershot/Burlington, VT: Ashgate.
- Hoffmann, Roland. 2018. *Lateinische Linguistik: Morphosyntax and Syntax in einzelsprachlicher und typologischer Perspektive*. Hamburg: Buske.
- Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties*. Washington D.C.: Georgetown University Press.
- Holmes, H. B. 1936. *An etymological vocabulary of 'Calila e Dimna'*. University of Wisconsin. (Doctoral dissertation).

- Hottinger, Arnold. 1958. »*Kalila u. Dimna*«. *Ein Versuch zur Darstellung der arabisch-altspanischen Übersetzungskunst*. Bern: Francke.
- Huffman, Henry R. Jr. 1973. *Syntactical influences of Arabic on Medieval and Later Spanish prose*. University of Wisconsin. (Doctoral dissertation).
- Ibn al-Muqaffaʿ. N.d. *Kitāb Kalīla wa-Dimna*. Paris. <https://gallica.bnf.fr/ark:/12148/btv1b10510991x/f1.item>.
- Kasten, Lloyd A. & Florian J. Cody. 2001. *Tentative dictionary of Medieval Spanish*. 2nd edn. New York: The Hispanic Seminary of Medieval Spanish.
- Lehmann, Christian. 2016. Latin causativization in typological perspective. In Muriel Lenoble & Dominique Longree (eds.), *Actes du 13ème Colloque International de Linguistique Latine, Bruxelles, 4–8 avril 2005*, 918–943. Louvain/Paris: Peeters (BEC).
- Lieber, Rochelle. 2009. *Introducing morphology*. Cambridge: Cambridge University Press.
- López García, Ángel. 2000. *Cómo surgió el español. Introducción a la sintaxis histórica del español antiguo*. Madrid: Gredos.
- Montaner Frutos, Alberto. 2011. *Cantar de Mio Cid*. Madrid: Real Academia Española.
- Müller, Bodo. 1987. *Diccionario del español medieval*. Heidelberg: Winter.
- Al-Munğid. 2005. *Al-Munğidu fī 'l-luğati wa-'l-ʔaʕlāmi*. Beirut: Dār al-Maʕriq.
- Neumann-Holzschuh, Ingrid. 1997. *Die Satzgliedanordnung im Spanischen. Eine diachrone Analyse*. Tübingen: Niemeyer.
- Payne, Thomas E. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.
- Pérez, Maurilio. 2010. *Lexicon latinitatis medii ævi regni legionis (s. VIII–1230) imperfectum*. Turnhout: Brepols Publishers.
- Pérez, Raoul Monserrate. 1943. *Vocabulario clasificado de 'Kalila et Digna'*. University of Chicago. (Doctoral dissertation).
- Pharies, David. 2002. *Diccionario etimológico de los sufijos españoles*. Madrid: Gredos.
- Real Academia Española. N.d. *Banco de datos [en línea]. Corpus diacrónico del español*. <https://corpus.rae.es/cordenet.html>.
- Rodríguez M. Montalvo, Sagrario. 1981. *Lapidario*. Madrid: Gredos.
- Saad, George N. 1974. Causatives and passives in Arabic. *An-Nashra* 7(1). 22–31.
- Sakel, Jeanette. 2007. Types of loan: Matter and pattern. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 15–29. Berlin/New York: Mouton de Gruyter.
- Sanaphre Villanueva, Mónica. 2010. *Analytic causative constructions in Medieval Spanish: The origins of a construction*. Rice University. (Doctoral dissertation).

- Shibatani, Masayoshi. 1976. The grammar of causative constructions: A conspectus. In Masayoshi Shibatani (ed.), *The grammar of causative constructions*, 1–40. New York: Academic Press.
- Shibatani, Masayoshi & Prashant Pardeshi. 2002. The causative continuum. In Masayoshi Shibatani (ed.), *The grammar of causation and interpersonal manipulation*, 85–126. Amsterdam/Philadelphia: John Benjamins.
- Simonet, Francisco Javier. 1983. *Historia de los mozárabes de España*. Madrid: Turner.
- Song, Jae Jung. 1996. *Causatives and causation: A universal-typological perspective*. London/New York: Longman.
- Song, Jae Jung. 2001. *Toward a typology of causative constructions*. München: Lincom Europa.
- Stinson, Bobby Ray. 1967. *An etymological lexicon of manuscripts A and B of Calila e Digna*. Chapel Hill: University of Carolina. (PhD Thesis).
- Tallgren, O. J. 1934. Acerca del literalismo árabe-español de la astronomía alfoncina. *Al-Andalus* II. 223–225.
- Väänänen, Veikko. 1978. *Introduction au latin vulgaire*. 3rd edn. Paris: Klincksieck.
- Wehr, Hans. 1952. *Arabisches Wörterbuch für die Schriftsprache der Gegenwart*. Wiesbaden: Harrassowitz.
- Wehr, Hans. 1976. *A dictionary of Modern Written Arabic*. Ithaca, NY: Spoken Language Services.
- Wright, Roger. 2000. *El tratado de Cabrerios (1206). Estudio sociofilológico de una reforma ortográfica*. London: Department of Hispanic Studies / Queen Mary & Westfield College.
- ʿAzzām ʿAbdu, ʿAbdu ʿl-Wahhāb & Ṭarḥīn Bek (eds.). 1941. *Kitāb Kalīla wa-Dimna*. Al-Qāhira: Maṭbaʿa al-Maʿārif wa-Maktabatuhā bi-Miṣr.

Discussion

Chapter 12

Reconciling the global and local in language contact

Yaron Matras

Aston Institute for Forensic Linguistics, Birmingham, and Department of Hebrew Language, University of Haifa

Contributions to this collection cover a continuum of phenomena broadly located between two poles of language contact perspectives: The idea of the multilingual repertoire as a dynamic and fluid pool of resources which individuals deploy at their meaningful discretion, and the social conventions and values that shape users' attitudes to languages and the hierarchical relations between them. Linked to those are the transformations in the use and shape of individual structural features: The replication of word forms with their phonetic characteristics from another language, the re-configuration of morphological and syntactic constructions blending features from more than one language, the re-combination of word forms with scripts, and more. This chapter deals with the dichotomy between 'named languages' and the practice that has been described as 'translanguaging'. It proposes an integrated model of language contact. At the core of the model is the view of contact as a balancing act of pull factors that impact strategies to manage an integrated repertoire of features. Structural changes in language (contact induced changes, convergence, or borrowing) are understood as new local practice routines that can be disseminated and shared across a practice community. As variables the model takes into consideration the functional properties of individual structural categories and the motivation to innovate practice around them, patterns of action routines and types of talk, as well as discourses about language.

1 Introduction

In 2011 a new question was introduced into the UK national census asking respondents to specify their 'main language'. The results gave the first statistical



picture of the diversity of languages spoken in the country. But the question was not designed to capture the country's multilingualism. In fact, it was worded and formatted in such a way that would effectively obscure valuable information on multilingualism (cf. Matras & Robertson 2015): A hierarchy was introduced by asking respondents to name another language only if their 'main language' was not English. That will have excluded the home languages of those who took 'main language' to mean the language spoken most hours of the day, for example, at work or a place of study, or the language favoured with peers or media. Respondents who opted to name a 'main language' other than English could only name a single language. The reality of multilingual repertoires could not be captured. Census takers were then asked to self-assess their proficiency in English. The overall purpose of the question was evidently to give authorities an indication of the extent to which speakers of other languages had lower levels of English. In the public discourse of the conservative political establishment that had been in government in the UK since 2010, use of other languages was linked to lack of integration and productivity and even to radicalisation and ideological extremism. Multilingualism was seen as a potential citizenship deficiency.

2 Counting languages vs translanguaging

At the same time the numbers tell a story. First, they represent the action of self-declaring another language in full awareness of the pressures to conform. Indirectly and unintentionally the census offers a platform for declarative agency or what Stroud (2018) calls 'acts of linguistic citizenship'. They also show some trends: In 2011 around 7.7 per cent of the population of England and Wales declared a 'main language' other than English. Ten years on, in the 2021 Census, the figure had risen to 8.9 per cent. Yet the proportion of those who declared that they did not know English remained exactly the same, at 0.3 per cent of the population. These figures demystify the notion that being multilingual necessarily comes at the expense of knowing English, particularly since it is likely that many respondents, for the reasons explained above, under-reported their use of other home languages.

Nevertheless, the representation of languages as countable, discrete entities skews the reality of many households. In 2013, we produced a short film about the languages of Manchester.¹ In an allusion to the census question we asked one interviewee what her 'main language' was. She responded with little hesitation: "Main language? Every day I speak three languages at the same time!" For

¹<https://www.youtube.com/watch?v=pmTDzsPrBp8> (Last accessed: 4 December 2023).

multilingual persons the question framed in a monolingual mindset is counter-intuitive. In the Northwest of England, the most widespread community language is Panjabi (and closely related variants Mirpuri, Pahari, and Potwari). Yet the highest census numbers for a ‘main language’ other than English in the region appear for Urdu. Forced to choose one over the other(s) people who originate from Pakistan preferred to declare a publicly recognised, official written language over regional varieties. Speakers of Romani almost never declared their home language. They listed the national languages of their respective countries of origin in central and eastern Europe, usually assuming that their language would not qualify as ‘main’ since it lacks institutional status and often not even knowing what the English term for the language is (they call their language *rroman*, lit. ‘in the manner of the Roms’).

The census is a blatant example of the way perception of language is shaped by nation-state ideologies. Some authors have been strongly critical of any enumeration of languages, referring to it polemically as ‘linguistic accounting’, ‘demolinguistics’, or ‘headcount of languages’ (see Pennycook & Otsuji 2015: 19–49; King 2016: 187–188; Stevenson 2017: 56–64). They juxtapose the listing of languages to first-hand investigations of linguistic practices, sometimes referred to as ‘linguaging’. The focus on practice is strengthened by an appreciation that increased mobility and new forms of mediality and institutional participation create ever more complex domains of interaction. Captured by concepts such as ‘ethnoscapes’ and ‘super-diversity’ (Appadurai 1992; Vertovec 2007) the multiplicity of interaction options leads to a lower degree of predictability of links between language, place, identity, and community, with methodological implications for the analysis of relations between linguistic forms, participants, place, and institutions (cf. Blommaert 2010; Blommaert & Rampton 2011; Arnaut et al. 2016; Arnaut et al. 2017). Terms such as ‘translanguaging’, ‘metrolingualism’, ‘heteroglossia’, and ‘crossing’ have been used to capture the dynamic fluidity of moves among linguistic forms (Rampton 1995; Blackledge & Creese 2010; García & Wei 2014; Pennycook & Otsuji 2015; Wei 2018). The notion of ‘translanguaging’ in particular has been celebrated almost with a sense of triumphalism: It stands for a paradigm shift that not only replaces the view of languages as fixed entities with clear demarcation boundaries but also calls for social engagement and intellectual resistance against ideologies that foster that view (cf. Creese & Blackledge 2018; Moore et al. 2020).

This critical, post-structuralist view of the links between language and social representations aligns itself with an established strand that theorises the use and processing of multiple languages: Multilingualism is not the added accumulation of several monolingual modes (Grosjean 1989). Instead, it is a complex set of

features (Jørgensen 2008) blended together in an individual's overall repertoire. That repertoire includes acquired norms and conventions according to which in a given interaction context features and sets of features are selected and others inhibited (Matras 2009/2020; see also Green 1998). Such notions of repertoire problematise 'language' as a pre-determined set of structures and view it instead as a dynamic, emerging pattern of practices, detaching it from fixed notions of pre-defined groups or speech communities and viewing groups as emerging and evolving networks of practice and people as moving in between and among them (Busch 2012; Blommaert & Backus 2013). The view of language contact as one closed system interfering with another has been replaced by a view in which plurality of form is the default and closed systems or 'named languages' are derived social constructions.

3 Contact, categories, and repertoire management

The very premise of contact linguistics questions a founding principle of modern historical linguistics, namely the idea that languages are only pre-destined to diverge from one another. Contacts between populations and the multilingualism that they create increase similarities between languages and can lead to convergence (Trubetzkoy 1928). Already the earliest examinations of the effect of contact questioned whether some forms were exempted entirely or partially from such processes (Whitney 1881). In due course attention was given to the likelihood that some components were more easily 'borrowed' from one language into another (Thomason & Kaufman 1988). The question was asked whether typological parameters could predict and account for ease of borrowing (Moravcsik 1978; Campbell 1993; Stolz & Stolz 1997; Matras 2007) and whether those in turn might reveal something about the inner functions of language categories and their status within the speech production process itself (Myers-Scotton & Jake 2000, Matras 2009/2020).

Frequency-based trends are not always meaningful: Nouns can be at the top of the borrowability hierarchy simply because they are the most common category in most languages and because they represent new objects and concepts that enrich semantic expression when cultures come into contact. But an implicational hierarchy such as 'but > or > and' (where '>' indicates greater likelihood of borrowing, and the implicational arrangement suggests that the presence of higher elements on the hierarchy is a pre-condition for lower elements) calls for an explanatory model (Matras 1998): If the semantic-pragmatic operation of contrast outranks that of disjunction, and the latter outranks addition, in their

respective susceptibility to borrowing, then that will have its roots in the ease with which bilingual users maintain a separation of forms by ‘language’ around the respective processing operation, or, instead opt to give up that separation in favour of generalising just a single form (usually the one that can be used in both in- and out-group communication, or simply in a wider set of interaction settings). The hierarchy, attested universally (cf. Hober 2022; Stolz et al. 2021; Grant 2012; Matras 2007), suggests that on the cline, the operation of contrast is more likely to serve as a trigger to users to give up separation of features within the repertoire. It is in other words more difficult to maintain such separation around the function of contrast, with implications for the processing of broken causal chains and turn management. The historical event that we regard as ‘borrowing’ between languages is thus seen and explained as triggered by factors that involve the management of a complex repertoire of features, across a complexity of ever evolving interaction settings. Borrowing is in reality a change in practice routines. It is the product of an innovation that is gradually propagated across a network of language users or a practice community. Through the inertia of our structuralist intellectual upbringing, we tend to view the outcome of that change in practice routine as a structural change in a particular named language which we label ‘borrowing’.

4 Towards an integrated model

I have so far alluded to two major innovative developments in the emergence of an epistemology of language contact: The move away from examining closed, self-contained ‘systems’ in contact and onto appreciating the existence of a complex and dynamic, wholesale repertoire of features; and the realisation that the fate of forms and features in linguistic settings that are complex and dynamic depends on their category status, i.e. on their function in the mental processing of information and knowledge and the structuring of interactional turns. Work within the translanguaging paradigm, innovative as it may be, has so far been rather reluctant to engage with the structural transformations that language contact brings about. Structural and typological approaches in their turn have by and large shied away from embracing critical approaches that seek to deconstruct the idea of language as a fixed ‘system’. The empirical contributions to the present volume show that there is a reality behind both: The dynamic fluctuation of features in an individual’s repertoire, and the metalinguistic perception of ‘languages’ as emblems of identity that can be enumerated, labelled, and evaluated and whose integrity can be either carefully maintained or intentionally disrupted and interrogated.

For this reason, we need an integrated theory of language contact. Such a theory must be explicitly equipped to account for contact-induced structural change in terms of the factors that motivate users to alter their practice routine when managing their complex repertoire of features. It must also account for the factors that motivate networks of users to converge around innovative practices. It must recognise that linguistic practice is driven not merely by aesthetic attributes that are associated with individual features but also by the function that different features and categories of features assume in the process of knowledge transfer and knowledge processing that is at the core of communicative interaction. At the same time, it must also acknowledge that users' practice routines can themselves become the subject of discourse – labelled, enumerated, and qualified.

In Matras (2009/2020; see also Matras 2021b) I outline the principles of such an integrated theory: Users manage their repertoire of linguistic features balancing three pull-factors: a) the wish to accommodate to context-bound expectations on the part of the listener by selecting those features that are deemed purposeful and permissible while inhibiting those that are not; b) exploiting the full expressive potential of the repertoire, making use of as many features (including word forms, constructions, suprasegmental features, and discourse-management routines) as possible to maximise expressiveness; and c) managing processing load effectively by reducing where possible the need to deploy the selection and inhibition mechanism by generalising features across interaction contexts and settings ('levelling'). The balancing act is a constant one, prompting users to negotiate and re-negotiate the choice of features locally, i.e. in each and every interaction and often utterance. Yet it is also guided by the conventions of established practice routines. The latter can be subject to more global meta-discourses that may contain and constrain individual users' flexibility to deploy features in a way that arises directly from the local balancing act. When altering practice routines users can draw on at least two distinct strategies: The deployment of linguistic 'matter' (phonological shapes or forms) and of linguistic 'pattern' (form-meaning relationship). For some constructions the choice is constrained by their very nature: Combinations of words are always 'patterns'. For others, both can be options: Definite articles and some lexical items can be replicated as word-forms or calqued through grammaticalisation or semantic extension, respectively. Attitudes to language may come into play: Nativising a form through pattern replication may be preferred as a way of preserving the integrity of a set of features that are associated with a particular interaction setting. Finally, the role of a feature and its category or function value in the processing of knowledge in communicative interaction will determine or partly determine its susceptibility to borrowing: Users are motivated to eliminate the need to select and inhibit, i.e.

to choose among functionally equivalent or near-equivalent items, and maintain the separation of sub-sets, when the processing burden is most intense. That accounts, among other things, for the high borrowability of expressions of contrast and of elements that help monitor and direct the interaction such as discourse markers ('pragmatic' markers). Similarly, the motivation to generalise features across the repertoire is greater when those features represent unique or particular knowledge spaces. This accounts, among other things, for the borrowability of so-called 'cultural loans', including names of institutions and culture-specific practices.

5 Usage and ideology on a continuum

Before I return to elaborate on the theory and its variables, I wish to reference the contributions to this volume and the way in which they demonstrate the existence of a continuum between flexible and dynamic repertoire management and the social reality of metalinguistic discourses about named languages and the social construction of demarcation boundaries around them.

Strict demarcation boundaries among languages are identified in interviews that elicit attitudes and in institutional settings where language choices are strictly defined. Sandra Schlumpf describes how metalinguistic discourses reveal the acceptance of hierarchies among named languages in Equatorial Guinea. Users associate languages with the practice routines in which they are deployed, attributing values accordingly. In effect, users differentiate (and label) different sets of features within their overall repertoires according to the communicative practices that they represent. Whether a language – an identifiable set of features that carries a label – has high or low prestige appears to be in part indexical to the settings and contexts in which it is deployed, particularly when compared to other languages whose deployment in institutional settings might be considered superior. Nancy Hawker discusses how Palestinian Arabic in Israel is viewed as an in-group language. The display of multilingualism, i.e. of proficiency in both the in-group language and in Hebrew, the majority and principal state language, is regarded as a valorisation of assets. In effect, it testifies to the user's greater ability to assume flexibility among different communicative practices in a variety of settings, particularly in institutional settings as well as addressing different user networks. In the social-political context of Israeli society, user networks are associated with different populations and a cultural and political boundary. Switching language can signal defiance or audience selection; in other words, it can disrupt or accommodate to established practice routines. As Fabio Gasparini

describes, repertoires are subject to historical changes and the Bəṭaḥrēt language is losing ground in Oman as Arabic infiltrates all domains of communication in the local community. This may be regarded as a consequence of the infiltration of nation-state ideologies into a community that has been situated on the fringe of such ideologies for much of its history. Changing ideologies leads directly to radical changes in practice routines.

The contrast of languages can be meaningful, in Gumperz' (1982) terms, also within a single interaction. Jacopo Falchetta describes the associations of repertoire elements with different social contexts. That allows users to exploit meaningful contrasts as a socio-pragmatic function in the use of Moroccan Arabic and French. While each set of features represents the sum of interaction settings with which they are linked, the alternation among them is itself a social determinant of a user's background, notably the user's ability to deploy a complexity of feature sets, testifying to their immersion in multiple networks of users and multiple practice settings. Similarly, Marta Rodríguez García describes how in Gibraltar Yanito is a permanent negotiation of repertoire components and feature sets. In fact, the label itself captures users' perception of their alternating deployment of sets of features that are otherwise, in institutional settings, considered to be separate languages, as an integrated whole. The practice routine of selecting features at users' discretion within the same interaction and network of users is acknowledged in the meta-discourse as a variety in its own right.

In her discussion of Sofia's linguistic landscapes Emilia Slavova shows how writing systems, normally subjected to more tightly regulated and institutionalised language use, can also be deployed at users' discretion as part of the resources of the complex linguistic repertoire. They become combinable in new ways with word forms associated with different written languages. If monolingualism is considered in public discourse to be the norm that is linked to nation states and language education, and a means of conformity, valorising the individual and good citizenship, then the absence of inhibition when deploying repertoire features to maximise creative expressiveness is in some ways an act of defiance (Slavova mentions usage "in unexpected ways"), one through which users assume agency to draw on past experience but subvert existing routines and give legitimacy to new forms of practice (cf. Emirbayer & Mische 1998; Liddicoat & Taylor-Leech 2020). I am reminded of a young trilingual child's theatrical use of a one-off mixed utterance, using English and German in an interaction context that is normally reserved for Hebrew (Matras 2009/2020: 38): "Aba, where do I get a *Lappen* so I can *wisch* my *Gesicht*?" – 'Daddy, where do I get a *wash cloth* so I can *wipe* my *face*?' The subversion of the routine is two-layered: First, in the unexpected choice of English as the predication language of the utterance when

addressing an interlocutor with whom Hebrew is the established routine; and second, in inserting lexical items from German into the English utterance. Below I will return briefly to this aspect of performativity and its role in explaining certain types of language contact outcomes.

Klaudia Dombrowsky-Hahn and Axel Fanego Palat discuss a German West African woman's deployment of various features of her repertoire in a way that does not always conform to the expected monolingual norm but enables communication in that it shows creative agency in forming constructions drawing on an array of resources. The replication of French impersonal constructions when speaking German, the generalisation of a preposition with a variety of verbs independently of the direction of motion, and the use of modifier-head juxtaposition to express possession can all be approached as an emerging practice that allows the user to navigate language-learning in settings shaped by frequent mobility and a range of participation networks. The outcome is a set of highly individualised innovations and features. Linda Bäumlér's chapter lends further insights into the potential cumulative effect of such choices shaped by participants' networks of practice: Whether or not English sounds are directly replicated in loanwords or replaced by Spanish equivalents has to do with a sense of 'affinity', that is, awareness of and appropriation of the wholesale repertoire. Affinity can emerge and be reinforced through factors such as interaction experience or exposure to media. It is in part an emotional state rather than a strictly objective circumstance such as geographical proximity, all the more so since in a globalised world immersion in practice routines can be remote and mediated. Miriam Neuhausen similarly shows how a process of phonological change in a minority or diaspora language aligns itself with a parallel process in the contact language: Users seem to blend together their 'management' of linguistic resources, adopting a change wholesale irrespective of the named language, in other words, indiscriminately in all settings and with all sets of interlocutors with whom the relevant sound pattern is used. Here too, the extent of the change in Pennsylvania German is linked to the extent of exposure to settings in which English is used.

Hans-Jörg Döhla's discussion of a 'learned language contact scenario' shows how such individual choices are triggered by a comparable sense of affinity. The translators seek to preserve the aesthetic appearance of the model text where a juxtaposition of factive and causative is used as a recurring template. It is the aesthetic format that becomes a repertoire feature, one that is preserved when addressing a separate audience of readers through the creative process of exhausting the expressive potential of Spanish verb derivational constructions. Codified

by the translated texts, innovation is then disseminated among the practice community of readers to become a characteristic feature of the genre. The process bears resemblance to Inga Hennecke's account of pragmatic markers in Manitoba French. Here, too, a practice community emerges, albeit in spoken discourse rather than through the reception of a written genre, characterised through the way certain features of the complex repertoire are managed. We have a good example of the inherent link between repertoire management and the role of certain functional categories in the motivation to generalise the mapping of meaning to form (French *comme* adopting the wider meanings of English *like*) and of actual word forms (English *so*) across the repertoire. The author hints that the wish to 'level' these features is in some sense pre-determined by the properties of the category of pragmatic markers: Their semantic and syntactic detachability as well as difficulties in translation equivalence.²

6 From practice to contact languages: Variables of a theory

The case studies show us how users' linguistic repertoires comprise word forms, form-meaning mappings and constructions, phonetic articulations as well as experiences, values and attitudes, all of which are associated with a range of experiences of various communicative interaction settings. These features can cluster in partly distinguishable sets that are subject to more or less strict selection control, but they can also be re-grouped into new configurations either at the level of local interaction or that of the individual user's emerging practice preference, and be disseminated among a practice network involving other users. Users may or may not exploit the contrastive potential of the affinity between features and associated interaction settings as a means to express difference and to replicate or interrogate social hierarchies.

The labelling and enumeration of languages, and the realisation that users engage in the practice of languaging drawing on their full repertoire of features, are therefore not at all mutually exclusive and so they need not constitute theoretically juxtaposed perspectives. They are instead complementary, provided we

²Hennecke takes issue with the explanation I provided in Matras (2020) where I traced lapses in the selection and inhibition mechanism, i.e. bilingual speech production errors, to cognitive factors. But, of course, we are dealing with different kinds of data here, and since Hennecke's chapter does not address such lapses in control, the cognitive triggers may not be evident. Still, I would argue that local shifts in meaning and distribution of individual markers have their roots in one-off lapses of control where repertoire features are blended. The stage that Hennecke examines is one in which such occurrences have become accepted and conventionalised.

can view both through a shared and integrated theoretical lens. Returning to the model of repertoire management briefly outlined above (cf. Matras 2009/2020), I would like to propose that such an integrated theory must give consideration to a differentiated set of variables. These include a) structural features (constructions, words, morphs, phones, suprasegmentals), b) linguistic action routines (institutional forms of discourse, pragmatic organisation of discourse modes, roles, turn taking, types of talks and distribution of illocutions), c) social and institutional settings that impact language and give rise to an array of language practice routines, and d) metalinguistic awareness and discourses about language. To a considerable extent these are also the elements that are alluded to in discussions of the super-diverse repertoire (Busch 2012; Blommaert & Backus 2013) and in those of language ecology (Pennycook 2010). The challenge as I see it is to link these dimensions explicitly to a theory of structural change in contact situations, explaining change as the product of innovations in the management of repertoire features, and explaining innovations as motivated by the goals of communication and the different procedures of mental processing of information that are triggered by different kinds of structural categories.

If we return to the ‘mixed’ utterance of the trilingual child quoted above, here the conscious defiance of the practice routines of feature separation is performative, aiming to achieve a particular effect on the listener and so on the relationship between speaker and listener at a given moment in the interaction. The structure of the utterance strongly resembles the conventionalised patterns that have been labelled ‘Mixed Languages’ (Bakker & Matras 2003): These are languages that display contact outcomes that are deemed to be unconventional, combining, for instance, grammatical inflection from one source language with core lexicon from another, or nominal inflection from one with verb inflection from another, or borrowing wholesale function word paradigms such as pronouns which are normally not borrowed wholesale from one language to another. For that reason they are deemed worth of the explicit label of being ‘mixed’. Traditional historical linguistic approaches have defined Mixed Languages in terms of the genetic tree-model of language diversification as languages whose genetic ancestry cannot be determined (Thomason & Kaufman 1988). Yet if our ‘critical’ or post-structuralist approach to language contact prompts us to abandon the tree-model as the principal prism through which we view language change and to think instead of convergence and re-configuration as a default, we must find a new way to conceptualise Mixed Languages. I propose that Mixed Languages arise from the performative practices of combining features that are not normally combined through the everyday pull factors (maximising expressive potential or easing the burden of processing load – the two principal motivations to

introduce permeations into the management of the complex repertoire). Instead they are quintessential expressions of agency, subverting established routines in order to perform new identities (cf. Matras 2021a). Our integrated model of repertoire management thus allows us to link so-called ‘translanguaging’ practice with the perception of language boundaries, the functional value of structural categories (and the ease with which they are generalised across interaction settings or ‘named languages’), forms of illocution, agency, and the perpetuation of new practices through dissemination across an emerging practice community. In other words, Mixed Languages are exceptional because they arise from utterances that are purposefully defiant of everyday patterns of mixing. They are products of a particular mode of repertoire management.

In a similar vein, Creoles, traditionally viewed in historical linguistics as the expansion of pidgins formed out of a need for restricted communication (but see critique of that view in Mufwene 2021) can be viewed as features of the repertoires of many individuals, adopted as an emerging shared practice in a newly formed practice community. It is not their ‘genetics’ that give them substance but rather their constitution as an assembly of features that enable communicative interaction in new and emerging settings. An integrated theory of language contact is challenged to link the global functions of individual structural categories to local processes of repertoire management, and in that way to account for new and changing action routines.

References

- Appadurai, Arjun. 1992. Global ethnoscaples: Notes and queries for a transnational anthropology. In Richard G. Fox (ed.), *Interventions: Anthropologies of the Present*, 191–210. Santa Fe: School of American Research.
- Arnaut, Karel, Jan Blommaert, Ben Rampton & Massimiliano Spotti (eds.). 2016. *Language and Superdiversity*. New York/London: Routledge.
- Arnaut, Karel, Martha Sif Karrebæk, Massimiliano Spotti & Jan Blommaert (eds.). 2017. *Engaging superdiversity. Recombining spaces, times and language practices*. Bristol: Multilingual Matters.
- Bakker, Peter & Yaron Matras (eds.). 2003. *The mixed language debate. Theoretical and empirical advances*. Berlin: Mouton de Gruyter.
- Blackledge, Adrian & Angela Creese. 2010. *Multilingualism: A critical perspective*. London: Continuum International.
- Blommaert, Jan. 2010. *Sociolinguistics of globalization*. Cambridge: Cambridge University Press.

- Blommaert, Jan & Ad Backus. 2013. Superdiverse repertoires and the individual. In Ingrid de Saint-Georges & Jean-Jacques Weber (eds.), *Multilingualism and multimodality. The future of education research*, 11–32. Rotterdam: Sense Publishers.
- Blommaert, Jan & Ben Rampton. 2011. Language and superdiversity. *Diversities* 132. 1–21.
- Busch, Brigitte. 2012. The linguistic repertoire revisited. *Applied Linguistics* 33(5). 503–523.
- Campbell, Lyle. 1993. On proposed universals of grammatical borrowing. In Henk Aertsen & Robert J. Jeffers (eds.), *Historical linguistics 1989*, 91–109. Amsterdam: John Benjamins.
- Creese, Angela & Adrian Blackledge (eds.). 2018. *The Routledge handbook of language and superdiversity*. London: Routledge.
- Emirbayer, Mustafa & Ann Mische. 1998. What is agency? *American Journal of Sociology* 103(4). 962–1023.
- García, Ofelia & Li Wei. 2014. *Translanguaging: Language, bilingualism and education*. London: Palgrave Macmillan UK.
- Grant, Anthony. 2012. Contact, convergence and conjunctions: A cross-linguistic study of borrowing correlations among certain kinds of discourse, phasal adverbial and dependent clause markers. In Claudine Chamoreau & Isabelle Léglise (eds.), *Dynamics of contact-induced language change*, 311–358. Berlin: De Gruyter Mouton.
- Green, David. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1. 67–81.
- Grosjean, François. 1989. Neurolinguists beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36. 3–15.
- Gumperz, John J. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.
- Hober, Nicole. 2022. On the borrowing of the English adversative connector *but*. In Natalya Levkovych (ed.), *Susceptibility vs resistance. Case studies on different structural categories in language-contact situations*, 183–257. Berlin: De Gruyter.
- Jørgensen, J. Normann. 2008. Polylingual languaging around and among children and adolescents. *International Journal of Multilingualism* 5(3). 161–176.
- King, Lid. 2016. Multilingual cities and the future: Vitality or decline? In Lid King & Lorna Carson (eds.), *The multilingual city: Vitality, conflict and change*, 179–202. Clevedon: Multilingual Matters.
- Liddicoat, Anthony J. & Kerry Taylor-Leech. 2020. Agency in language planning and policy. *Current Issues in Language Planning* 22(1–2). 1–18.

- Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36. 281–331.
- Matras, Yaron. 2007. The borrowability of grammatical categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 31–74. Berlin: Mouton de Gruyter.
- Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.
- Matras, Yaron. 2020. *Language contact*. 2nd edn. Cambridge: Cambridge University Press.
- Matras, Yaron. 2021a. Repertoire management and the performative origin of Mixed Languages. In Maria Mazzoli & Eeva Sippola (eds.), *New perspectives on mixed languages*, 362–403. Berlin: de Gruyter Mouton.
- Matras, Yaron. 2021b. Theorizing language contact. In Richard Janda, Brian Joseph & Barbara S. Vance (eds.), *The handbook of historical linguistics*, 375–392. Oxford: Wiley Blackwell.
- Matras, Yaron & Alex Robertson. 2015. Multilingualism in a post-industrial city: Policy and practice in Manchester. *Current Issues in Language Planning* 16(3). 296–314.
- Moore, Emilee, Jessica Bradley & James Simpson (eds.). 2020. *Translanguaging as transformation. The collaborative construction of new linguistic realities*. Bristol: Multilingual Matters.
- Moravcsik, Edith. 1978. Universals of language contact. In Joseph H. Greenberg (ed.), *Universals of human language*, 94–122. Stanford: Stanford University Press.
- Mufwene, Salikoko. 2021. Creoles and pidgins. Why the latter are not the ancestors of the former. In Evangelia Adamou & Yaron Matras (eds.), *The Routledge handbook of language contact*, 300–324. London: Routledge.
- Myers-Scotton, Carol & Janice Jake. 2000. Four types of morpheme: Evidence from aphasia, code-switching, and second language acquisition. *Linguistics* 38. 1053–1100.
- Pennycook, Alastair. 2010. *Language as a local practice*. Milton Park: Taylor & Francis.
- Pennycook, Alistair & Emi Otsuji. 2015. *Metrolingualism: Language in the city*. Abingdon: Routledge.
- Rampton, Ben. 1995. *Crossing: Language and ethnicity among adolescents*. London: Longman.
- Stevenson, Patrick. 2017. *Language and migration in a multilingual metropolis. Berlin lives*. Houndmills: Palgrave Macmillan.

- Stolz, Christel & Thomas Stolz. 1997. Universelle Hispanismen? Von Manila über Lima bis Mexiko und zurück: Muster bei der Entlehnung spanischer Funktionswörter in die indigenen Sprachen Amerikas und Austronesiens. *Orbis* 39. 1–77.
- Stolz, Thomas, Deborah Arbes & Christel Stolz. 2021. *Pero* – Champion of Hispanization? On the challenges of documenting function word borrowing in Mesoamerican languages. In Danae Maria Perez & Eeva Sippola (eds.), *Post-colonial language varieties in the Americas*, 17–53. Berlin: De Gruyter Mouton.
- Stroud, Christopher. 2018. The multilingual citizen: Towards a politics of language for agency and change. In Lisa Lim, Christopher Stroud & Lionel Wee (eds.), *The multilingual citizen: Towards a politics of language for agency and change*, 17–39. Bristol: Multilingual Matters.
- Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Trubetzkoy, Nikolai S. 1928. Proposition 16. *Acts of the First International Congress of Linguists, Leiden*. 17–18.
- Vertovec, Steven. 2007. Super-diversity and its implications. *Ethnic and Racial Studies* 30(6). 1024–1054.
- Wei, Li. 2018. Translanguaging as a practical theory of languages. *Applied Linguistics* 39(1). 9–30.
- Whitney, W. D. 1881. On mixture in language. *Transactions of the American Philological Association (1869–1896)* 12. 1–26.

Name index

- Aalberse, Suzanne, iv
Aaron, Jessi Elana, 248
Abbassi, Abdelaziz, 40, 44
Aboh, Enoch Oladé, iii
Abu Arshid, S., 172
Acuña Ferreira, A. Virginia, 88
Ahlers, Wiebke, 103
Aijmer, Karin, 245, 246, 261
Aikhenvald, Alexandra Y., 136, 310
Aixelà-Cabré, Yolanda, 216, 227
Al-Munġid, 312
Alemany y Bolufer, José, 320
Alfonso Vega, Milagros, 320, 324
Allan, Kori, vii
alMahri, Abdullah M., 141
Alonso, Martín, 315
alTabuki, Salim B., 132
Alward, Ali S., 307
Amer, Faten, 157
Amnesty International, 175
Andersen, Gisle, 245
Anderson, Vicki Michael, 100
Appadurai, Arjun, 140, 335
Applbaum, Kalman, 174
Aranda Ortiz, Antonio, 320
Arends, Jacques, iii
Arnaut, Karel, 335
Ash, Richard, 104
Asociación de Academias de la
 Lengua Española, 276, 292
Atanassova-Divitakova, Svetlana, 8
Auer, Peter, 66, 67, 74, 81, 92
Ayalon, Hanna, 174
Backus, Ad, 336, 343
Baker-Smemoe, Wendy, 100, 109
Bakker, Peter, iii, 343
Barbosa da Silva, Diego, 231
Barnes, Sonia, 103
Barontini, Alexandrine, 56
Barry, Alpha O., 188, 189
Bates, Douglas, 112, 283
Bäumler, Linda, 277
Beale-Rivaya, Yasmine, 303
Beeching, Kate, 261
Bek, Tarḥīn, 305, 315, 317
Ben-Rafael, Eliezer, 157
Beniak, Edouard, 246–248, 255, 260,
 267
Benor, Sarah Bunin, 101
Bentahila, Abdelâli, 38, 40–42, 44, 56
Berruto, Gaetano, 66, 138
Bettega, Simone, 131
Blackledge, Adrian, vii, 175, 335
Blench, Roger, 190
Blestel, Élodie, iv
Block, David, 161
Blommaert, Jan, iv, v, 3–5, 8, 220,
 335, 336, 343
Boersma, Paul, 282
Bohmann, Axel, 103
Bolden, Galina B., 255
Bolekia Boleká, Justo, 227, 233
Bolinger, Dwight, 261

- Bolker, Ben, 112
Bolly, Catherine, 254
Bonfiglio, Thomas P., 143
Bossong, Georg, 302, 303, 305, 306, 313
Bouchard, Marie-Eve, 221
Boumans, Louis, 47
Bourdieu, Pierre, 153, 154
Boutieri, Charis, 39
Bowie, David, 100, 109
Bras, Myriam, 254
Briz, Antonio, 275
Brody, Jill, 247, 267
Broselow, Ellen, 306
Brunot, Louis, 39
Buchholtz, Mary, 103
Buchstaller, Isabelle, 251
Burdick, Christa, 221
Burridge, Kate, 99–101
Busch, Brigitta, 187
Busch, Brigitte, 336, 343

Caballero Fernández-Rufete, Salvador, 281
Cabanillas, Isabel De La Cruz, 274
Cameron, Deborah, 135
Campbell, Lyle, 336
Canagarajah, A. Suresh, iv, 6
Carlson, Lauri, 261
Carmeli, Yoram S., 174
Carter, Paul, 102, 104
Casado-Fresnillo, Celia, 217
Cassano, Paul Vincent, 277
Castillo Barril, Manuel, 226
Castillo Rodríguez, Susana, 233
Casu, Laura, 141
Caubet, Dominique, 40, 42, 47, 56
Chakrani, Brahim, 39, 41
Chambers, Jack K., 281
Chambers, Jack K., 102, 105, 116
Chaudron, Craig, 277
Cheikho, Louis, 309
Chevasco, David, 68
Chirilă, Elena Magdalena, 217
Cissé, Ibrahima A. H., 188, 189
Clyne, Michael, 247
Cody, Florian J., 315, 318
Collins, James, v
Comrie, Bernard, 310
Corriente, Federico, 303, 308
Coupland, Nikolas, v
Creese, Angela, vii, 335
Creissels, Denis, 206
Croft, William A., 196
Cuenca, Marie-Josep, 261

D’Arcy, Alexandra, 251
Dailey-O’Cain, Jennifer, 102
Danov, Milen, 9
Davidse, Kristin, 251
Davies, Eirlys E., 38, 40–42, 44, 56
Davies, Mark, 324
Davis, Mark, 153, 175
de Fina, Anna, 193
de Ruiter, Jan Jaap, 40
De Zarco, Mariano, 233
Dean, Kathryn, 154, 161
Degand, Liesbeth, 254
Del Valle, José, 220
Dewaele, Jean-Marc, 88
Diallo, Abdourahmane, 189, 192, 205
Diallo, Mohamed Larabi, 206, 207
Dietrich, Günter, 305
Dixon, R.M.W., 310
Dnevnik, 9
Döhla, Hans-Jörg, 302, 305, 306, 309, 310, 313–318, 320–323, 325
Dombrowsky-Hahn, Klaudia, 197

Name index

- Dostal, Walter, 131
Dostie, Gaétane, 246
Draper, Barb, 101
Drißner, Gerald, 306
Du Bois, John W., 103
Dumestre, Gérard, 188, 206
Dupoux, Emmanuel, 294
Durante, Charles, 66
- Eberhard, David M., 188, 189
Eckert, Penelope, 54, 100, 103, 153
Eco, Umberto, 325
Eelen, Gino, 159
Ehlich, Konrad, 249
Emirbayer, Mustafa, 340
Ennaji, Moha, 39, 40
Epp, Marlene, 99, 102, 106
Eychenne, Julien, 277
- Facebook, 162
Fagard, Benjamin, 254
Falchetta, Jacopo, 42, 43
Faulhaber, Charles B., 302
Feijóo Rodríguez, Sonia, 67, 72
Ferguson, Charles A., 40
Filho, Wilson Trajano, iii
Fischer, Wolfdietrich, 307, 308, 319
Fishman, Joshua A., 66
Fleiss, Joseph L., 282
Ford, David C., 323
Frantz, John B., 101
Fridland, Valerie, 113
Fruehwald, Josef, 102
Fussell, Elizabeth, 281
- Gadanidis, Timothy, 103
Gal, Susan, 161, 172
Galabov, Stoyan, 9
Gallois, Cindy, 219
- Galmés de Fuentes, Álvaro, 302, 303, 319, 320
Galtier, Gérard, 188
García, Ofelia, iv, 335
Gardner-Chloros, Penelope, iii
Gasparini, Fabio, 130, 131, 134
Gass, Susan M., 208
Ghanem, Asad, 161
Gold, Peter, 64
Gómez Capuz, Juan, 274–276, 293, 295
Gómez Vélez, Martha Isabel, 221, 231
Gordon, Matthew, 278
Goria, Eugenio, 66, 68, 72, 75, 89, 91, 92
Gorter, Durk, 4, 8
Goss, Emily L., 247, 267
Grandguillaume, Gilbert, 143
Grant, Anthony, 337
Gray, Jonathan, 277
Green, David, 336
Grocott, Chris, 64
Grosjean, François, 335
Grupo Val.Es.Co., 275
Güldemann, Tom, 190
Gumperz, John J., 49, 50, 67
Guthrie, Malcolm, 222
- Hachimi, Atiqa, 40
Hall, Kira, 103
Hamm, Bernd, 277
Handelman, Don, 152
Hannerz, Ulf, 140
Hartung, Cathrine, 154
Haspelmath, Martin, 274, 310
Haßler, Gerda, 302, 305, 325
Haugen, Einar, 274, 304
Hawker, Nancy, 151, 154, 157–160, 162, 163, 172, 173, 175

- Hazen, Kirk, 109
 Heine, Bernd, iii, 248, 249, 253, 254
 Heller, Monica, 153, 163, 174
 Henkin, Roni, 174
 Hennecke, Inga, 246, 249, 254, 266
 Hickey, Raymond, iii
 HighViewArt, 9
 Hill, Jane H., 247
 Hill, Kenneth C., 247
 Hilty, Gerold, 304
 Hitchcock, Richard, 303
 Hlavac, Jim, 246, 248, 268
 Hober, Nicole, 337
 Hoffmann, Roland, 312, 323
 Holes, Clive, 306–308
 Holes, Clive D., 140
 Holmes, H. B., 315
 Hottinger, Arnold, 305
 Hualde, José Ignacio, 286
 Huang, Jason L., 39, 41
 Huffines, Marion Lois, 100
 Huffman, Henry R. Jr., 302, 305

 Ibn al-Muqaffaʿ, 305, 314, 315, 317–319, 321, 322
 Iliev, Ivan, 10, 11
 INE, Instituto Nacional de Estadística, 278
 INEGI, Instituto Nacional de Estadística Geografía e Informática, 278
 Irvine, Judith T., 172

 Jake, Janice, 336
 Janzen, Jörg, 131
 Johnstone, Barbara, 100, 103
 Joos, Martin, 102
 Jørgensen, J. Normann, iv, 336
 Jucker, Andreas, 261, 262

 Juffermans, Kasper, iv
 Jørgensen, J. N., 5, 7

 Kaltenböck, Gunter, 246
 Kang, Yoonjung, 274
 Karsh, Efraim, 173
 Kasten, Lloyd A., 315, 318
 Katriel, Tamar, 173
 Kaufman, Terrence, iii, 210, 336, 343
 Kaye, Alan S., 139
 Kellermann, Anja, 64, 67, 68, 75
 Kendall, Tyler, 113
 Kerswilll, Paul, iii
 Khalidi, Rashid, 129, 144
 Khoumssi, Khawla, 41
 Kilgarriiff, Adam, 219, 279
 King, Lid, 335
 Klein, Wolfgang, 192, 209
 Knesset, 155, 160–162, 164, 165, 167, 169–171
 Knörr, Jacqueline, iii
 Koch, Gary G., 282, 283, 286
 Kopp, Achim, 100
 Krusteva, Asya, 8
 Kuipers, Giseline, 159
 Kuteva, Tania, iii, 248, 249, 253, 254

 Labatut, Roger, 198
 Labov, William, 43, 69, 102–104
 LaCharité, Darlene, 275, 294
 Lahlou, Moncef, 38, 40, 44
 Lakoff, George, 261
 Lander, Edgardo, 232
 Landis, J. Richard, 282, 283, 286
 Lankiewicz, Hadrian, 6
 Lantto, Hanna, 88
 Lara Delgado, Jonnhy, 221, 231
 Le Draoulec, Anne, 254
 Leal, Iris, 173

Name index

- Lehmann, Christian, 310, 312, 323
Leonti, Marco, 141
Levey, David, 64, 66, 68
Levey, Stephen, 265, 266
Levshina, Natalia, 111
Liddicoat, Anthony J., 340
Lieber, Rochelle, 306
Lipski, John M., 64, 216, 232, 233, 248, 255, 256
López de Ayala, Ignacio, 64
López García, Ángel, 313
Lorente, Beatriz P., vii
Lorenzo, Emilio, 275
Loureiro-Porto, Lucía, 64
Lozanova, Denitza, 8
Lüdecke, Daniel, 111
Lüpke, Friederike, 186

Maddy-Weitzman, Bruce, 143
Makoni, Sinfree, iv, 186
Maly, Ico, 4, 8
Manfredi, Stefano, 45, 46
Mariscal Ríos, Alicia, 65
Martín Rojo, Luisa, 216
Martínez Celdrán, Eugenio, 276
Martinez, Cristina Tejedor, 274
Maschler, Yael, 246
Matras, Yaron, iii, v, 185–187, 192, 246, 256, 334, 336–338, 340, 342–344
Max Planck Institute for Psycholinguistics, 104
Mbomio Bacheng, Joaquín, 216
Mcdossi, Oded, 174
McElhinny, Bonnie, vii, 153, 174
Mearns, David L., 133
Meechan, Marjory, 100
Meehan, Teresa, 250
Meisenburg, Trudel, 275

Meyer, John C., 159, 165
Mihatsch, Wiltrud, 250, 254
Miller, Catherine, 129, 130, 143, 144
Milroy, Lesley, 104, 278
Mische, Ann, 340
Moatassime, Ahmed, 39
Mohamadou, Aliou, 188
Molin, Paul Marie, 207
Montaner Frutos, Alberto, 323
Moore, Emilee, 335
Moore, Emma, 102, 104
Morales de Walter, Amparo, 226
Moravcsik, Edith, 336
Moreno Fernández, Francisco, 218, 220
Moreton, Elliott, 105
Morris, Miranda J., 130, 131, 143
Moscato García, Francisco, 234
Mosegaard Hansen, Maj-Britt, 245, 254, 255, 260, 261
Moseley, Christopher, 130
Mougeon, Raymond, 246–248, 255, 260, 267
Moustauoui Srhir, Adil, 231, 234
Moyer, Melissa G., 64, 65, 67, 68, 72, 75, 82, 89
Mualem, Mazal, 172
Mufwene, Salikoko, iii, 344
Mufwene, Salikoko S., iii, iv
Müller, Bodo, 318
Müller-Bardey, Thomas, 310
Muysken, Pieter, iii, 65, 71
Myers-Scotton, Carol, iii, 46, 67, 336
Myles, John, 153

Negev Coexistence Forum, 166
Neuhausen, Miriam, 109
Neumann-Holzschuh, Ingrid, 313
Niedrig, Heike, 188, 189

- Niedzielski, Nancy, 102
Nikolova, Milana, 10
Norman, Chelsey, 11
Nowicka, Magdalena, 154
Nycz, Jennifer, 100, 102, 103, 105, 107,
108, 113

OECD, 170, 174
Oncins-Martínez, José, 274, 275, 277,
295
Otsuji, Emi, 335
Owens, Jonathan, 143

Pappé, Ilan, 158
Paradis, Carole, 275, 294
Pardeshi, Prashant, 310
Pavlenko, Aneta, 88, 175
Payne, Thomas E., 310
Peirce, Charles Sanders, 54
Peled, Yoav, 161
Pellegrini, Chloé, 39
Pennycook, Alastair, iv, 186, 343
Pennycook, Alistair, 335
Peperkamp, Sharon, 294
Pérez, Maurilio, 303
Pérez, Raoul Monserrate, 315
Peterson, John E., 145
Pharies, David, 318, 319
Pichler, Heike, 265
Pietikäinen, Sari, iv
Piller, Ingrid, vii
Podesva, Robert J., 100, 103
Polinsky, Maria, 310
Politov, Alexander, 8
Poplack, Shana, 266
Post, Rebekah Elizabeth, 40, 41, 48,
57
Potowski, Kim, 246–248, 255, 260,
267, 268

Pratt, Chris, 274, 275
Prince, Ellen, 250
Pustka, Elissa, 275, 277, 278, 293–295
Python software foundation, 105

Quijano, Aníbal, 232
Quilis, Antonio, 217

R core team, 104
Raith, Joachim, 101
Rampton, Ben, iv, 3–5, 335
Real Academia Española, 276, 279,
292
Redouane, Rabia, 39
Rehbein, Jochen, 249
Repinecz, Martin, 216
Richter, Manfred, 100
Riochí Sifá, Juan, 216
Roberts, Julie, 102
Robertson, Alex, 334
Robinson, David, 112
Rodríguez García, Marta, 66
Rodríguez González, Félix, 274–277,
294, 295
Rodríguez M. Montalvo, Sagrario,
304
Romaine, Suzanne, 140
Rosa, Jonathan, 221
Rosen, Nicole, 100, 101
Rosenfelder, Ingrid, 102, 104
Rouhana, Nadim N., 152, 157
Rovisco, Maria, 154

Saad, George N., 323
Sadlier-Brown, Emily, 105
Sakel, Jeanette, 302
Salmons, Joseph C., 247, 267
Sanaphre Villanueva, Mónica, 302,
320, 321, 323–325

Name index

- Sánchez Moreano, Santiago, iv
Sandøy, Helge, 281
Sankoff, Gillian, 266
Saraceni, Mario, 5
Sauvant, Emile, 207
Sayahi, Lotfi, 39
Schieffelin, Bambi B., 233
Schiffrin, Deborah, 255, 261
Schlumpf, Sandra, 216, 217, 228, 229
Schweickard, Wolfgang, 274
Selinker, Larry, 208
Shalhoub-Kevorkian, Nadera, 152, 157
Shepherd, Nick, 221
Shibatani, Masayoshi, 310
Shields, Kenneth, 100
Shohamy, Elana, 155
Shohat, Ella, 157
Siegel, Jeff, iii
Signorell, Andri et mult. al., 282
Silva-Corvalán, Carmen, 248, 255
Silverstein, Michael, 54
Sim, Julius, 282
Simon-Vandenberg, Anne-Marie, 245, 261
Simonet, Francisco Javier, 303
Skriver, Crystal, 100, 101
Smandych, Russell, 277
Smith, Sara, 261, 262
Smolicz, Jerzy J., 145
Solomon, Julie, 247
Song, Jae Jung, 310
Springer, Otto, 100
Stanford, James N., 109
Stanley, Joseph A, 105
Stevenson, Patrick, 335
Stinson, Bobby Ray, 315
Stockey, Gareth, 64
Stolz, Christel, 336
Stolz, Thomas, 336, 337
Storch, Anne, 186
Stroud, Christopher, 334
Suárez-Gómez, Cristina, 64, 66
Suleiman, Yasir, 144, 157, 158
Swain, Merrill, 6
Tabouret-Keller, Andrée, 145
Taha, Habatalla, 175
Tallgren, O. J., 305
Talmy, Leonard, 196
Tamer, Youssef, 39
Taylor-Leech, Kerry, 340
Textor, Johannes, 284
Thomas, Bertram, 131
Thomas, Erik R., 102, 105
Thomason, Sarah Grey, iii, 210, 336, 343
Tilmatine, Mohand, 143, 235
Torres, Lourdes, 246–248, 255, 260, 267, 268
Tosco, Mauro, 143
Trubetzkoy, Nikolai S., 336
Trudgill, Peter, iii, 210, 281
Trujillo, José Ramón, 216
Underhill, Robert, 250
Väänänen, Veikko, 313
van Coetsem, Frans, iii
Van Den Avenne, Cécile, 188
Vandelanotte, Lieven, 251
Velupillai, Viveka, iii
Vendelin, Inga, 294
Veronelli, Gabriela A., 221
Vertovec, Steven, 335
Vigouroux, Cécile, iii
Vildomec, Veroboj, 208
Vydrin, Valentin, 206, 207

- Waltereit, Richard, 260, 261
Wąsikiewicz-Firlej, Emilia, 6
Waters, Sophia, 261
Watson, Janet C. E., 141
Weaver, Simon, 159
Weenink, David, 282
Wehr, Hans, 306, 311
Wei, Li, iv, 335
Weinreich, Uriel, iii, 186
Weston, Daniel A, 68, 75, 92
Whitney, W. D., 336
Wickham, Hadley, 288
Williams, Ann, iii
Williamson, Kay, 190
Winford, Donald, iii
Wolff, H. Ekkehard, 216
Woolard, Kathryn, 159
Woolard, Kathryn A., 233
Wright, Chris C., 282
Wright, Roger, 303
- Yadgar, Yaacov, 157
Yakpo, Kofi, 216, 232–235
- Zaborski, Andrzej, 143
Zezeza, Paul Tiyaambe, 231
Ziamari, Karima, 38, 40, 41, 44, 56
- ʿAzzām ʿAbdu , ʿAbdu ʿl-Wahhāb,
305, 315, 317

Global and local perspectives on language contact

This edited volume pays tribute to traditional and innovative language contact research, bringing together contributors with expertise on different languages examining general phenomena of language contact and specific linguistic features which arise in language contact scenarios. A particular focus lies on contact between languages of unbalanced political and symbolic power, language contact and group identity, and the linguistic and societal implications of language contact settings, especially considering contemporary global migration streams. Drawing on various methodological approaches, among others, corpus and contrastive linguistics, linguistic landscapes, sociolinguistic interviews, and ethnographic fieldwork, the contributions describe phenomena of language contact between and with Romance languages, Semitic languages, and English(es).