

# Heritage Languages

A language contact approach

Suzanne Aalberse, Ad Backus  
and Pieter Muysken

STUDIES IN BILINGUALISM

58

JOHN BENJAMINS  
PUBLISHING COMPANY

# Heritage Languages

# *Studies in Bilingualism (SiBil)*

ISSN 0928-1533

The focus of this series is on psycholinguistic and sociolinguistic aspects of bilingualism. This entails topics such as childhood bilingualism, psychological models of bilingual language users, language contact and bilingualism, maintenance and shift of minority languages, and socio-political aspects of bilingualism.

For an overview of all books published in this series, please see [benjamins.com/catalog/sibil](http://benjamins.com/catalog/sibil)

## **Editors**

Tihana Kraš  
University of Rijeka

Jason Rothman  
UiT The Arctic University and Universidad Nebrija

## **Advisory Editorial Board**

Fatih Bayram  
UiT the Arctic University of  
Norway

Sarah Bernolet  
Ghent University

Ellen Bialystok  
York University

Elma Blom  
Utrecht University

Kees de Bot  
University of Groningen

Marc Brysbaert  
Ghent University

Cécile De Cat  
University of Leeds

Annick De Houwer  
University of Erfurt

Cheryl Frenck-Mestre  
Aix-Marseille Université

Belma Haznedar  
Bogaziçi University

Erika Hoff  
Florida Atlantic University

Aafke Hulk  
University of Amsterdam

Judith F. Kroll  
University of California, Irvine

Tanja Kupisch  
University of Konstanz

Terje Lohndal  
Norwegian University of  
Science and Technology

Gigi Luk  
Harvard University

Viorica Marian  
Northwestern University

Loraine K. Obler  
CUNY

Johanne Paradis  
University of Alberta

Michael T. Putnam  
Pennsylvania State University

Ute Römer  
Georgia State University

Monika S. Schmid  
University of Essex

Ludovica Serratrice  
University of Reading

Darren Tanner  
University of Illinois Urbana  
Champaign

Enlli Thomas  
Bangor University

Ianthi Maria Tsimpli  
Cambridge University

Sharon Unsworth  
Radboud University Nijmegen

Marilyn Vihman  
University of York

Li Wei  
University College London

Marit Westergaard  
UiT, the Arctic University of Norway

Stefanie Wulff  
University of Florida

## **Volume 58**

Heritage Languages. A language contact approach  
by Suzanne Aalberse, Ad Backus and Pieter Muysken

# Heritage Languages

A language contact approach

Suzanne Aalberse

University of Amsterdam

Ad Backus

Tilburg University

Pieter Muysken

Radboud University Nijmegen

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/sibil.58

**Cataloging-in-Publication Data available from Library of Congress:  
LCCN 2019038038 (PRINT) / 2019038039 (E-BOOK)**

ISBN 978 90 272 0471 4 (HB)

ISBN 978 90 272 0470 7 (PB)

ISBN 978 90 272 6176 2 (E-BOOK)

© 2019 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Company · [www.benjamins.com](http://www.benjamins.com)

# Table of contents

List of tables	XI
List of figures	XIII
List of abbreviations and grammatical glosses	XV
Preface	XVII
CHAPTER 1	
<b>Heritage speakers and heritage languages</b>	<b>1</b>
1.1 Introduction	1
1.2 Characterizing heritage speakers	2
1.2.1 Unofficial language	3
1.2.2 Language dominance shift	4
1.2.3 Divergent grammars and other possible effects of the dominance shift	5
1.2.4 Personal and cultural ties to the language	7
1.2.5 Age of onset and acquisition in a naturalistic setting	8
1.2.6 Are HLs community languages?	9
1.2.7 Summary	10
1.3 The contact scenario approach to HLs	11
1.3.1 Typical contact scenarios	11
1.3.2 An example: Turkish as a HL in Northwestern Europe	13
1.3.3 Evaluating the scenario approach	15
1.4 Overview of the book	18
CHAPTER 2	
<b>History of the field of heritage language studies</b>	<b>23</b>
2.1 Introduction	23
2.2 The perspective of the diaspora languages	24
2.2.1 Dutch from a diaspora perspective	24
2.2.2 Other diaspora varieties	29
2.2.3 Diaspora studies in a broader perspective	30
2.3 The perspective of the country of immigration	30

- 2.3.1 The United States 30
- 2.3.2 Early studies on ethnolects and Canadian HL research 32
- 2.3.3 Case studies of HL languages in the United States 34
- 2.3.4 HLLs in Australia 39
- 2.3.5 The European context 40
- 2.3.6 Summary 41
- 2.4 Summary and introduction of the speakers' perspective 41

## CHAPTER 3

**Social aspects of heritage languages 43**

- 3.1 Introduction 43
- 3.2 The scenario approach: Attending to social and linguistic factors 43
- 3.3 Maintenance 45
  - 3.3.1 Indigenous minorities 46
  - 3.3.2 Immigration 47
  - 3.3.3 Social factors that affect maintenance 48
  - 3.3.4 Investigating language choice 51
  - 3.3.5 Ways of influencing language choice 57
- 3.4 Shift 58
  - 3.4.1 Shift and acculturation 58
  - 3.4.2 When shift reaches its endpoint 59
  - 3.4.3 Power versus solidarity 61
  - 3.4.4 Language shift and ethnolects 63
- 3.5 When language choice is not clear-cut 63
- 3.6 Summary 65

## CHAPTER 4

**Bilingual language use 67**

- 4.1 Introduction 67
- 4.2 Codeswitching and borrowing 68
  - 4.2.1 Early stages: Just foreign content words 69
  - 4.2.2 Intermediate stages: Diversified codeswitching patterns 72
  - 4.2.3 Shift stage: Development towards HL status in the narrow sense 78
- 4.3 Language change 81
- 4.4 Codeswitching in its social context 82
  - 4.4.1 Does codeswitching represent a third language? 82
  - 4.4.2 Linguaging 85
- 4.5 Summary 86

## CHAPTER 5

<b>Methods for collecting heritage language data</b>	<b>87</b>
5.1 Introduction	87
5.2 Criteria for evaluating a particular method: Validity	88
5.3 Overview of methods used	91
5.3.1 Spoken data	91
5.3.2 Written documents	97
5.3.3 Survey data and questionnaires	99
5.3.4 Experimental data	100
5.3.5 Judgment tasks	103
5.4 Summary and conclusion: Which method to choose?	107

## CHAPTER 6

<b>Studying variability in heritage language speaker populations and the base line</b>	<b>111</b>
6.1 Introduction	111
6.2 Establishing the baseline and the problem of monolingual bias	112
6.2.1 Standard language grammar	113
6.2.2 Exchange students and other recently arrived native speakers	113
6.2.3 Transnational research design	113
6.2.4 Vary subject populations	115
6.2.5 Cross-generational family studies	116
6.2.6 Multiple baselines	117
6.2.7 Bilingual baselines	118
6.2.8 Summary	118
6.3 Factors in individual variation in the acquisition perspective: Timing, quality and quantity of the input	119
6.4 Speaker characteristics, language use and language output	124
6.5 Social embedding in the multilingual speech community and the larger society	126
6.6 Identity work, style shift, variation, and change	131
6.7 Measuring proficiency and assessing linguistic profiles	132
6.7.1 Cloze test	133
6.7.2 Fluency measures	134
6.7.3 Lexical proficiency tasks	135
6.7.4 Sociolinguistic background questionnaires	137
6.8 Conclusion	139



## CHAPTER 7

**Heritage language phenomena and what triggers them**

141

- 7.1 Introduction 141
- 7.2 Phenomena studied 141
  - 7.2.1 Phonology 142
  - 7.2.2 Lexicon 143
  - 7.2.3 Morphology 145
  - 7.2.4 Syntax 145
- 7.3 Language internal factors: Changes in the input for new generations of speakers 146
  - 7.3.1 Order of acquisition 149
  - 7.3.2 Frequency 149
  - 7.3.3 Optionality 151
  - 7.3.4 Restricted use 154
- 7.4 Cross-linguistic influence: External factors 155
  - 7.4.1 Filter of grammatical categories via the dominant language 155
  - 7.4.2 Convergence through a shift in distribution 155
  - 7.4.3 Loan translations and semantic extensions 156
  - 7.4.4 Contact induced grammaticalization or additive borrowing 157
- 7.5 Comparing internal and external factors 158
- 7.6 Summary 160

## CHAPTER 8

**Grammatical models and research paradigms**

161

- 8.1 Introduction 161
- 8.2 Generative grammar 162
  - 8.2.1 Outline 162
  - 8.2.2 Case study 165
- 8.3 Variationist sociolinguistics 169
  - 8.3.1 Outline 169
  - 8.3.2 Case study 170
- 8.4 Optimality theory 172
  - 8.4.1 Outline 172
  - 8.4.2 Case study 173
- 8.5 Usage-based models 178
  - 8.5.1 Outline 178
  - 8.5.2 Case study 179
- 8.6 Summary discussion: Integrating the models 181

## CHAPTER 9

<b>Language processing in multilingual speakers</b>	<b>183</b>
9.1 Introduction	183
9.2 Core notions in research on language processing in bilingual speakers	184
9.3 Core findings	189
9.3.1 Cross-language interactions	190
9.3.2 Processing differences	192
9.3.3 Language switching and inhibition	194
9.3.4 Summary of preceding discussion	195
9.4 Factors influencing language processing in bilinguals	196
9.5 The issue of age of acquisition	199
9.6 Concluding remarks and perspectives for codeswitching research	200

## CHAPTER 10

<b>Heritage languages in a post-colonial setting: Focus on Papiamentu</b>	<b>203</b>
10.1 Introduction	203
10.2 Early history of Papiamentu	204
10.3 Background on Papiamentu and its status nowadays	205
10.4 A brief history of Papiamentu-Dutch contact	206
10.4.1 The 18th century	207
10.4.2 Increase of Dutch influence on Curaçao in the 19th and 20th centuries	207
10.4.3 The current situation	208
10.5 Dutch influence on Papiamentu	209
10.5.1 Quantity and quality of Dutch loans	210
10.5.2 Phonological adaptation of Dutch loans	211
10.5.3 Calques	211
10.5.4 Discourse markers and modal particles	212
10.5.5 Prepositions and verb particle combinations	213
10.5.6 Passive and the agent phrase	214
10.5.7 Other function words borrowed from Dutch	215
10.6 Morphological integration of Dutch nouns and verbs	215
10.6.1 Nouns and nominalizations	215
10.6.2 Verbs and inflection	217
10.7 Papiamentu in the Netherlands	218
10.8 Summary and conclusion	223

## CHAPTER 11

<b>The political dimension of heritage languages: Endangered languages, language rights, and the preservation of diversity</b>	<b>225</b>
11.1 Introduction: The politics of diversity management	225
11.2 Frames of reference	226
11.2.1 The Babylon frame	226
11.2.2 The Tsunami frame	227
11.2.3 The Heritage frame	227
11.3 Reversing language shift and indigenous language revival	229
11.4 HL education	232
11.4.1 Organization and support	232
11.4.2 Varieties of the HL taught	234
11.4.3 Dominant language from home country or home vernacular language?	234
11.4.4 HL proficiency as a learning resource within the mainstream classroom	236
11.5 Documentation of heritage varieties and language death	237
11.6 Codeswitching in HLLs and language loss	239
11.7 Linguistic human rights and HLLs	242
11.8 Conclusion and overview	243

## CHAPTER 12

<b>Technical terms used in this book related to heritage languages</b>	<b>247</b>
<b>References</b>	<b>263</b>
<b>Language index</b>	<b>297</b>
<b>Subject index</b>	<b>299</b>

# List of tables

- Table 1.1** Key dimensions of HLs cited by different authors 11
- Table 1.2** Well-known scenarios from the language contact literature, with potential links to the study of HLs 12
- Table 2.1** Features of different varieties related to the Dutch diaspora, including the Netherlands itself 28
- Table 2.2** Canadian Italian vocabulary in Toronto (based on Danesi, 1984) 33
- Table 2.3** Canadian Portuguese vocabulary in Winnipeg (based on Mota, 1997) 33
- Table 3.1** Generational shift in Turkish Dutch communities 53
- Table 5.1** Some earlier studies and the data collection methods used 95
- Table 6.1** The types of stability and change possible on the level of the individual and of the community (based on Labov, 1994) 132
- Table 6.2** American Russian: Continuum of speakers (based on Polinsky, 2006, p. 253) 136
- Table 8.1** The predictions of the umbrella hypothesis (adapted from Laleko & Polinsky (2016) 168
- Table 8.2** Utterances used progressively with less context 174
- Table 8.3** Utterances used progressively with more context 174
- Table 8.4** Bilingual grammar compared to monolingual grammar 177
- Table 8.5** Tableau for monolinguals in a less context situation 177
- Table 8.6** Continuum between different models 181
- Table 10.1** Census of 1981 and 2000/01 (adapted from Maurer, 1988, p. 143 and Kester & Fun, 2012, p. 238) 209
- Table 10.2** Papiamentu past participle formation strategies 218



# List of figures

- Figure 4.1** Stages in the status and use of a language in a minority community 69
- Figure 6.1** Connections between different subject populations 115
- Figure 6.2** Cross-generational family studies 116
- Figure 6.3** Multiple baselines that can be used 117
- Figure 7.1** The relation between the base line and the HL in the incomplete acquisition model 147
- Figure 8.1** Continuous ranking scale (adapted from Koontz-Garboden, 2004) 176
- Figure 8.2** Relative constraint rankings vary (adapted from Koontz-Garboden, 2004) 176
- Figure 9.1** A schematic presentation of the processing levels involved in language production and comprehension 186
- Figure 9.2** An example of sub-networks of languages within a general network of linguistic representations 189
- Figure 11.1** Parties engaged in the field of HL education 233



# List of abbreviations and grammatical glosses

ABL	ablative	INF	infinitive
ACC	accusative	INT	interjection
AOR	aorist	IP	inflection phrase
ASP	aspect	L1,2,3	first, second, third language
BLP	bilingual proficiency test	LOC	locative
EU	euphonic buffer element	NEG	negation
CAN	be able	OT	optimality theory
CL	classifier	OV	object verb
COND	conditional	PST	past tense
CP	complementizer phrase	PERF	perfective
D	Dutch	PL	plural
DAT	dative	PLUPF	pluperfect
DEM	demonstrative	POSS	possessive
DER	derivational affix	PRES	present tense
DET	determiner	PRO	pronoun
DIM	diminutive	PROG	progressive aspect
EMPH	emphatic marker	PRTC	past participle
EVI	evidential	REFL	reflexive
EXIS	existential	SES	socio-economic status
F	feminine	SG	singular
FOC	focus	(s)OV	(subject) object verb
FUT	future tense	(s)VO	(subject) verb object
GEN	genitive	T	Turkish
HAB	habitual	TMA	Tense Mood Aspect
HL	heritage language	TP	tense phrase
ILL	illative	V2	verb second
IMP	imperfective	VOT	Voice Onset Time
IND	indicative	WPM	word per minute
INDET	indeterminate		





# Preface

Heritage languages are non-dominant languages, sometimes with little prestige; their speakers also speak one or more dominant languages of the country they live in. Heritage languages have received a lot of scholarly attention and provide a link between the academic concerns of linguistics and those of the wider community. In this book, we will consider heritage languages from the perspective of their history, the social setting they appear in, their structural properties, and their interaction with other surrounding languages. The book grew out of the realization, as we were engaged in several projects on heritage languages, that there is wide-spread interest in this field, in several continents, and in various traditions but also that there was no systematic book length treatment available that outlines the main issues and problems from the perspective of language contact in this field in a systematic manner and in a global perspective. The contact perspective entails that we position heritage languages studies in a long tradition of studies in related language contact domains such as immigrant languages, contact induced language change, maintenance and shift, language death, bilingual language use, bilingual processing and human rights.

This volume differs theoretically from existing publications in the heritage languages field (e.g. Montrul, 2016 and Polinsky 2018a) by taking a very broad language contact perspective. This means that we will focus not just on reduction, simplification and loss, but also on addition of or changes in linguistic features, due to contacts with other languages spoken in the heritage language context. Thus the heritage languages in a particular country can also be seen as constituting a linguistic area due to earlier processes of convergence. Furthermore, heritage languages are viewed as sociolinguistic phenomena, involving agency and creation in addition to acquisition constraints and processing costs. We emphasize language use in addition to language acquisition. This means that if heritage speakers diverge in their speech from homeland speakers, we interpret these differences as possibly resulting from identity work and natural change and adaptation to a new environment, rather than as proficiency issues only. It also means that there is attention for social meaning conveyed through language choice, bilingual games and codeswitching. Also, data from recorded language corpora are taken into account, in addition to data from tests and controlled experiments. A final

distinctive feature of this book is that it takes a global perspective, rather than only focusing on North America and Europe, although studies from those regions will be amply cited.

Our own expertise is complementary. Suzanne Aalberse had studied the connection between language acquisition and language change partly inspired by the generative framework. Ad Backus works on usage-based models, codeswitching, and contact effects in the Turkish spoken in immigrant communities in Europe. Pieter Muysken's main expertise lies in creole studies and contact-induced language change.

We want to acknowledge the research groups at our respective universities, Amsterdam, Tilburg and Nijmegen for their support. Amsterdam has a lively group working on bilingual acquisition and heritage languages, working primarily from a generative perspective. Tilburg has a long tradition in this area, with a number of studies particularly on Moroccan Arabic and Turkish as heritage languages. The Radboud University Nijmegen Languages in Contact group has added to this a series of studies on yet another set of heritage languages.

We also want to acknowledge the co-authorship of several colleagues for specific chapters: Anne Verschik for the chapter on bilingual language use, Gerrit Jan Kootstra for the chapter on Processing, and Bart Jacobs for the chapter on Papiamentu. For the rest, each of us had the primary responsibility for several chapters, as listed below, but all of us contributed bits here and there.

At our different universities we have had stimulating discussions about the issues discussed here with Sible Andringa, Aafke Hulk, Jan Hulstijn, Folkert Kuiken, Caitlin Meyer, Brechje van Osch, Judith Rispens, Jeannette Schaeffer, Petra Sleeman, Josje Verhagen and Fred Weerman in Amsterdam; Zeynep Azar, Rik Boeschoten, Derya Demircay, Seza Doğruoz, Nadia Eversteijn, Guus Extra, and Pelin Onar Valk in Tilburg; Mariam Hadji, R. van Hout, Pablo Irizarri van Suchtelen, Gerrit Jan Kootstra, Francesca Moro, Hülya Sahin, Sophie Villerius, and Kofi Yakpo in Nijmegen.

We are grateful to other colleagues in the US, Canada, Europe, and Australia who have been working in this area for a much longer period, including the late Michael Clyne, Naomi Nagy, Carol Pfaff, Carol Myers-Scotton, Maria Polinsky, Silvana Montrul, and Anne Verschik. Two anonymous reviewers and the Series Editor for this volume, Jason Rothman, have helped enormously to improve the quality of this book. Needless to say, none of these friends and colleagues is in any way responsible for the remaining shortcomings.

We also want to acknowledge the support of our institutions and several granting agencies for research reported on here.

- Amsterdam Centre for Language and Communication (ACLC)
- Centre for Language Studies (CLS), Radboud University Nijmegen
- European Research Council with the ERC Advanced Grant Traces of Contact to Pieter Muysken
- Netherlands organization for scientific research (NWO) with a Postdoc grant to Ad Backus for the project ‘Synchronic register variation in monolingual and bilingual Turkish’, 2003–2007
- The Tilburg School of Humanities, Tilburg University
- Stellenbosch Institute of Advanced Studies (STIAS) for the opportunity for Pieter Muysken to work on editing on this volume

The primary authorship for the individual chapters was as follows:

1	Heritage speakers and heritage languages	SA & PM
2	History of the field of heritage language studies	PM
3	Social aspects of heritage languages	AB
4	Bilingual language use	AB & A. Verschik
5	Methods for collecting heritage language data	AB & PM
6	Studying variability in heritage language speaker populations and the base line	SA
7	Heritage language phenomena and what triggers them	SA
8	Grammatical models and research paradigms	SA & PM
9	Language processing in multilingual speakers	G. J. Kootstra & PM
10	Heritage languages in a post-colonial setting: Focus on Papiamentu	B. Jacobs & PM
11	The political dimension of heritage languages: Endangered languages, language rights, and the preservation of diversity	PM & AB
12	Technical terms	PM

Finally, a note on terminology. We have used the abbreviation ‘HL’ to refer to ‘heritage language’, and we use the general term ‘codeswitching’ where we also find code-switching and code-mixing in the literature.



# Heritage speakers and heritage languages

## 1.1 Introduction

You inherited language from your parents, but the fact that you did does not directly turn you into a heritage speaker. Common to most definitions of heritage speakers is that they learned a language at home that is not the dominant language of the country. The language that is not the dominant language of the country but that does connect you to your roots is the heritage language (HL). The term heritage refers to knowledge and goods from the past that can be used now and in the future. Because the term heritage is frequently used in the context of endangered heritage, for example by UNESCO, some people associate the term ‘heritage’ with vulnerability and threat, rather than its more neutral meaning of something associated with the past.

The observation that only languages that are in competition with a dominant language are referred to as HLs strengthens the association of threat, but this association is not felt by all. Some HLs like heritage Mandarin Chinese have a vibrant and long future ahead from the global perspective. The language may not be maintained by all speakers however. The extent to which HLs are under threat varies across languages and speakers. The possibility of language loss is considered in Chapter 3 as one of the options of intergenerational transmission of HLs.

Different researchers have defined heritage speakers in slightly different ways, depending on their theoretical background. This chapter shows what features belong to the core of the definition of heritage speakers within a certain tradition and explains this core as related to the main goal of these linguistic traditions.

Whereas a sociolinguist may be interested in patterns of maintenance and shift (e.g. “do Italian Americans still speak Italian or do they shift to English?”), a generative linguist may focus on age of acquisition effects: HLs are learned in a naturalistic setting at an early age; if age is the most important factor in acquiring grammatical categories like gender, why do heritage speakers sometimes diverge from like monolingual natives? In a usage-based approach, the focus will be on actual use of the language in daily life. The focus of these research interests feeds the way a heritage speaker is defined.

This chapter illustrates distinctions and similarities between approaches to the definition on the basis of speaker bibliographies. For example, Xiaonan is a second-generation Chinese, her parents spoke Dutch to her at home, but she was sent to Chinese school when she was six years old and she speaks Chinese with her grandmother and she speaks Chinese in the local community center. Is Xiaonan a heritage speaker? If you think the core of being a heritage speaker is the connection of the language to your ancestral ties the answer would be yes. If learning the language in a naturalistic setting at a young age is the core of your definition, the answer is less clear.

Throughout this chapter you are asked to reflect on this type of question and to actively use the criteria presented such as age of onset and ancestral affiliation to identify heritage speakers. After first comparing definitions of HL speakers, the second section situates HLs in the more general field of language contact studies. The third section provides an overview of the remainder of the book.

**Main goals of this chapter:**

To understand the different ways in which heritage speakers are defined.

To situate HLs in the more general field of language contact studies, and discuss them in terms of the contact scenario approach.

To present an overview of the book.

## 1.2 Characterizing heritage speakers

Who are heritage speakers? The term heritage speakers was developed in Canada in the seventies in the context of the Ontario Heritage Languages program, a government program that provided funding for school systems for 2.5 hours of HL instruction. The term began to be used by American language policy makers in the nineties (Cummins, 2005, p. 585) and started to gain currency with theoretically oriented linguists in the beginning of this century (Van Deusen Scholl, 2003, p. 212). The term is now gaining currency in Europe as well. (cf. Benmamoun, Montrul, & Polinsky, 2013a, b; Kupisch, 2013). The Canadian government defined HLs as follows (Nagy, 2011): “a mother tongue that is neither an official language, nor an indigenous [i.e. Aboriginal] language” (Harrison, 2000; Cummins, 2005).

We will take this first definition as a starting point to discuss the six central questions that play a role in the characterization of HLs and the study of heritage speakers, namely:

1. Does the language have official status in the country where it is spoken?
2. Did the speaker undergo a shift in language dominance?
3. What is the language proficiency of the speaker (upper and lower bounds)?
4. Does the speaker have personal and ethnic or ancestral ties to the language?
5. What was the age of acquisition of the heritage language and was it learned in a naturalistic setting?
6. Is the language a community language?

These questions will be discussed one by one in the following sections. When reading literature on heritage languages it is important to keep in mind that authors differ in who they define as heritage speakers and that answers to these six questions can therefore differ. We include work based on many different definitions of heritage speakers.

Although the term ‘heritage’ requires has many components that might be viewed differently by different authors, the term always includes the passing on of patterns, beliefs, and ways of living from one generation to the other. Heritage can be tangible (such as buildings, objects, and landscapes) and intangible (such as cultural practices or language). Heritage is something of the past, since it has been passed on from before, but also of the present, since it has been accepted, even if it often undergoes changes. Heritages are valued, as we will also show in Chapter 11, as part of the treasure of a community.

### 1.2.1 Unofficial language

One aspect in the Canadian definition of HLs is the unofficial status of the language in the country where it is spoken. The Canadian government definition is very clear in this respect: only languages that are not official or indigenous are HLs. In this respect the term HL overlaps, but by no means coincides, with the term ‘minority language’ (for discussion see Loudon, 2016). A consequence of the unofficial status of the language is that exposure to the language is usually limited to more informal settings such as the home or community center. Similar situations of limited use play a role in other languages that are not the dominant language of the country such as indigenous languages (e.g. languages that were spoken by the indigenous people, but that are not the dominant language of the country (anymore) such as native American languages in the Americas or Gaelic in Ireland) and colonial languages (languages brought along by colonists, but not the official language of the country such as German or Norwegian in the US). Note that it is possible that some languages have an official status such as Frisian in the Netherlands as a second official language next to Dutch (speakers have the right to communicate with government representatives in Frisian for example) without



being the dominant language of the country. Fishman (2001a, p. 81) refers to all these languages as HLs because speakers are connected to these languages through their ancestry and because they are in competition with the dominant language of the country. We follow this broader interpretation of heritage languages by Fishman (2001a) in this book and include indigenous languages and official minority languages in the discussion of heritage languages, because they are comparable in terms of language contact effects. Note however, that from the Canadian government perspective these languages are not heritage languages.

Let us now delve deeper into the six components of being a heritage speaker by comparing speaker biographies. Consider the following two speaker bibliographies – taking in mind that HLs tend to be the languages that are not spoken in official settings. Which speaker is a heritage speaker and why?

#### *Maria*

Maria was born and raised in Miami, Florida. Her parents were born in Cuba and they read newspapers in Spanish and listen to the radio in Spanish. She uses Spanish on Facebook to connect with her Cuban cousins. She speaks Spanish at home and she can read Spanish. English is her dominant language outside the home, and she considers herself Cuban-American (see Carreira, 2004).

#### *Danique*

Danique has French parents and they speak French to each other and to Danique. They live in the Netherlands and the dominant language outside is Dutch. Danique goes to a French track in the international school and it is her ambition to go to a French university when she is older. She identifies herself as French and as a “world citizen”.

Both Maria and Danique speak a language at home that is not an official language of the country. This aspect makes both speakers “heritage speakers”. But there is a crucial difference between them? What is it?

### 1.2.2 Language dominance shift

A crucial difference between the two speakers is language dominance: Danique is dominant in her L1 French and she perceives herself as French. In contrast, Maria has shifted dominance: whereas she began her life speaking Spanish, she now feels more comfortable speaking English and she identifies as Cuban-American rather than only Cuban or only American. Many definitions of heritage speakers consider shifted language dominance as a crucial point for heritage speaker status. A frequently quoted definition of heritage speakers that includes dominance shift is Valdés (2000), though its usefulness is somewhat limited as it is formulated to

apply to Spanish in the United States. Below we included a more general version of this definition by Benmamoun, Montrul, & Polinsky (2013a).

*Roughly, we define heritage speakers as asymmetrical bilinguals who learned language X – the ‘heritage language’ – as an L1 in childhood, but who, as adults, are dominant in a different language*

Although language dominance shift is often assumed as characteristic for heritage speakers, especially in the American literature on immigrant HLs, not all researchers assume it (see Nagy, 2015; Kupisch & Rothman, 2018). To some researchers any language that is not the dominant language of the country counts as an HL, even if it is the dominant language for a certain individual. However, most researchers do assume a dominance shift, and a feature that is directly related to dominance shift, is proficiency. A shift in dominance is often associated with a decrease in proficiency in the HL.

### 1.2.3 Divergent grammars and other possible effects of the dominance shift

The dominance shift described in 1.2.2. can affect language development. One effect of less use is decreasing confidence in speaking the heritage language and less confidence can in turn lead to even less use (Sevinç & Backus, 2017). Lack of use might also result in (temporary) disfluencies such as a lower word per minute rate or a higher uh-rate (Irizarri van Suchtelen, 2016). Grammatical structures might also be affected by less or halted exposure. According to some definitions only those speakers who feel insecure about their heritage language, who sound disfluent and/or who diverge noticeably from monolingual natives are considered heritage speakers. Heritage speakers are sometimes not just perceived as different but also as less proficient speakers. Nagy (2015) refers to this perspective as the ‘deficiency perspective’. We discuss problems with this perspective in Section 7.3.

The idea that heritage speakers are speakers whose language has changed under influence of the dominance shift is expressed by, again, Valdés (2000), and also Polinsky (2011), who states that:

*Heritage languages are spoken by early bilinguals [...] whose L1 (home language) is severely restricted because of insufficient input. [...] they can understand the home language and may speak it to some degree but feel more at ease in the dominant language of their society.*

Polinsky’s (2011) definition implies that the heritage language is the weaker language in the sense that speakers feel less at ease in speaking the heritage language than in the dominant language of the country. The idea that a heritage language is a weaker language is also expressed in her 2018 book which describes the heritage

speakers as ‘unbalanced bilinguals whose heritage (weaker) language is their first language’ (Polinsky 2018a, p. 4). Polinsky (2018a, p. 9) notes that early definitions of heritage speakers such as Valdes (2000) stress differences between heritage native speakers and monolingual native speakers. More recent research includes an interest in what is stable across heritage and homeland speakers (Polinsky 2018a, p. 9).

Note that Polinsky’s (2011) definition implies that heritage users at least understand the language well, but that they do not necessarily have to be able speak it. Users who can understand a language but cannot or dare not speak it are also referred to as ‘overhearers’ (Au, Knightly, Jun, & Oh, 2002; Knightly, Jun, Oh, & Au, 2003). In practice, part of the way heritage users are generally conceived of is some communicative competence in order to function inside their speech community. For example, if heritage users can speak this makes it easier to gather data from them and some universities require a minimum amount of linguistic knowledge in order for students to register for an HL class rather than for the general second language learning class. However, central in many definitions are not the lower proficiency levels, but the higher ones: heritage speakers lack something when compared to monolingual native speakers. Sometimes the notion of ‘receptive bilingualism’ is used (Sherkina-Lieber, 2015) to characterize speakers who cannot actively speak the HL but who do comprehend it.

A tradition that uses lesser proficiency as a defining part of speakers is the study of semi-speakers. The label “semi-speaker” was used by Dorian (1981) and it partially overlaps with “heritage speaker”. Semi-speakers speak a language that is not the dominant language of the country; it may be an indigenous or aboriginal language or an immigrant language, as in the case of Cape Breton Gaelic, and Pennsylvania German. Indigenous languages are explicitly included in Fishman’s (2001a) definition of HLs. Dorian studied language loss of Gaelic in isolated language communities in Scotland. Within these communities, levels of linguistic proficiency in Gaelic were unequally distributed. While some speakers were fluent, most were dominant in English, the language with more prestige. Speakers whose proficiency in the indigenous language was low were referred to as semi-speakers. Whereas the notion semi-speakers has less proficiency as a core part of the definition, this is certainly not true for all traditions. Note that some authors explicitly include all speakers with a home language other than the dominant language as heritage speakers including those cases where no divergence is observed between these speakers and homeland speakers (cf. Nagy, 2015; Kupisch & Rothman, 2018).

If we look at the speaker biographies above from the perspective of divergent proficiency, Danique in case study 2 would not be a “real” heritage speaker because her French is very monolingual-like and not divergent enough to be considered a heritage speaker.

Apart from proficiency, other factors play a role as well in being considered a heritage speaker. The idea that you can be a “real” or not so “real” heritage speaker also touches on your identity. Identity issues related to the label “heritage speaker” play a prominent role in language pedagogy. Compare the following three speaker profiles and think about what it means to these three speakers to be given the label “native speaker”, “heritage speaker” or “second or foreign language learner”.

#### *Rose*

Rose grew up in an Ambon Malay community in the Netherlands. Her parents had already learned Dutch in Ambon (see Chapter 10) and they feel Dutch is the language that will enhance career chances; therefore they speak Dutch at home. Later in life, Rose took courses in Ambon Malay and as soon as she could speak well enough she began to speak the language with her family and within her community.

#### *Safiya*

Safiya was born and raised in the United States just like her parents, grandparents and great-grandparents. She grew up speaking mostly English at home and at school. She overheard Spanish from other children and knows a few phrases in that language. She refers to herself as an African American. When she is in college she enrolls in a class for Swahili to connect with her African roots.

#### *Ernest*

Ernest has a Dutch father and an Indian mother. He identifies himself as 100% Dutch with an Indian mother. He feels comfortable in India, he likes the food, the liveliness, but he does not speak the language. He explains that his mother refused to speak Hindi to him except when she was very angry. He says that he knows seven words that mean ‘illegitimate child’ in Hindi because of the outbursts of his mom.

The label “native speaker” would be problematic for all three speakers, because they did not speak the language of their ancestors since birth. Some would refer to Rose and Safiya as heritage speakers or as learners with a heritage motivation, because of their personal attachment to the language (Van Deusen Scholl, 1998, 2003; Fishman, 2001a; Carreira, 2004). Because Ernest does not identify as Indian and because he hardly speaks the language, he would be less likely to qualify as a heritage speaker. These observations bring us to the fourth point of the definition: “personal and cultural connection” to the language.

### 1.2.4 Personal and cultural ties to the language

The motivation to learn a new language (Ambon Malay and Swahili respectively) in Rose and Safiya is related to their ethnic and cultural heritage. Either directly as in the case of Rose who is learning the language of her family and more indirectly

to Safiya (because the language she learns is associated with her ethnicity but not necessarily the language any of her direct ancestors ever spoke). Carreira (2004) shows that learners who feel a personal tie to a language have special needs as language learners. Referring to these learners as second or foreign language learners may make the learners feel like outsiders of the group they feel connected to. Authors like Carreira (2004) and Polinsky and Kagan (2007) refer to speakers who have limited knowledge of the language but who have a personal or cultural connection to the language as “heritage learners broad”. Since Ernest is not speaking or learning Hindi and because he does not identify himself as Indian, there is less reason to label him as a heritage speaker. A typical heritage speaker feels a personal and emotional connection to the HL and identifies with the language. The strong connection between identity and the HL is expressed well by Josh in the example below.

*Josh*

“My home language is Chinese. My parents are from China. They praised me, scolded me, all in Chinese. ... My Chinese is really bad. I can’t read and I can only write my name. But when I think of Chinese, I think of my mom, dad, and home. It is the language of my home, and my heart.” He (2010)

Although Rose and Safiya have a personal connection to the language they are learning just like Josh, not all definitions of heritage speakers would include them as heritage speakers, because of the age of onset of exposure to the HL.

### 1.2.5 Age of onset and acquisition in a naturalistic setting

To some approaches, the age of onset and the setting in which the languages are learned are crucial in distinguishing heritage speakers from second language speakers (but note that heritage speakers in the broad sense, that is speakers who have personal attachment to the language, have special identity needs when learning the language). Consider the statement formulated by the original steering committee of the National Heritage Language Resource Center at UCLA in 2000:

*“A defining distinction between heritage language and foreign language acquisition is that heritage language acquisition begins in the home, as opposed to foreign language acquisition which, at least initially, usually begins in a classroom setting”*

UCLA Steering Committee (2001), *The Family* (taken from Lynch, 2014)

One reason for the interest in heritage speakers is that they can inform us about age of onset issues (see Chapter 7) and the role of language use in language proficiency; HLs are usually learnt at an early age in a naturalistic setting just like monolingual L1 acquisition, but the acquisition process differs between speakers

who keep using the language in most domains the rest of their lives and those who shift to another language. Comparing heritage speakers to adult second language learners can inform researchers about the role of language use (input and output) on the one hand, and the effect of early age effects and acquisition processes in a naturalistic setting on the other hand. The definition of heritage speakers that most strongly focuses on age of onset issues rather than ancestral connections comes from Rothman (2009, p. 156) who defines heritage speakers as follows:

*“A language qualifies as a heritage language if it is a language spoken at home or otherwise readily available to young children, and crucially this language is not a dominant language of the larger (national) society. Like the acquisition of a primary language in monolingual situations and the acquisition of two or more languages in situations of societal bilingualism/multilingualism, the heritage language is acquired on the basis of an interaction with naturalistic input and whatever in-born linguistic mechanisms are at play in any instance of child language acquisition. Differently, however, there is the possibility that quantitative and qualitative differences in heritage language input and the introduction, influence of the societal majority language, and differences in literacy and formal education can result in what on the surface seems to be arrested development of the heritage language or attrition in adult bilingual knowledge.”*

To Rothman (2009) the crucial fact about HLs is the way they are acquired: at a young age in a naturalistic setting without being the dominant language of the country.

### 1.2.6 Are HLs community languages?

Rose speaks the HL in her own community e.g. with her family and people in her neighborhood. Some research on HL supports the idea that heritage speakers form a language community, with their own linguistic norms. The perspective of a coherent community of speakers is especially important for sociolinguists. For example, Nagy (2015) writes: “The grammar of each language variety (e.g., heritage vs. homeland, Generation 1 vs. Generation 2) is first examined as a complete variable system that stands on its own. Comparisons between systems (e.g., between generations or between heritage and homeland varieties) are then made using the same methods for each group of speakers.”

Two examples of research on heritage speakers as a community of speakers are the studies performed by Doğruöz & Backus (2009) on NL-Turkish (Turkish in the Netherlands) and by Tahitu (1989) on Melayu Sini (“Malay from here” referring to Malay in the Netherlands). If heritage speakers form a tight-knit community it is very well possible that a special new heritage variant arises with its

own norms. Note, however, that not every HL develops in a community setting. Some HLs are spoken at home only, or the social networks are too varied to create a more fixed set of community norms. For example, Lynch (2013), cited in Lynch (2014, p. 231) states with regards to US Spanish that there is no ‘U.S. Spanish’ variety to speak of in the traditional sociolinguistic sense, for two principal reasons: (1) a lack of community-based norms of usage owed fundamentally to the lack of generational continuity of Spanish in the U.S.; and (2) readily apparent patterns of discontinuity according to the variationist model, that is: lack of a shared “set of norms for the interpretation of language, as reflected in the treatment of linguistic variables: patterns of social stratification, style shifting, and subjective evaluations” (Labov, 1989).

When HLs are spoken in a tight-knit community, it opens up the possibility of a new variant emerging with added complexities such as new verbal inflections in heritage German as spoken in Northern Italy (Dal Negro, 2004) or finiteness and definiteness marking in Ambon Malay in the Netherlands (Moro, 2016).

### 1.2.7 Summary

In short we have discussed six dimensions cited as characterizing HLs and heritage speakers, namely, the unofficial status of the language (1), a dominance shift (2), a divergent grammar (3), personal, ancestral and cultural connections to the language (4), early age of onset of acquisition of the language in a naturalistic setting (5) and the limited use of the language within a language community (6). The weight of each criterion differs per definition, up to the point that a criterion may be irrelevant to some researchers, depending on whether their approach is primarily sociological, sociolinguistic, or psycholinguistic. When you read an article it is important to keep in mind who counts as heritage speakers/users for the author, because different authors may use different definitions. In Table 1.1 we summarize the previous discussion.

For those researchers that assume that dominance shift is a key part of the characterization of heritage speakers, one becomes a heritage speaker only after dominance shift, so after the age of entering school around the age of four to six (cf. Aalberse & Hulk, 2018). Much research on heritage learners focuses on (young) adults rather than children. Putnam, Kupisch, and Pascual y Cabo (2018) stress that ‘HL’ is actually an umbrella term that covers language use in various age and speaker groups. While we agree with many of their arguments to include bilingual child acquisition (sometimes labeled 2L1) in HL studies, this is not our primary focus here, since most of the language contact literature deals with older speakers.

It is useful at this point to mention ethnolects, broadly speaking language varieties associated with specific ethnic subgroups within a larger society.

**Table 1.1** Key dimensions of HLs cited by different authors

	Fishman (2001a)	Valdés (2000)	Polinsky (2011) Polinsky & Kagan (2007)	Carreira (2004)	Rothman (2009)	Nagy (2015)
no official status	x	(x)	(x)	(x)	x	(x)
a shift in language dominance		x	x			
divergent grammar		x	x			
personal and ethnic or ancestral ties	x	(x)	(x)	X		(x)
age of acquisition		x	x		x	
community language						x

Typically, ethnolects are seen as variants of the dominant language, e.g. Italian American English, but sometimes a broader definition is used, touching upon HLs. This is further discussed in Chapter 2 in relation to the work of the Canadian researcher Danesi.

### 1.3 The contact scenario approach to HLs

In this section we will situate HLs in the more general field of language contact studies, and discuss them in terms of the contact scenario approach. The scenario approach received its major impetus with the distinction between ‘maintenance’ and ‘shift’ scenarios in Thomason & Kaufmann (1988) and ‘source’ versus ‘recipient language agentivity’ in Van Coetsem (1988, 2000). It has since then been further developed and elaborated by various others (cf. e.g. Muysken, 2010a). The key idea in this approach is that there are **specific socially determined language contact settings that have specific linguistic outcomes**. We will call these settings “scenarios”.

#### 1.3.1 Typical contact scenarios

Typical scenarios include linguistic borrowing into an HL from a dominant language or shift to that language, but there are many other scenarios as well. These are outlined in Table 1.1., which presents some well-known scenarios from the language contact literature, arranged in terms of how frequent they are in different communities, and taking into account whether they assume a power asymmetry between the languages involved or not and what their grammatical properties are.



**Table 1.2** Well-known scenarios from the language contact literature, with potential links to the study of HLs

Scenarios for HLs	Frequency	Symmetry configuration	HL properties affected in this scenario
<b>Linguistic borrowing:</b> Speakers of an HL adopt words and word related patterns from a dominant language	Frequent	Asymmetrical: from a dominant superstrate to a socially subordinate language	Words, sometimes word endings, in short relatively concrete features, rather than patterns
<b>L2 learning, shift and substrate formation:</b> Speakers of an HL gradually stop speaking this language, and shift to a new variety of the dominant language, which may contain patterns of their original language	Frequent	Asymmetrical: from a subordinate language to a socially dominant language	Relatively abstract language features from the L1, including semantic and pragmatic ones, as well as morphological, phonetic and morpho-phonemic distinctions from the L2
<b>Codeswitching:</b> Bilinguals mixing their HL and other languages more or less on an equal level	Frequent	Often asymmetrical, but some symmetrical patterns have been reported as well	Complete fragments from both languages interact
<b>Attrition:</b> Speakers of an HL continue speaking their language but through limited use it loses many original features	Relatively frequent	Asymmetrical: a subordinate language is used less, a new language is dominant	Particularly lexicon, more complex inflection and derivation, and complex grammar from the HL are affected
<b>Leveling:</b> Speakers of originally different dialect varieties of the HL arrive at a (sometimes slightly simpler) compromise variety	Relatively well-attested	No assumption of asymmetry, although some varieties involved in the leveling process may be more important than others	Mostly morphological, lexical, and phonological features
<b>Grammatical convergence under language maintenance:</b> As part of prolonged bilingualism patterns in the HL may start resembling patterns of the dominant language	Relatively frequent, though not as frequent as borrowing	Potentially symmetrical	May lead to surface convergence, e.g. in semantic categories, word order, and intonation

Table 1.2 (continued)

Scenarios for HLs	Frequency	Symmetry configuration	HL properties affected in this scenario
<b>Relexification:</b> HL speakers maintain their language, so to speak, but massively replace the word forms of this HL with forms from a dominant language	Highly infrequent in its stable forms	Asymmetrical in the division of labor, but both languages play an important role	Typically, the grammar of one language is combined with the lexicon of another one
<b>Metatypy:</b> This would involve an extreme case of convergence. Affected are basic grammatical and constituent order patterns	Relatively infrequent	Asymmetrical: a subordinate language changes its typological properties under the influence of a dominant language	the HL gradually changes its grammatical character under strong influence of the dominant language

The logic in this approach to linking contact settings with their outcomes is roughly the following: Assume that if a prototypical social setting involving language contact *A* has been well studied and produces linguistic properties *p* and *q*, then a social setting under study, *B*, resembling *A* in crucial ways, will be likely to also have these properties *p* and *q*, assuming also roughly the same types of languages involved. For example, immigrant group X in the US (with e.g. Finnish as HL) will resemble another immigrant group Y (with e.g. Russian as HL), and thus is likely to show the same linguistic phenomena as Y in the contact setting, for instance, the same lexical borrowing patterns, and the same general attrition patterns in case marking.

Following this logic, language contact scenario models could be used in two ways:

- i. They could predict, given a specific language contact setting and a specific language pair, what the linguistic outcome is most likely to be.
- ii. They could help understand, given a specific linguistic outcome, what would be the most likely contact setting leading to that outcome has been.

### 1.3.2 An example: Turkish as a HL in Northwestern Europe

To give a more elaborate contemporary example, consider codeswitching between Turkish as an HL in Northwestern Europe and the Germanic languages spoken there. We have some examples such as the following from Turkish-Dutch and Turkish-German bilingual speech, typically referred to as codeswitching.

- (1) *iki gün önce işte bioscoop-a, vragen yap-tıydı-m*  
 two day before INT cinema-DAT ask do-PLUPF-1SG  
 ‘but two days before I had asked her along to the movies’ (Backus, 1998, p. 9)
- (2) *ben feiern yap-mi-yca-m ki ama feiern etm-iyor-um ki*  
 I celebrate do-NEG-FUT-1SG EMPH but celebrate do-PROG-1SG EMPH  
 ‘I will not celebrate my birthday, but I won’t celebrate.’  
 (Kallmeyer & Keim, 2003, p. 36)

The examples share the use of infinitives embedded into the Turkish structure, often with the inflected verb *yapmak* ‘do’ (sometimes *etmek* is used, as (2) shows). We also typically find non-Turkish nouns with Turkish case endings.

Given such examples, we may be able to predict that the following Turkish-Norwegian code-switched utterance is also a possible outcome, as indeed it is (cited from Türker, 2000, p. 113; see also Backus & Türker-Van der Heijden, 2009):

- (3) *bu-n-lar-ı nasıl avslutt-e yap-ıyor-sun bu-r-da?*  
 this-EU-PL-ACC how finish-INF do-PROG-2SG this-DER-LOC  
 ‘How do you close these down here?’ (Erdal, intermediate generation)

The authors argue that here Standard Turkish would have *bitirmek* ‘finish’ or *kapatmak* ‘close/bring to an end’. This is possibility (i), listed above.

Possibility (ii) would involve interpreting, on the basis what we know from other situations, a specific linguistic outcome as the result of a particular type of scenario. Thus, we may be confronted with a bilingual utterance in Copenhagen, say, involving Danish and Turkish (Steensig, 2001), and then with some confidence interpret it as an instance of codeswitching, as opposed to some other type of language contact scenario, such as second language learning or language attrition.

- (4) *Oğlan kız-a fri yap-sın*  
 boy girl-DAT propose do-CAN.3  
 ‘The boy can propose to the girl.’

However, things are not as straightforward as suggested by these examples. First of all, as stressed by Backus (e.g. 1996) and Muysken (2000, 2013), codeswitching may take different forms and have different linguistic outcomes. The type of codeswitching pattern we see in (1)–(4) is only one of the possibilities, falling under conventionalized insertional codeswitching. A more detailed discussion of codeswitching follows in Chapter 4.

Second, there is the risk of circularity. We need to be careful that we do not interpret the patterns found in (1)–(4) simply in terms of their linguistic characteristics, and then assume that they all result from the same underlying process or scenario. Different processes may have identical or at least similar outcomes.

Third, the scenario model is necessarily incomplete because there may always be bilingual practices not covered by any list.

Nonetheless, adopting these scenarios makes concrete comparisons with well studied other settings possible, on the basis of prototype modeling. The proposed model has the advantage of being fairly concrete, and allowing for a diversity of outcomes, but always based on concrete situations. Thus, for instance, convergence has been related to language loyalty, in a maintenance scenario, by Matras (2009) and Silva-Corvalán (1994). Speakers flag language loyalty through using the lexicon of the HL and reduce their processing load by making the grammars of two languages more similar.

### 1.3.3 Evaluating the scenario approach

The scenario model helps us to understand different types of HLs. These can be characterized in two perspectives: one is the more sociolinguistic perspective of language variation and change, which will be further elaborated in Chapters 3 and 4, and the other that of child language acquisition and early bilingual acquisition in general, which will be discussed in detail in Chapters 6 and 7. Here we will briefly sketch the basic issues involved.

If we limit ourselves to prototypical heritage speakers, namely early bilingual speakers who shifted to the dominant language during the school age, we find widely diverging results on their reported language. Some authors (Irizarri van Suchtelen, 2016; Kupisch, 2013; Nagy, 2015) find little to no difference between heritage speakers and homeland speaker in many domains, whereas other researchers report divergence in a number of domains (Benmamoun, Montrul, & Polinsky, 2013a, b; Montrul, 2008; O'Grady, 2011; Polinsky, 2006, 2008a, 2008b). We can understand these differences in outcome better if we take the scenario into account under which the HLs developed.

Three sets of factors that play a role in acquisition processes and as such contribute to the acquisition scenario and that we will discuss in more detail below are (1) positive versus negative attitudes towards bilingualism in society (cf. Kupisch, 2013, p. 206; Benmamoun, Montrul, & Polinsky, 2013a, b), (2) the type of social network the language is used in (Carreira, 2004; Chau, 2011; Extra, 2002) (3) the family situation in which the language is learned including factors like literacy, visits to the home country, sequential versus simultaneous language acquisition (Montrul, 2008; Kupisch, 2013; Kupisch & Rothman, 2018), caretaker background (Montrul & Sánchez-Walker, 2013) and the presence of (multiple) siblings. Kupisch (2018) summarizes the evidence for the idea that simultaneous bilinguals might be disadvantaged compared to sequential bilinguals in their HL because their HL has been in contact with the majority language for a longer period of time, so there has

been more pressure from the dominant language and comparatively less exposure to the HL during the crucial years (assuming that the input is roughly split in half for the simultaneous bilinguals).

Let us begin by looking at the role of societal attitudes towards bilingualism. Directly connected to the bilingualism attitudes are the concepts of additive versus subtractive bilingualism. These terms were coined by Lambert (1975, 1977, 1981) who contrasts bilingual settings in which the acquisition of another language is seen as an addition, as something extra and of worth, referred to as additive bilingualism, with subtractive bilingualism. This is a 'form of bilingualism experienced by ethnolinguistic minority groups who, because of national educational policies and social pressures of various sorts, feel forced to put aside or subtract out their ethnic languages for a more necessary and prestigious national language' (Lambert, 1981, p. 12).

The first language is most likely to be replaced by the dominant language when a country has a negative attitude towards bilingualism in general or in specific situations. In many cases the negative attitudes depend on the language pair. In the Netherlands, for example, if a young child knows English this is seen as an asset and learning English is promoted in schools (Lobo, 2013). In contrast, having a Berber background can be perceived as a problem (Nortier, 2012). Because most children would like to belong and to feel accepted, negative attitudes towards their home language in the school environment and in the wider community in general can cause them to give up the language (Wong Fillmore, 1991, p. 323). Carreira (2004) shows that attitudes about Spanish vary considerably in different parts of the US. Attitudes are much more positive in Florida than in Indiana, for example. Subtractive bilingualism would therefore be more likely to occur in Indiana than in Florida.

Studies suggest that subtractive bilingualism negatively affects the first language as much as direct influence from the dominant language (Wong Fillmore, 1991). Many reports on negative outcomes for HLs come from the United States and one possible reason for this is the effect of subtractive bilingualism. Many heritage speakers in the US feel pressure to shift to English completely and this negatively affects HL proficiency.

Another set of factors that various authors have proposed in relation to the heritage acquisition scenario is related to the social network that the HL is used in. Sometimes it is only used at home, while in other cases the language fulfills a much broader function in a community and in the surroundings. For Cantonese speakers in the Netherlands, Chau (2011) shows that those who live in Amsterdam have a broader Chinese network than those who live in the smaller town of Venlo. The Amsterdam Chinese hang out with their Chinese friends, watch Cantonese soaps and listen to Cantopop, whereas the Venlo-Chinese mostly have white Dutch

friends and are focused on Dutch culture. She shows that the Amsterdam Chinese know more slang and master more grammatical structures than the Venlo-Chinese who used more old-fashioned words and who have a more limited grammatical range. Moro (2016) shows that Moluccans who live in special wards with fellow Moluccans have a different pattern of language change than Moluccans who live in mixed neighborhoods. Carreira (2004) shows that Spanish in Florida, apart from being more valued, is also easier available in various domains such as in newspapers and multimedia. This availability of Spanish may boost heritage performance.

The last set of factors that affects the outcome of HL learning is the family situation. How many languages are spoken at home? Language input in the family setting has both a qualitative and a quantitative dimension, which should be kept apart indeed are taken into account in much of the bilingual acquisition search. Montrul (2008) finds that simultaneous bilinguals outrank sequential bilinguals in their linguistic proficiency and she relates this difference to the amount of input the learners receive in the sensitive period.

Montrul and Sánchez-Walker (2013) show that caretakers have an impact on heritage outcome: children who go to English speaking daycare develop and retain less of their HL than children who speak Spanish at home. Geense and Tsui (2001, p. 94) report on Chinese children in the Netherlands who stay with Dutch speaking guest families.

HL families differ from each other in the possibility and the willingness to make their children literate in the HL. Research by Tarone & Bigelow (2005) and Tarone, Bigelow & Hansen (2009) has shown that language intake in illiterates is different than in alphabetically literate speakers. Heritage speakers are usually literate in the dominant language and in that sense unlike the participants in Tarone's research but it is possible that no alphabetic script is available for the HL or that the heritage speaker is not educated in the alphabetic variant and that this causes a different kind of (meta) linguistic awareness when compared to heritage speakers who are literate in their HL (cf. Oller & Eilers, 2002; Benmamoun, Montrul, & Polinsky, 2013a). Moreover, if reading material is available this may enhance the richness of the input for the heritage speaker. Finally, the presence of older siblings has an effect on language use. Once older siblings go to school, they tend to switch to the dominant language. They bring the dominant language home, making high proficiency in the HL less likely in young children than in older children.

The take home-message of this section is that social settings such as the cultural norms concerning language mixing, caretaker effects and other aspects mentioned in this section matter for the outcome of language contact. Many differences between heritage speakers are reported and the contact scenario approach may explain part of these differences. This is further taken up in Chapter 6.

## 1.4 Overview of the book

This book is organized as follows. This first chapter has introduced the term HLs and related it to the broader field of language contact studies, including the scenario approach.

**Chapter 2** shows how the study of HLs developed. HLs have been studied for almost sixty years now – at least since Haugen (1953). Recently it has come to be studied from a wide variety of angles, bringing together linguists with different interests and theoretical orientations. However, it is worthwhile to delve into the history of the field, to better understand the way its mental frames are structured, and its embedding in the academic world. We will first turn to the **DIASPORA** perspective: what happens to a language when spoken beyond its original borders, through political developments or emigration? Notice that this perspective involves the fate of a single language. Then the very prominent **IM-MIGRATION** perspective will be explored: how do all kinds of different languages come together in a country with many immigrant groups? The chapter points to a third perspective, which characterizes the field at present: the **SPEAKER** perspective, further explored in the book.

**Chapter 3** places HLs in a social context. Speakers of the language at some point in history started using one or more other languages in daily life. The resulting bilingualism produced a situation in which people have to choose which language to use in any given communicative context. The long-term effect of myriad everyday language choices by all these speakers is the tug-of-war between language maintenance and language shift. If the language is maintained, it has entered a stage in which it is generally seen as an HL, meaning it is associated with the community's history and traditions, is actively used up to a point and primarily in the home domain, and undergoes changes as a result of domain restriction, loss of practice, and influence from the other language. Language shift is the possible endpoint of this process, when language choice favors the other language in all domains. If a language is not actively used anymore, it becomes an HL in the narrow sense of the word, as a language that is remembered as the vehicle of communication in the past, and of which people may remember the occasional word or ritual formula. The push and pull between maintenance and shift is the central point of discussion in this chapter.

In order to understand what causes languages to be maintained in a situation of bilingualism, or to be let go of, it is crucial to understand what determines language choice. Communicative aspects of language choice and bilingual language use are the topic of **Chapter 4**, which thus also analyzes situations in which the choice between languages is avoided. The clearest case is when both languages are combined, in the pattern of speech called 'codeswitching'. Explanations for why

people choose this way of speaking sometimes emphasize that HLs simply lack particular vocabulary and hence have to make use of the resources offered by the other language. Mixing in material from the other language often happens at a subconscious level in such cases. Other explanations appeal to more intentional decisions made by speakers, especially for pervasive codeswitching in which the two languages constantly alternate. Speakers shift back and forth between more monolingual and more bilingual modes of speaking, and individual switches are often made for particular reasons, having to do with the pragmatics of the conversation and with the social values indexed by the languages involved.

The superordinate level of 'language' is not always the right level at which to address such issues. Often, the focus is just on features such as words, pronunciation details, a particular construction or a discourse style which index particular values and connotations. Indexicality entails that people have attitudes about linguistic forms and may approve or disprove their use in different communicative settings. In the case of HLs, this purism may target any use of the other language, or the use of particular features in the HL that betray foreign influence or are believed to betray it. Such disagreements play a role in intra-community conflicts which in turn influence people's linguistic choices, and thus indirectly affect the battle between maintenance and loss. Chapter 4 presents several bilingual contexts and focuses on the motivations for people's linguistic choices and the consequences of these choices.

**Chapter 5** moves to a practical question. How do we collect data on HLs? Inherited from the study of language contact and minority languages is the practice of recording samples of spontaneous conversation and interviews. Such data are immensely valuable as a window on how the language is used in everyday life and what its structure and vocabulary look like. It can only be used, however, in settings in which the HL is still in use. As in sociolinguistics, HL researchers also make use of questionnaires, especially to gain information on attitudes and use patterns. A final methodological strand is informed by traditions in the study of second language acquisition and elicits data in controlled settings. This ranges from pencil-and-paper judgment tasks to laboratory-based psycholinguistic experiments.

Together these different ways of collecting data on the status and structure of HLs make it possible to triangulate findings and state conclusions with greater confidence. On the other hand, data do not always converge, and in this sense the use of different methodologies to study the same HL setting can also expose the full complexity in ways that would be impossible if only one method is used.

The chapter will start with an overview of the various methods that have been employed, with examples taken from various studies. It will show how recordings are handled, and deal with issues in transcription and encoding. Several judgment



and experimental studies will be reviewed, mostly with an eye towards when which methods seem most effective. Ethical concerns will also be raised. One concern throughout will be the degree of ecological validity, as it is not always easy to find a balance between the richness afforded by recordings and the control afforded by experiments.

If one has gathered data on heritage speakers a central question is whether the HL has undergone change. To establish change, one needs to have a baseline comparison, the topic of **Chapter 6**. This chapter, which contains a number of case studies, concerns two central questions. The first question is who we should compare heritage speakers to and thus who could provide the baseline. The second question is how we can compare heritage speakers to baseline participants. One challenge related to the how-question is inter-speaker variability in bilingual speaker groups and thus also in heritage speaker groups. This variability results from differences in time of exposure to other languages and loss of contact with original speech community norms. How can we compare two groups if variation in one of those groups (the heritage group) is so large? The first part of this chapter is devoted to making the comparison between heritage speakers and the baseline as neat as possible. It discusses the role of dialect variation, selective register input, attrited input and lack of literacy as factors that need to be controlled for when comparing heritage speakers to monolingual speakers. It also looks at how we can have the most effective comparison between second language learners and heritage speakers. The second part of the chapter concerns ways of dealing with inter-speaker variation. It discusses methods of categorizing speakers in order to deal with inter speaker variation such as speech rate correlations, vocabulary tests and speaker biographies.

Taking a linguistic perspective, **Chapter 7** moves to phenomena in HLs more generally. The chapter gives an overview of the types of variation and change that can occur in HLs. The chapter is divided into two main sections and one comparative final section. The first section discusses change and variation resulting from reduced input and reduced use, independent from the characteristics of the dominant language. The second section discusses change related to interaction with the dominant language. The third section concludes the chapter by comparing the effects of the two different sources of change. To what extent are the two sources really separate and to what extent do they interact or yield the same results?

Once data have been gathered, it is important to interpret them correctly. **Chapter 8** provides an overview of theories on variation and change in HLs and on the models they use to interpret and to predict change. A brief history and context per model are provided and all models are accompanied by a case study that illustrates the type of questions asked and the methodology chosen. The generative framework is illustrated based on work by Laleko & Polinsky (2016)

on topic and case marking in Japanese and Korean. The variationist approach is illustrated by a study by Nagy (2015) on Voice onset time (VOT) in three HLs in Toronto. The optimality paradigm is illustrated by a study from Koontz-Garboden (2004) on the use of progressive aspect in Spanish-English bilinguals. The usage-based model is illustrated through a study of contact induced grammaticalization by Backus, Doğruöz and Heine (2011), and a related cognitive model by a study on the competition between analytic and synthetic constructions in Moroccan Arabic in the Netherlands.

**Chapter 9** discusses HL speakers from the perspective of multilingual processing and tries to bring to bear some of the findings of experimental psycholinguistics on HL research. Many HL speakers are reported as speaking slower than non-HL speakers, often with pauses, with word finding difficulties, etc. Other speakers show no evidence of slow speech rate etc. but do reveal ‘interference’ or ‘co-activation’ of the dominant non-HL. Of course, both speaker types can be studied from a psycholinguistic perspective.

There is a growing research literature that needs to be considered when we address language processing in heritage speakers. First, we consider recent models of language processing in bilingual speakers, focusing on the models of Levelt and Dijkstra. The next section concerns the relation between production, comprehension, and acquisition, including different types of learning, etc. Subsequently, we discuss different types of HL speakers: how can psycholinguistic models account for these different types? Then the concept of ‘priming’ is explained and applied to HL speakers. What are its basic properties, and what is its role in language learning and language change? Finally, we discuss the issue of language mode and very briefly touch upon the relation between multilingualism and cognitive processing in HL speakers.

**Chapter 10** looks at a specific subtype of HLs, namely HLs in a post-colonial setting and it focuses on one such HL, namely Papiamentu. In contrast to many HLs spoken in immigrant communities that do not have a history of contact with the dominant language, languages spoken in a post-colonial setting do have a longer history of language contact. Hindi, for example, has been in contact with English since the British invasion of India and the subsequent founding of Imperial India as a British colony. In a similar vein, the complex relation between British English and West-Indian English and Caribbean Creole in the UK is rooted in a similar set of relationships in the Caribbean. The effect of long-time contacts on an HL will be illustrated in this chapter with the case of Papiamentu or Papiamento, the Creole language of the Caribbean islands Aruba, Bonaire and Curaçao.

Lexical influence is very substantial, while some loanwords are quite old, and in the case of creole languages, may date from the period of genesis of the HL. When the dominant language is also the original lexifier language of the creole

(as in the case of English and London Jamaican), it is hard to distinguish between loans and original lexicon. There may be several varieties of the HL from the very beginning of migration, which differ in the degree of influence from the dominant language. The dominant language may also exert structural influence on the HL.

**Chapter 11** moves to the political dimension of HLs. Their status and recognition are part of what may be called *diversity management*: the response from institutional entities, ranging from state governments to school administrations, to the existence of language diversity and multilingualism within their political or institutional boundaries. Diversity management can take various perspectives: HL exclusive, only the dominant language; somewhat exclusive, the dominant language and traditional community languages, but not immigrant languages; inclusive, all languages spoken.

Topics addressed in this chapter include reversing language shift and indigenous language revival, HL education, HL competence as a learning resource within the mainstream classroom, documentation of heritage varieties, and overall language policies regarding HLs.

**Chapter 12**, finally, contains an alphabetical list explaining the concepts and technical terms used in this book.

# History of the field of heritage language studies

## 2.1 Introduction

The topic of HLs has been studied by specialists for almost sixty years now – at least since Haugen’s *The Norwegian Language in America* (1953) but before – under various labels such as Minority language studies, Immigrant language studies, Spanish in the US, and Turkish in Europe. As noted in our introduction, recently it has come to be studied from a wide variety of angles, bringing together linguists with different interests and theoretical orientations. However, it is worthwhile to delve into the history of the field, to better understand the way its mental frames are structured, and its embedding in the academic world.

We will first turn to the DIASPORA perspective: what happens to a language when spoken beyond its original borders, through political developments or emigration (Section 2.2); notice that this perspective involves the fate of a single language.

Then the very prominent IMMIGRATION perspective will be explored: how do all kinds of different languages come together in a country with many immigrant groups (Section 2.3).

Currently, as will become clear in this book, the SPEAKERS’ perspective is focused upon: what are the characteristics and the experiences of the individuals who speak a HL and do we characterize the language use of these speakers?

We will notice many differences, at least at the outset, between North America and Australia on the one hand, and Europe on the other. Often these are concerned with the following question: where do you belong? The recognition of HLs in North America and Australia rather than in Europe is due in part to different conceptions of citizenship, defined in legal terms but with ramifications which go much beyond the law, and rooted in deeply held convictions. In Europe there is a tradition of *ius sanguinis* ‘law of blood’: citizens are all those with a particular ancestry, while in North America and Australia we find *ius soli*, ‘law of the soil’. All newcomers who are resident are recognized as belonging. In Europe, a small or large minority originating from another country, e.g. people of French or Italian descent living in Germany, would be perceived primarily as French or Italian, rather than new Germans. In contrast, people of German ancestry or people of

Russian ancestry living in Australia would be perceived as Australians (for these issues see the discussion e.g. in Extra & Verhoeven, 1993, 1998).

Geography also plays a role. In Europe many of the national territories are relatively small, and thus people from a neighboring state may be perceived as still more or less belonging to the same nation, since they live next door. National boundaries are more fluid, while ethnic boundaries seem more rigid. In North America, the dividing line between Mexico and the US is sharp, but Mexican Americans are just that, a special kind of Americans.

#### Main goals of this chapter

To give a historical overview of the field of HL studies.

To present the perspective on HLs as diaspora languages.

To present HLs from the perspective of the country of immigration.

## 2.2 The perspective of the diaspora languages

One perspective on HLs is that of the diaspora: this perspective focuses on the spread of a language beyond the borders of the country where it was originally spoken and compares its different varieties, for example it compares English spoken in India to English spoken in Kenya. This perspective includes HLs, as well as for example creole languages with a European lexifier, such as Jamaican Creole. Initially, this perspective was rooted in nationalism and post-colonial nostalgia on the part of the European nations. Just like the remains across the world of the Portuguese and Dutch trading forts and of the Spanish missionary churches bear witness to an earlier global presence, diaspora varieties of the European languages testify to former and in part still existing colonial networks. Thus, the fact that there are Spanish words in several Oceanic languages testifies to the Spanish colonial presence in the Philippines and surrounding areas.

Sometimes the diaspora HLs were studied with the expectation that they would turn out to be archaic versions of dialects of the homeland languages, a kind of living language museums as it were, since migration often took place a century or more ago. This expectation has not been borne out in many cases. Diaspora varieties often cannot be traced back to specific dialects and have undergone all kinds of other changes.

### 2.2.1 Dutch from a diaspora perspective

To take just take a few examples of the older European powers, in the colonial period Dutch underwent expansion as a colonial diaspora language to several

different regions. It was brought to Indonesia, the Cape province of South Africa, the Hudson Valley and New Jersey in the United States, as well as to coastal regions of continental South America, including Northern Brazil, Suriname, and parts of (former British) Guyana.

In certain parts of the eastern United States Dutch (later labeled Jersey Dutch) survived into the 19th century as an HL, and there are some documented samples. The following fragment reflects this variety (Prince, 1913). We first give the original Jersey Dutch form (spelled in a modified phonetic way) in line 1, then the gloss in line 2, the modern Dutch form in line 3, and the translation in line 4. Features commented on are marked bold:

- (1) *En kääd'l had twi jongers; de ene blêv täus;*  
a man had two boys the one stayed home  
Een kerel had twee jongens; de ene bleef thuis  
‘A man had two boys; the one stayed home;’
- (2) *de andere xöng vort f'n häus f'r en stât.*  
the other went forth his house for a town  
de andere ging weg his huis naar een stad  
‘the other went away from home to town’
- (3) *Hâi doğti ôm dât täus en z'n vâders pläk.*  
he thought about that home and his father's place  
Hij dacht aan dat thuis en zijn vaders plaats.  
‘He thought of that home and his father's place.’
- (4) *Tû zâide: äk zâl na häus xâne.*  
then said I shall to house go  
Toen zei ik zal naar huis **gaan**  
‘Then he said: I shall go home.’
- (5) *Mâin vader hât plânti.*  
my father had plants  
Mijn vader heeft planten  
‘My father has plants.’
- (6) *En tû de vader zâg 'm komme, hâi(S) xöng(V) äut*  
and when the father saw him come he went out  
En toen de vader hem zag komen, **ging(V) hij(S)** uit  
‘And when the father saw him come, he went out.’
- (7) *en mûten 'm en boste z'n zön*  
and met him and kissed his son  
en ontmoette hem en kuste zijn zoon  
‘and met him and kissed his son’

- (8) *en tû brogt 'm in h'm häus*  
 and then brought him in him house  
*en bracht hem toen zijn huis in*  
 'and then brought him into his house.'

Without pretending to do an exhaustive analysis of this fragment from the Prodigal Son, we notice that it contains many features typical of what we find in HLs.

- i. Even though the text runs on like ordinary Dutch, it is very limited in the lexical choices and the treatment of the theme, for example the word *kääd'l* for *man* in (2) suggests a limited awareness of register effects, since the term is associated with a more informal context than a biblical story.
- ii. It contains archaic or dialectal expressions, such as *xöng vort* [ging weg] 'went away' in (2); *boste* [kuste] 'kissed' in (7) (see also Abraham, 2011).
- iii. The lexicon is entirely Dutch, but the use of the pronoun in *h'm häus* [zijn huis] 'his house' in (8) is English-like, in that *h'm* resembles 'his' in its initial consonant; another possibility is that this is a simplification.
- iv. The word order has become more regular SVO rather than the SOV and verb in second position of ordinary Dutch, as evidenced by *zâg'm* in (6) where the verb *zag* 'see' is used before the object *'m* (him) rather than after it, as in Dutch. Whereas Standard Dutch has subject-verb inversion after an adverbial clause, as illustrated in (6), inversion is absent in Jersey Dutch as absent as illustrated by *t hâi xöng* 'he went'. Likewise, in (8) modern Dutch would have a postposition *in* where Jersey Dutch has a preposition. Lexical and structural reduction, archaic and dialectal forms, influence from the dominant language, and structural readjustment are often found in HLs. However, there is variability; in (4) traditional SOV order is maintained for the lexical verb *xâne* 'go' at the end.

Let us now move to other diasporic variants of Dutch and compare their contexts of use. In the Republic of Suriname, the original colonial language Dutch is very much alive, with many first language speakers. Since it is an official language of the country there are no heritage speakers. In Suriname Dutch competes with the vernacular lingua franca, Sranantongo, as well with a number of languages specific to different ethnic groups.

Dutch is also spoken on several islands in the Caribbean, such as Aruba and Curaçao (Dutch Antilles, now still part of the Kingdom of the Netherlands), and it was spoken on St. Thomas and neighboring islands (now part of the US Virgin Islands). In the various islands that together constitute the former Dutch Antilles (they have changed constitutional status in various ways), Dutch is only a first language for a small minority, and there are no HL speakers. In Aruba and Curaçao it is spoken as an elite and administrative language alongside Papiamentu (see Chapter 11).

In several places in the Caribbean, now extinct Creole languages with a predominantly Dutch lexicon have emerged: Virgin Islands Creole Dutch (also called ‘Negerhollands’) on St. Thomas and St. John (Sabino, 2012), and Berbice Dutch Creole (Kouwenberg, 1994), and Skepi Dutch Creole (Robertson, 1983) in Guyana. Here the languages brought to the New World by the enslaved Africans constitute the heritage for many descendants of enslaved Africans. Here are some examples of the now extinct Skepi [Essequibo] Dutch:

(9) *Ek da lo nau*  
 I PROG go now  
 ik daar loop nu  
 ‘I am going now.’

(10) *Ende ni lafe nag*  
 2.PL NEG walk yet  
 jullie Niet lopen nog  
 ‘You not gone yet?’

Just like in the case of Jersey Dutch we find archaic dialectal forms such as *ende* [< Zealandic *jender*] ‘2.PL’ in (10), but in Skepi Dutch only the word ‘heritage’ lexical remnants can be identified. The structural make up of this new language is totally different from Dutch, even in the older form that gave rise to the Creole. Although from the diaspora perspective Skepi Dutch could be grouped together with Jersey Dutch and other extra-terrestrial varieties, the classifying this variety as a HL would be a misnomer.

In Asia there were Dutch colonists in Sri Lanka and Indonesia, and of course Indonesia remained under Dutch control until 1948. Of the earlier settler communities, no speakers are left in Sri Lanka. There are some elderly people in Indonesia who still speak the language as an HL, but there has been no serious study of their language.

Dutch was also spoken in the Cape region of South Africa. In South Africa, no more Dutch is spoken (except by more recent immigrants to the Republic of South Africa), but the Dutch varieties of the early settlements have undergone a process of koiné formation and have been incorporated into a new language (along with Khoikhoi, Portuguese Creole, and Malay words and structures), Afrikaans. Thus, in South Africa, varieties of Dutch were transformed into Afrikaans.

Then in the 19th and 20th centuries Dutch migrants left for the American Mid-West and Canada and later to Brazil, New Zealand, Canada, and Australia. In the eastern US the language has died out, but the more recent immigrant communities in the American Mid-West, particularly in Iowa (e.g. Smits, 1996), preserve some HL speakers, as is the case in Brazil, New Zealand, Canada and Australia.



In Table 2.1 different varieties related to the Dutch diaspora are contrasted on some sociolinguistic and structural dimensions. We have selected three structural features, which are often discussed in the scholarly literature: (a) are verbs placed at the end in subordinate clauses (SOV), and in second position in main clauses (V2), or is the pattern of SVO (verb always after the subject) followed? (b) Are verbs inflected for person, tense, and number, or left bare and uninflected? (c) Do pronouns have different case forms when they occur as subjects, objects, and possessors?

**Table 2.1** Features of different varieties related to the Dutch diaspora, including the Netherlands itself

	Netherlands, Flanders, Suriname	Australian Dutch	Jersey Dutch	Afrikaans in South Africa	Virgin Islands Creole Dutch	Berbice Dutch, Skepi Dutch
<i>Profile</i>	Original language	Recent HL	Older HL	Restructured koiné variety	Creole	Mixed creole
<i>Clausal word order</i>	OV, V2	OV, V2 some SVO	Some SVO	OV, V2	SVO	SVO
<i>Verbal inflection</i>	Person and TMA inflection	Present but erratic	Present but erratic	Some remnants of inflection	No inflection	Some TMA inflection, but not like Dutch
<i>Pronominal case</i>	Regular case distinctions	Regular case distinctions	Regular case distinctions, some anomalies	Case distinctions partly maintained	No case	No case

When we compare the varieties in Table 2.1 we can observe that some parts of Dutch are more vulnerable in contact situations than others and that variants differ in how far they have diverged from Dutch as spoken in the Netherlands, Flanders, and Suriname. In this comparison, the diaspora perspective may be valuable, answering questions such as: are there generalizations to be made about the way Dutch interacts with other languages in a diaspora context, either grammatically or sociolinguistically? Thus, for instance, Australian Dutch shares many features with the older Jersey Dutch, but the latter shows more differences from the original languages. However, the perspective in Table 2.1 is only partially relevant to HL studies in the strict sense, since many of the languages involved are not HLs but new languages that have Dutch as their lexifier. Still, the contact processes that HLs undergo can often be profitably studied in a broader context of language contact studies, including creole studies.

### 2.2.2 Other diaspora varieties

Parallel to the European colonial expansion, and in many cases even preceding it, is the diaspora of various Middle Eastern and Asian languages, especially Arabic, Chinese, Hindi, and Japanese. Arabic underwent a tremendous expansion, of course, with the rise of Islam, but is generally not an HL, except for Arabic communities in the United States, and the post-World War II Moroccan and Algerian diaspora to Western Europe (Versteegh, 2014). In the early twentieth century, Lebanese Arabic speakers migrated to many areas of the world, including Brazil, Argentina, the United States, Venezuela, Australia, and Mexico.

Chinese was brought to parts of the Pacific and Indian Ocean starting in the 10th century, was brought further afield in subsequent centuries, and is now a truly global language (He & Yun, 2008). It will be further discussed in this book as an HL in the Netherlands.

North Indian languages spread throughout the southern hemisphere, following the Indian mercantile presence around the Indian Ocean, reaching as far as East Africa from the early Middle Ages onward. Particularly, the recruitment of contract laborers from northern India in the wake of the abolition of slavery in the mid-nineteenth century led to the establishment of speakers of Indian languages in many places around the world: Fiji, Mauritius, Trinidad, Guyana, Suriname. The development of Indian languages outside India has been documented in Barz & Siegel (1988), where several case studies are presented including one on Sarnami Hindustani in Surinam, which is quite typical HL in a diaspora setting.

The Japanese likewise have a long history of emigration. From the 12th century onward there were Japanese communities in the Pacific, but the major emigration of Japanese, principally to Brazil and the US, only started in the early twentieth century. There are now Japanese rural and urban communities in several South American countries, and the Japanese American community is likewise very large. In the 1980's the Japanese government tried to bring members of the Japanese diaspora back to Japan, an event which attracted both public and academic attention. It brought into relief how much both the culture and the language of the diaspora Japanese had changed, and the event was widely discussed in diaspora cultural studies. There have been numerous programs promoting the teaching of Japanese as an HL, for instance.

In addition to the Asian and European post-colonial diaspora HLs, there are also cases of diaspora languages resulting from other political developments. After the redrawing of national boundaries in the wake of World War I, many Hungarian speaking communities were located outside of the borders of Hungary, e.g. in Romania, Serbia, Slovakia, and Austria. In addition, many Hungarians migrated to the US and Australia. These different varieties have been investigated from

various angles. Altogether, Hungarian has been studied as a diaspora language in some detail (Vazsonyi, 1995; Fenyvesi, 2005; Kontra, 2006). Kontra (1998) reports on the *Sociolinguistics of Hungarian outside Hungary* project, based on a survey gathering sociolinguistic data in Slovakia, Ukraine, Serbia, Austria, Slovenia, Romania, and Croatia. Fenyvesi (2005) is a collective volume with studies by all major researchers in the field, while Kovács (2001) has written on Finnish and Hungarians in Australia, including an analysis of their codeswitching practices. Among the grammatical constructions studied in a number of these diaspora communities, we find particularly nominal plural marking, case marking, and some verbal constructions.

### 2.2.3 Diaspora studies in a broader perspective

In the domain of cultural studies, diaspora studies have developed into a separate discipline, with a dedicated journal, *Diaspora: A Journal of Transnational Studies*, introductions such as Cohen (1997) and specialized volumes, such as Vertovec & Cohen (1999), and Evans Brazziel & Mannur (2003). These studies are oriented on a much wider domain than language alone and study the depiction of for example heritage stories and heritage food in literature and movies. An important message that follows from these studies is that in diaspora, languages and cultures always undergo changes and transformations that can be studied together in a single analytical framework.

## 2.3 The perspective of the country of immigration

Diametrically opposed to the diaspora perspective is that of the immigration perspective where multiple immigrant languages are studied from the perspective of the immigrant country. In this perspective, the United States, Canada, and Australia have taken the lead.

### 2.3.1 The United States

HL studies of migrants have a particularly long history in the US. A good starting point is Haugen (1953; see also Haugen, 1956 and Muysken, 1997). This study was the beginning of a series of works on immigrant languages in the US, which often went under names like 'American Finnish' or 'American Swedish.' Sometimes more localized terms are used, such as New Jersey Dutch, Pennsylvania Dutch (a form of German), Iowa Dutch, or Texas German. With increased attention to the large Hispanic population of the US and their language use, a considerable

literature has developed on Spanish as an HL. In recent years also Hindi, Russian, and the Chinese languages have become the object of scholarly research.

This older tradition of studies appeared in the United States typically with titles such as *The X language in America*. The series of serious investigations of migrant languages in the United States has a not strictly academic precursor in Mencken, someone who would now be called an ‘independent researcher’ – a term used for the many academics, linguists and others, who want to carry on with their research even though the academic job market has no place for them. H. L. Mencken’s *The American Language* (1936) contains thumb-nail sketches of the status and features of 28 immigrant languages in the United States, from German to Gypsy (Romani, to be sure), from Ukrainian to Japanese (pp. 616–697). Included are data from the 1930 census, loan words, syntactic interferences, phonological adaptations, particulars of dialect origin, data on bilingualism, language use, the media, etc. Haugen (1953, p. 13) finds it ‘entertaining,’ though something less than reliable’. Interestingly, Mencken’s sketches are included in an appendix called ‘Non-English Dialects in American,’ suggesting that ‘American is a far wider concept than English’. Mencken’s work had the merit of arousing interest in the topic, even if his presentation is somewhat anecdotal. He gives examples of loan blends such as *brandman* for ‘fireman’ in American Swedish (Swedish has *brandsoldat* ‘fire soldier’), parallel to *postkort* ‘postcard’ (<Sw. *breffkort* ‘letter card’) and *familjemedicin* ‘family doctor’ (<Sw. *husläkare* ‘house doctor’). His many examples of loan words with phonological and morphological adoptions from the HL, such as *abbordare* ‘to board’ and *giumpare* ‘to jump’ in American Italian suggest highly frequent borrowing from English in the immigrant varieties. Here Haugen (1953, p. 65) puts in a word of caution: ‘Most accounts of borrowing exaggerate the extent of it, since they are written either to discourage it or to awaken amusement’ (see also Louden, 2016). The tradition started by Mencken is characterized by the following features:

- a. The native language of migrants has a central place;
- b. Most space is devoted to phonetics/phonology and lexicon;
- c. The original dialect background of the immigrant varieties is studied;
- d. There is much attention to publications and media in the native languages.

To be sure, many of the languages considered were European immigrant languages with a writing tradition. Of Mencken’s 28 surveyed languages, 22 fit into this category directly.

The reason why this tradition has become popular in the United States may lie in cultural history and the historically determined self-image of the country. Haugen (1956) distinguishes four types of languages in the Americas:

1. native [such as Navaho]
2. colonial [such as English, Spanish and French]
3. immigrant [German, etc.]
4. creolized [Gullah, Haitian]

The concept of ‘native’ is relatively clear, it refers to the language of the original or native inhabitants of the country, but the distinction between ‘colonial’ (languages from European colonizers) and ‘immigrant’ – the language of more recent immigrants- is less clear, since some languages (Dutch is an example) are both present in the colonial period and have substantial groups of later immigrant speakers. In addition, of course English is ultimately also an immigrant language; the crucial difference, however, is that English-speaking immigrants came to an English-speaking country, whereas speakers of, say, Polish, had to learn English. The concept of ‘creolized’ is clearly demarcated only insofar as we concentrate on the clear cases such as Haitian.

It is more difficult to use a term such as ‘official’ or ‘national’, since legislation varies widely with respect to the legal status of the languages. Furthermore, the status of a language may differ from area to area. French in northern New England has a different status from that in Quebec, Spanish in Monterrey (Mexico) has a different status from that in Texas, and English in Costa Rica has a different status from English in Belize. In any case in the European context there is no such separate category of colonial languages. Perhaps this explains the rather early American interest in immigrant varieties.

### 2.3.2 Early studies on ethnolects and Canadian HL research

In the Canadian context the term ‘ethnolect’ was used for a type of language that most people would now refer to as HLs. For example, Danesi (1984, 1985) referred to Toronto Italian also known as *italiese* (< *italiano* + *inglese* ‘English’) or *Italo-Canadian*, as an ethnolect. Toronto Italian differed from Italian from Italy (original Italian) according to Danesi. This variety is mostly characterized, according to Danesi, by lexical innovations. Examples are given in Table 2.2. It is important to distinguish ethnolects, in the sense that the term is used in most current literature, from HLs, and to indicate possible links between these two types of varieties.

Verbs all have the first conjugation, are inflected as in Italian in the normal way: e.g., *puscio* ‘I push’; *ho pusciato* ‘I have pushed’; *puscerò* ‘I will push’.

Danesi’s work set an example for related other studies of HLs in the Canadian context. A good example is Portuguese as a HL in Winnipeg, as studied by Mota (1997). There similar tables are presented (here some examples of the words in his list starting with *b* and *c*):

**Table 2.2** Canadian Italian vocabulary in Toronto (based on Danesi, 1984)

Canadian English equivalent	Toronto Italian	Original Italian
store	storo	Negozio
sink	sinco	avandino/acquaio
cake	checca	Torta
mortgage	morgheggio	ipoteca/mutuo
fence	fenza	Recinto
ticket	ticchetta	Biglietto
to push	pusciare	Spingere
to paint	pintare	Verniciare
to freeze	frisare	Congelare
smart	smarto	Intelligente
cheap	cippe	Economico

**Table 2.3** Canadian Portuguese vocabulary in Winnipeg (based on Mota, 1997)

Canadian English equivalent	Winnipeg Portuguese	Original Portuguese
bike	a baique	bicicleta
boring	borim	mascador
backlane	a becline	a rua traseira
business	a bisnas	o negocio
carrot	o carrote	a cenoura
cheap	chiparia	barato; fraco
corner	a coma	a esquina; o canto

The tradition that started with Danesi focuses on HL use (often based on questionnaires) and lexical innovation. The latter is indeed striking in the several cases mentioned. Apart from work by Nagy and Danesi another Canadian perspective is given in the studies on early bilinguals by Genesee, Nicoladis, & Paradis (1995) and Paradis & Genesee (1996) who look at the simultaneous acquisition of French and English in young children. Here the relation between the two languages is quite different from that sketched by Danesi, however: the two are the official languages of Canada.

The way Danesi uses the term *ethnolect* differs from how it is most commonly used nowadays, as denoting a variety of a dominant language used by a specific non-dominant ethnic group and characteristic of that group. The first use of the term with this meaning, as far as known to us, was in Carlock & Wölck (1981). In their case, reference was to English as spoken in the Polish-American community

in Buffalo. The Polish ethnolect of English, although its lexicon is heavily English-based, followed the prosody of Polish, and was recognizable as typical of the way Polish Americans spoke English.

At first sight, it may just seem like a matter of confusion that Danesi (1984) uses the term ethnolect to describe Toronto-Italian. But if we think of ethnolects as all the language variants that a specific ethnic group in a specific country may use to signal their ethnicity than Toronto-English is just as much an ethnolect as Italian accented English or English-Italian codeswitching, designating the result of language maintenance (see Chapter 3), distinct from the standard use designating the result of shift of a migrant group to the dominant language. However, it may be that in some immigrant groups, language maintenance and codeswitching, including the creation of an HL, are the functional equivalent of shift-related ethnolects that characterize other communities. This would suggest two alternative definitions of ethnolect:

*Ethnolect narrow:* Variety of a dominant (often national) language spoken by a specific (non-dominant) ethnic group

*Ethnolect broad:* The varieties in the repertoire of a non-dominant ethnic group used in a larger context (includes HL, codeswitching, etc.)

In the broad definition, then, HLs would be part of the ethnolect (see also Clyne, 2000).

### 2.3.3 Case studies of HL languages in the United States

We will now discuss as an example a few studies, carried out by Pap (1949), Lehtinen (1966), and Seaman (1972) of immigrant languages in the United States. The best-known study is Haugen's (1953) already mentioned two volume study on the Norwegian HL communities in the United States. Haugen's work has been widely praised since it is both a cross-generational historical sociolinguistic study of the bilingual community of successive waves of Norwegian immigrants to the United States (Volume I), and a detailed linguistic study of the gradual incorporation of English elements (pronunciation features, words, phrases, grammatical patterns) into American Norwegian (Volume II).

Other relevant studies on immigrant varieties in the United States include Benson (1960, Russian), Dudek (1925, Czech), Espinosa (1930, 1946, Spanish in New Mexico), Green (1961, Yiddish), Henzl (1982, Czech), and Lyra (1962, Polish). Hernandez Chavez, Cohen, and Beltrano (1975) and Pfaff (1979) constitute follow up examples with studies on Mexican Americans, precursors to the enormous amount of current scholarship in this area. By way of illustration, our

focus will be on the way English nouns (in Finnish) and verbs (in Portuguese and Greek) have been incorporated into these HL varieties.

### *American Portuguese*

Pap (1949) has analyzed the Portuguese of immigrants residing mostly in New England, basing himself on observations, newspaper articles, and recordings. To gain an impression of his work it is worthwhile to consider his table of contents, which shows concern for sociolinguistic and historical aspects as well as language contact phenomena:

1. Introduction
  - Settlement history
  - Economic conditions
  - Cultural and social traits
2. Speech conditions in general
  - The position of Portuguese in relation to English Agencies in the defense of Portuguese
  - Transformation of Portuguese Speech
3. Survival of regional and popular speech traits in immigrant Portuguese
  - Phonology and grammar
  - Vocabulary
4. The influence of English on Immigrant Portuguese
  - Phonology and grammar
  - Vocabulary: formal analysis
  - Vocabulary: semantic analysis
5. Proper nouns
  - Portuguese-American personal names
  - Names of Portuguese-American organizations
  - American geographic names in Portuguese Immigrant speech
6. Conclusion

Turning to English verbs, Pap shows that many verbs are incorporated into the (quite productive) Portuguese *-ar* conjugation class. However, a few verbs in this class receive an extra *-e-* between the verb stem and the ending, resulting in *-ear*:

<i>bord-ar</i>	'board, live in a boardinghouse'
<i>fris-ar</i>	'freeze'
<i>chinj-ar</i>	'change'
<i>jamp-ar</i>	'jump'



<i>raid-ear</i>	'ride'
<i>fait-ear</i>	'fight'
<i>pamp-ear</i>	'pump'

The *-ear* ending is also common with English verbs adopted into Spanish as well, as in *boxear* 'to box' or *parquear* 'to park'. We do not know why there are fixed channels for verb integration, like the affixation of *-ear* in Spanish, which parallels the use of *-ear* in American Portuguese. Part of the answer is that the *-ar* conjugation is the most productive one, but why then the extra *-e-* vowel? It may be that a marked subclass is created to flag the non-native character of the elements involved. It could also be due to the factive or causative character of the *-ear* suffix ('to make something X'), as in monolingual Spanish *blanquear* (< *blanco* 'white') 'to whiten, to make white'. This makes it a derivational morpheme parallel to the use of a 'do' verb, to which we turn now.

Pap provides several examples (1949, p. 114–117) of verbs integrated with *fazer* 'do, make' (in addition to many cases where the verb has been adapted directly to Portuguese inflection). In standard Portuguese, this construction occurs, but it is only used occasionally in specific lexicalized combinations.

<i>fazer o chinche</i>	'change (money)'
<i>fazer o telefone</i>	'telephone'
<i>fazer o save</i>	'save'
<i>fazer o find-out</i>	'find out'
<i>fazer o give up</i>	'give up'
<i>fazer o fool</i>	'fool'
<i>fazer o boda</i>	'bother'
<i>fazer o spoil</i>	'spoil'

In American Portuguese sometimes, the embedded verb is nominalized with the Portuguese masculine article *o*. Pap does not raise the issue of why sometimes a helping verb or light verb is used, and at other times the verb has Portuguese inflection. It is possible that syntactic and phonological factors sometimes preclude verb integration into Portuguese. In the cases of *give up* and *find out*, there is no V + particle construction in native Portuguese. In the case of *bother*, the final liquid is not a proper base for suffixing Portuguese *-ear*. In the Portuguese of this community, *fazer* is also used with nouns, as in *fazer dinheiro* 'make money'. The use of *o* makes a nominalization analysis for verbs quite plausible. Furthermore, Pap (1949, p. 106) notes: 'Since the English word in these combinations appears as a noun, the direct object of a transitive phrase of this kind is preceded by the

preposition *de* (of).’ Absorption of transitivity results from the fact that the English verb becomes the object of *fazer*, which then leads to a new intransitive verb, making a preposition necessary.

In addition, there are also a few cases of bare verbs used as the complement to *fazer* (1949, p. 105), as in *fazer box* ‘to box’. Finally, in rare cases there are bare English verb stems, as in the following infinitive construction:

- (11) *Ainda nao é tarde para insulate as vossas casas.*  
 yet not is late for insulate the your houses  
 ‘It is still not too late to insulate your houses.’

### *American Finnish*

One of the earliest studies with extensive discussion of codeswitching data is Lehtinen (1966). Meri Lehtinen used recorded speech sample of over ten hours (consisting mostly of narratives) from only one speaker, a third generation Finnish-English bilingual immigrant from Minnesota, who was a graduate student in linguistics but had no formal training in Finnish. In addition, she had ample fieldwork experience with immigrant Finnish in general. Lehtinen, a Finland-born bilingual, conducted the interview herself.

A first result is that most switches are English bare nouns without determiners, often though not always accompanied by a Finnish case-marker (Lehtinen, 1966, p. 226):

- (12) *sitte meni ... Kansassii*  
 Then I.went ... Kansas.ILL  
 ‘then I went to Kansas ...’

A second common feature in Lehtinen’s corpus, which was recorded with a non-fluent bilingual, is the high incidence of flagging and hesitations, as seen in the hesitation pause in (12), which may signal a switch to the other language.

A third feature is the high incidence of semantically empty Finnish nominal determiners, such as *semmonen* (‘such’), as in (13), a typical feature of bilingual speech in Finnish when foreign items are ‘flagged’ (Lehtinen, 1966, p. 225). Here *semmonen* does not really mean ‘something like a hay stack’ but is simply a marker of a switch:

- (13) *me teki semmonen hay stack ja ...*  
 we made such hay stack and ...  
 ‘We made a hay stack and ...’

Lehtinen argues that English nouns and adjectives incorporated into Finnish obligatorily carry a stern formant *i* (if they end in a consonant), except where

the appropriate Immigrant Finnish case marker would be null, in which case *i* is optional (1966, p. 180). Furthermore, objects in Immigrant Finnish are often marked null (1966, pp. 44–46). Thus, it would be difficult to see whether the object was case-marked or not. Although borrowed objects frequently have null marking with regards to case, all English nouns in the corpus are preceded by a determiner such as *semmonen* ('such'), in (13) and *se* ('that') in (14) (Lehtinen, 1966, p. 225):

- (14) *ne ne tee.. teki se hospital*  
 they they ma.. make that hospital  
 '... they made the hospital.'

The typological differences between the languages in terms of case marking and agreement lead to a very limited insertion strategy, which is complicated further by the phonological differences between the languages. Nonetheless, case and verb agreement are suffixal and can be added to English nouns/adjectives and verbs, respectively, once these are morphologically adapted to Finnish. Nouns are often incorporated into Finnish noun phrases using a Finnish determiner, which carries the case marking. Consider a case such as (15) (Lehtinen, 1966, p. 182):

- (15) ... *se oli semmosen typical general store tiäkkö, ja ...*  
 ... it was such typical general store ya-know, and ...  
 ... it was such a typical general store you know, and ...

It may be that the true head of the predicate noun phrase here is Finnish *semmosen* ('such'), with the English expression *typical general store* simply adjoined, without any evidence that it is integrated into the noun phrase (cf. Muysken, 2000, p. 94). Recall that in Finnish there is agreement between the determiner and the head.

Altogether, the typological differences between the languages involved a specific HL setting play an important role in the nature of changes that the HL undergoes.

### *American Greek*

American Greek is well represented by the work of Seaman (1972), who worked in the Chicago area. Seaman deals with the socio-cultural background of Modern-Greek and with American-English contact, both on the phonological level and on the morpho-syntactic level of Greek. There is an extensive discussion of the lexicon.

Seaman did not conduct a quantitative study of American Greek. Nonetheless, the ample exemplification allows us to compare the frequency of the different patterns of adapting English verbs. Cases typically have the form of (16) where an English verb (participate in (16) is preceded by the semantically underspecified action verb *kani* ('make' or 'do')

- (16) *O Petros kani participate.*  
 DET Petros do/make.3SG participate  
 ‘Petros is participating.’

Again, the original use of *kani* is that of a verbalizer ‘do’ or ‘make’ used with noun phrases, as in (17) and (18):

- (17) *O Petros kani mia prosfora.*  
 DET Petros do/make.3SG one remark  
 ‘Petros makes a remark.’

- (18) *O Petros kani banyo.*  
 DET Petros do/make.3SG bath  
 ‘Petros takes a bath.’

Seaman shows that two verbs participate in the construction: *jino* ‘be, become’ and *káno* ‘do, make’, and notes (1972, p. 169):

This practice of avoiding complex or little-known Greek verbal constructions by the substitution of hybrid predicates is universal in the Chicago Greek communities and occurs in the speech of most Greek-American bilingual speakers of all generations.

Some examples of the patterns encountered are:

<i>káno cover up</i>	[make.1SG cover up]	‘I cover up.’
<i>káno delivery work</i>	[make.1SG delivery work]	‘I do delivery work.’
<i>kánune fishing</i>	[make.3PL fishing]	‘they fish’
<i>kánune feast</i>	[make.3PL feast]	‘they feast’

Such forms have also been noted in other varieties of diaspora Greek and reflect an almost universal tendency to use light verbs in verbal compounds in bilingual contexts (Muysken, 2016).

### 2.3.4 HLs in Australia

The interest in HLs in Australia notably came with the pioneering work of Michael Clyne (e.g. 1982; Clyne, 2003; Clyne & Kipp, 1999). In Australia, work of researchers like Haugen was picked up by Clyne in the 1960s, who studied the German and Dutch of immigrants to Australia. Issues he picked up on were attrition phenomena, codeswitching, and grammatical convergence towards English (1967, 1976). One of Clyne’s contributions in this area is the notion of triggering as a factor facilitating codeswitching. In later work he extended his research to all immigrant

groups of Australia (2003). Clyne and Kipp (1999) compare many heritage communities in Australia with regard to the degree with which the HL is maintained, showing a surprising spread, with Vietnamese being highly retained, while a language like Dutch was rapidly abandoned. Greek and Italian are intermediate in this respect as HLs. In recent years, Clyne's project has been taken up by many other scholars, including Hlavac (2012) on HL Croatian-English codeswitching in Australia, Guanglun (2014) working on the role and value of HL Chinese in Australia, and Alvanoudi (2018) on contact phenomena in HL Greek in Australia.

### 2.3.5 The European context

In the European context, the study of the linguistic situation of adult migrants has tended to concentrate on patterns of second language acquisition of the dominant language rather than use of the HL. The focus of these works has been on syntax and semantics, to a lesser extent on the lexicon and morphology, pragmatics, and phonology. Studies that do look at both languages of bilinguals focus on simultaneous or early sequential bilinguals, such as Meisel (1989, 2014) or Muller & Hulk (2001).

A good example in the tradition of the study of the second language acquisition of the dominant language is the large international European Science Foundation project (Klein & Perdue, 1992). This project was directed at untutored second language acquisition by adults, following earlier German research by Klein & Dittmar (1979) and Clahsen, Meisel, & Pienemann (1983). There has also been much research, often with a more immediate educational background, directed at migrant children in schools. Here literacy skills, grammar, and increasingly lexical and productive skills are the subject of study. The child studies are often carried out in primary schools, and the adult studies in training centers etc. In the meanwhile, there was little knowledge of home language use, and of in-group interaction in naturalistic settings. There was less attention for the linguistic characteristics of the languages involved, processes of dialect leveling, nor contact-induced language change. In the domain of phonology, there is the work on Spanish-German child development of Lleó & Kehoe (2002), Kehoe, Lleó, & Rakow (2004), and Kehoe, Lleó, & Rakow (2011).

We will begin the survey of the European history of studies in HL with the Linguistic Minorities Project survey of minorities in England (1985) as an early European example. Notice the shift to speakers from languages. In fact, individual properties of languages do not play a role in this sociological study. Its contents reflect the changes that have taken place. After a discussion of the history and background of linguistic minorities in England and of the role of bilingualism in society, there is a discussion of surveys of adult language use. The focus is on

patterns of language use, the provision of mother tongue teaching, and bilingualism and education. The latter is studied in *Schools Language Surveys*. The main recommendations concern conditions for sharing languages in the classroom. Aghinotri's (1987) work on the Sikhs in Leeds, Li Wei's work on the Chinese (1994) in Tyneside, and related studies are much more focused on the native languages as such of migrants in England.

In continental Europe, the interest in immigrant HL varieties gained ground later. There were occasional studies of Italian and Spanish immigrants in northern European countries in the 1970s, but the main work has been on Turkish and Arabic. An example is the work in Germany by Rehbein, Herkenrath, & Karakoc (2009a, b). Some of this work will be discussed in detail in the following sections, but the number of references is large. Work on Spanish as an HL has continued in Sweden, e.g. in the work of Bylund & Díaz (2012). Johansson (1992) has collected preliminary work by others on Turkish in Europe and elsewhere. Vermes & Boutet (1987) contains some work on Arabic in France. Most of the earlier work on migrant languages in Europe has been done in the context of language-contact and codeswitching research, however. Good examples are Nortier (1990) on Moroccan Arabic and Backus (1996) on Turkish in the Netherlands. Extra & Verhoeven (1993) contains useful overviews of a number of early studies immigrant languages in the Netherlands.

### 2.3.6 Summary

The purpose of this section was to draw attention to the country of immigration perspective, which has led to a rich but not often cited series of sources from the 1940s and 1950s for immigrant languages, as a partial model for new HL research. To be sure, these sources represent somewhat antiquated types of research, in terms of data gathering methodology and analytical tools. Nonetheless, they constitute a different and often original perspective and could well inspire new research in the global context.

## 2.4 Summary and introduction of the speakers' perspective

In this chapter the two main perspectives were introduced that guided the development of HL studies in the past: that of the country of origin of immigrants, the diaspora perspective, and that of the country of destination, the immigration perspective. These two approaches have enabled broad comparative studies of commonalities and differences as languages spread over the world or as different languages had to adapt to the same immigration environment. However, these

approaches are less successful in accounting for the processes that lead to the formation of HLs at the level of the individual speaker. Therefore, current HL studies, as noted in Chapter 1, depart from neither of these two approaches. Rather they adopt a perspective taken from sociolinguistics and psycholinguistics in abroad sense and focus on the speakers themselves. The perspective taken in many contemporary HL studies is in line with the cognitive concerns of modern linguistics. It is also the main perspective taken in this book. The speaker perspective is broad, focusing both on the identity of HL speakers and on their language processing and proficiency.

## Social aspects of heritage languages

### 3.1 Introduction

Our objective in this chapter is to place HLs in their social context. Once speakers of a language are bilingual, they must choose many times a day which language to use. The long-term effect of myriad everyday language choices by all these speakers is the tug-of-war between *language maintenance* and *language shift* (Fishman, 1972). If the language is maintained, it is likely to undergo many changes because in particular contexts it is not used anymore and because it is impacted by the other language. Language shift is the possible endpoint of the process when language choice more and more favors the other language in all or most domains. The push and pull between maintenance and shift is the central point of discussion in this chapter.

It is organized as follows. Section 3.2 discusses the importance of attending to both social and linguistic factors when studying language contact, while 3.3 focuses on maintenance and 3.4 on shift. Finally, Section 3.5 looks at the most eye-catching aspect of language use in many bilingual communities, when people do not choose to use just one of the languages in a particular communicative setting but use both, in the communicative mode called codeswitching.

#### Main goals of this chapter

To understand why bilingualism always involves the push and pull between maintenance and shift.

To understand the factors that influence the outcome.

To understand why the social environment in which an HL is spoken determines much of what happens to the HL.

To understand what changes occur in languages caught up in language contact settings.

### 3.2 The scenario approach: Attending to social and linguistic factors

Like all languages, an HL exists in a social world, and it is the reality of the world that determines what happens to the language. It is, therefore, useful to examine HLs in their social contexts. This is typical of contact linguistics more generally,



as it was recognized long ago that the linguistic phenomena found in language contact situations are brought about by the interaction between social and linguistic factors. As we saw in Chapter 1, the scenario approach focuses on the idea that there are specific socially determined language contact settings with specific linguistic outcomes. That is if language A and B come into contact, the outcome of this contact depends on the social situation. The fact, for instance, that Pennsylvania German has Anglicized grammatical patterns such as the simplification of its case marking has as much to do with the fact that the language is spoken in a social setting in which American English is dominant as with the possibilities that German offers for incorporating lexical and grammatical features from English. A modern German expat in London may show contact effects in his German that are very different, because the social setting is different.

Not only social factors shape the outcome of language contact. The typological characteristics of the two languages are also important. Descendants of Finnish emigrants to the US may speak a version of American Finnish, and this too may be characterized by extensive grammatical influence from English, yet, the precise forms this influence takes will differ from those of Pennsylvania German. The social setting may be similar, but the typological characteristics of the participating language are not. Nevertheless, the field has been characterized by separate socio-linguistic and linguistic accounts, enough so for Thomason & Kaufman (1988), and also Johanson (1992), to emphasize that a full account of contact-induced change requires attention to both social and linguistic aspects, and to the relationship that holds between them. This can be analyzed with a usage-based approach to linguistic competence (Bybee, 2010; Backus, 2013) or within a generative framework with an approach such as Feature Reassembly (Putnam 2019; Putnam et al. 2019).

This chapter starts out with a consideration of the all-important continuum between language maintenance and language shift, describing the factors that determine whether a minority language is maintained as the main language of communication within the minority community.

We will look briefly at the main global phenomena that bring bilingual settings into being, immigration and colonialism. In keeping with the idea that it is people's actions that determines the fate of the languages they speak, we zoom in on what the implications are of societal or community bilingualism for the daily life of the actual speakers of the HL. The central topic here is language choice, and the degree to which choosing one or the other language alludes to aspects of the speaker's identity, with all its ramifications concerning how languages index, or 'stand for', particular values. In addition to the social message speakers aim to achieve through language choice, their language use is often determined by factors that have little to do with ideology and identity, but more with what they can do, and what they feel comfortable with, in either language. Throughout the chapter it

will be seen that a bilingual's proficiency in each language also impacts language choice. Finally, the chapter will close off with a consideration of what happens to the other language in the bilingualism equation, the language that is not the HL. Sometimes bilingualism produces an ethnic variety of the majority language, a so-called 'ethnolect'. See Chapter 2 for a broader discussion on the relation between ethnolects and HLs.

### 3.3 Maintenance

The main social question regarding bilingualism in any context is whether bilingual usage is going to endure or not. If a language is completely given up, the community is said to have undergone *language shift*: it has swapped its original language for another one as the basic vehicle for communication. If, on the other hand, the ancestral language continues to be used in this function, even if only for communication within the community, we speak of *language maintenance*. Discourse about minority languages is often about whether language shift can be avoided. This is only natural: the language has become a minority language precisely because it is spoken by a minority population, usually one that finds itself surrounded by a majority that speaks a different language. In many cases, that majority language is the language that is used in situations that provide socio-economic opportunity: school, work, the media, health care, public services, and more generally as the lingua franca of social life. Therefore, minority language speakers usually have to be able to speak the majority language. Native speakers of the majority language, on the other hand, usually have no need to know the minority language. Given this social dynamic, the minority language will always be under threat, and the sociolinguistic literature reflects this: one of the main questions is what determines whether a language is maintained or not, and the main applied question to go with it is what a community can do to ensure that its language survives.

Maintenance and shift are usually conceptualized as involving a continuum, of which shift is the endpoint. Of course, one could also see it as a simple dichotomy: as long as the language is used, it is maintained, and when it is not used anymore, shift has occurred. However, there are many gradations of being maintained, depending on the degree to which a language is used in the various domains of life. It makes a big difference whether parents still use a language to talk to their children or only to communicate with their own parents. In Ethnolinguistic Vitality Theory (Giles, Bourhis, & Taylor, 1977), languages are given a quantitative Vitality Index score, suggesting a relatively precise indication of the degree to which they are being maintained. An example is Yagmur & Kroon (2003), who studied the minority

language Bashkir in this respect. Conceptualizing the state of the language in terms of vitality puts a positive spin on things of course; the other side of the coin is that a low vitality index means the language is vulnerable.

Irrespective of whether one looks at all this from an activist perspective or not, studies of maintenance and shift describe the struggle of minority languages to stave off extinction. Traditionally, if the language is near death, it is referred to as 'moribund', 'obsolescent' or 'relic' (Riehl, 2015), and the term 'Heritage Language' also sometimes conjures up this late stage in the process of shift. Until recently, healthier minority languages were not referred to as 'HLs', perhaps because it feels strange to refer to a cultural commodity that is still in daily functional use as an object of heritage. However, the meaning of the term has gradually been extended to cover most minority languages, the rationale being that any behavioral feature that links a group to its past serves to link it to its heritage, and that an effort to avoid shift is also an effort to protect heritage. Leaving aside the question whether the term 'heritage' is accurate, the underlying motivation for much of this book is that it is a good thing to study all minority languages within the same general descriptive framework.

### 3.3.1 Indigenous minorities

Many languages caught up in the struggle for survival are indigenous minority languages. This normally means that they are being spoken in the same place where they have always been spoken, or at least since many generations, but that this homeland is also part of a larger context, such as a nation state, in which another language plays a dominant role. The group concerned is, therefore, an indigenous minority in that larger context, though it may well be the majority population in its homeland.

In many cases, some form of post-colonialism provides an adequate socio-cultural framework within which to examine the social circumstances of indigenous minority languages, though only if we define it in a sufficiently broad way, including any configuration in which a group has been subjugated socio-politically by an invading group or by a powerful neighbor (see Chapter 2 for more information). This way, it will include African languages such as Luo (in Kenya) or Tswana (in South Africa), forced into a relationship with English because of the legacy of colonialism, as well as European languages such as Frisian (in the Netherlands), Sorbian (in Germany) and Irish. Most such groups speak a language of their own, or at least once did, and many of these indigenous languages are in the process of being edged out by the more powerful language. Many others have reached the end of that struggle and are not spoken anymore. As is typical for most language contact situations, the languages in contact tend to have an asymmetry in status.

The relationships between languages can change if the socio-political circumstances change. The most famous reversal is probably the resurgence of English after it was dominated by Anglo-Norman for a couple of centuries after the Norman Conquest of 1066. Catalan was in a subordinate position during the Franco years, but nowadays its position in Catalonia is strong. In Estonia, many people are bilingual in Estonian and Russian. This is the result of the fact that the country was part of the Soviet Union until 1991, with Russian as the *lingua franca*. Given understandable political motivations, the status of Russian changed overnight from societal dominance to the native language of part of the population and a not so popular second language for everybody else. More than twenty years later, there is a sizable Russian-speaking minority in Estonia, and many of the Russian speakers are bilingual.

### 3.3.2 Immigration

As Chapter 2 also illustrated, the other main source of bilingual settings across the world is immigration. Movement of peoples is, of course, not only something we witness in modern times. Witness, for instance, the enormous distances covered by large groups of people in relatively short time spans during the Great Migrations in the early Middle Ages, when for example the Germanic Vandals and various tribes of Gothic people moved from what is now Pomerania and Denmark to Southern Spain and Northern Africa, by way of Hungary, the Balkans, southern Ukraine, Greece and Italy. The world is awash with languages spoken at a location that is far removed from where the ancestral language is supposed to have been spoken. The Navaho, for instance, live in the US Southwest, while most other Athabascan languages are found in northern Canada. Another product of immigration is of course the Anglo-Saxon invasion of Britain, which brought speakers of what was to become English, and led to the Celtic population exchanging its language for English, as the Gauls in France had done earlier when they switched to Latin because of Roman colonialism.

As these examples show, though, it often was military conquerors that immigrated and imposed their language on the native population. Typical of immigration in modern times, however, is labor migration, a phenomenon that only really took off with the advent of the Industrial Revolution. Such immigrants tend to have a lower position on the social ladder, so there is no chance of imposing their language on the natives. In fact, labor migrants tend to shift to the majority language relatively quickly. Until they do, though, they form bilingual communities. Once shift is underway, their native language turns into an HL.

Some examples of modern immigration are very well known and have been studied extensively in the literature on bilingualism. Primarily, perhaps, this holds

for Hispanic immigration to the United States, but there is also a sizable literature on Asian immigration to the UK, mostly from the Indian subcontinent, and on immigration from Mediterranean countries into Western Europe, primarily from Turkey (to Germany, the Netherlands, Switzerland, etc.) and from Northern Africa (especially in France). However, there are numerous examples, especially if we adopt a relatively broad definition of labor migration, something along the lines of 'seeking a better life, with better economic opportunities and free from persecution'. This way, the category covers such diverse populations as Pennsylvania Germans, Japanese in Peru, Germans in Chile and Argentina, Greenlanders in Denmark, Texas Czechs, Dutch in Australia, Hmong in the US, Mozambicans in South Africa, and hundreds more. Pushing back in time, throughout history people have moved to places where economic opportunities or political freedom were greater. In the past, this often involved relatively short distances (e.g. the emergence of the Dutch colonial, maritime and trade power in the early 16th century was fueled in part by the influx of enormous numbers of religious refugees from Flanders).

A general pattern is that immigration often starts off with the idea of a temporary stay; just long enough to earn sufficient sums of money that can be sent back home as remittances, but at some point along the way the residence in the new country becomes more permanent. The first migrants, often men, are joined by their families and soon enough children are growing up in the new country, which helps the family taking root. Often there is a prolonged period in which the older generation dreams of returning and the younger generation feels at home in the place where they have grown up. These are just general patterns, of course, and the tension between wanting to stay and wanting to go back plays out in different ways for different individuals.

### 3.3.3 Social factors that affect maintenance

By and large, we refer to a situation involving maintenance if the HL continues to be used as the main language for informal in-group communication, or at least as one of the languages for this function. It will normally do so in addition to the majority language, but the modes in which the languages co-exist vary enormously. In many cases, some form of a diglossia applies, a pattern identified by Ferguson (1959) to describe the relation between two languages in a country and adopted for cases of bilingualism by Fishman (1972). Diglossia is characterized by a rather strict division of labour between two languages. One language is used by speakers in formal settings and the other language is used in informal settings, as is typical, for instance, of the division between dialects and standard languages. In HL contexts, there is often an almost exclusive use of the HL in private domains

and equally unrivaled use of the majority language in public ones. In many other cases, however, there is much more variation, and often extensive codeswitching, i.e. the use of both languages side by side in the same conversation. In some cases, languages are mixed so thoroughly that few utterances in everyday conversation are in just one language.

This reference to everyday language use indicates that there is a tight link between the higher-order question of what happens to a language in the long run and the everyday issue of what language is chosen for use in concrete communicative situations. There is a large literature in sociolinguistics about what determines whether or not a language is maintained or not; as we will see, the factors that are mentioned can usually be interpreted as an indication of how likely it is that a speaker will choose the HL or the majority language in a particular communicative exchange (for overviews, see, e.g., Clyne, 2003 and De Houwer, 2009).

Perhaps the main conclusion of many years of research on maintenance and shift is that there are so many factors that play a role, and they play it so differently in concrete cases, that it has proven impossible to construct a single widely accepted theory that accounts for maintenance and shift and that predicts what will happen in any given bilingual setting. Factors of importance include the size of the community, the duration of the contact setting, the degree to which the community is replenished with new speakers (e.g. in case of continuing immigration), the degree to which the language receives institutional support, and many others. The reasons why it is hard, perhaps impossible, to combine these factors into a reliable predictive model are that the factors interact in various ways, and it is possible for the same factor to facilitate maintenance in one setting and shift in another. Despite the empirical complexity, there are various models that combine many of these factors into a single framework, such as the model of ethnolinguistic vitality (Giles & Johnson, 1987).

### *Case study: Turkish in the Netherlands*

It is difficult to order the various factors that influence the outcome of bilingualism in a clear way that has predictive value for other or new cases of bilingualism. However, clearly there are recurrent patterns. We will start with a case study, Turkish as an immigrant language in Western Europe, and explore which aspects of this situation are typical and which are unique.

Surveying the various studies that have been done on ‘Immigrant Turkish’, Backus (2013) identifies the following factors as having been reported to determine why Turkish is maintained remarkably well over the fifty years of bilingualism between 1960 and 2010.

1. There are few exogamous marriages: until recently, more than 90% of spouses came from Turkey, rather than from within the immigrant community, let alone from non-Turkish backgrounds (Yagmur & Akinci, 2003)
2. The Turkish language is often reported to be a 'core value' for Turkish identity (Extra & Yağmur, 2010, p. 131)
3. There is a strong element of transnationalism, meaning that there is a lot of contact with relatives and friends in Turkey, nursed through visits, summer vacations, and communication through Skype, etc.
4. There is some marginalization and physical segregation, with parts of towns mostly populated by ethnic Turks. Usually this has to do with the availability of cheap housing, e.g. in the area of Berlin called Kreuzberg (Pfaff, 1994, p. 77). In such areas, Turks encounter relatively few members of the majority culture in their daily lives, limiting the need for the use of the majority language as a lingua franca. Many services, such as banks and hairdressers, are Turkish in such parts of town.
5. There is easy access to, and much use of, Turkish media, such as satellite television, newspapers, and the Internet (Sauer, 2010), reflecting an on-going desire to stay in touch with developments in Turkey.
6. Despite the limited availability of government-sponsored home language education, children often receive some form of education in standard Turkish in schools or in some type of extracurricular form.
7. There is quite an abundance of Turkish organizations, ensuring many opportunities for 'intra-group' contact.
8. Turks have their own mosques in which to worship (cf. Eversteijn, 2011).

For some of these factors, it is easy to see how they would promote or hinder maintenance in any bilingual setting. If exogamy is virtually non-existent, few people will enter the family domain whose lack of knowledge of the ethnic group language would force the use of a lingua franca. Similarly, if you're not considered a member of the ethnic group if you can't speak the language, it will feel like a duty to maintain it, especially if ethnic pride runs strong. Perhaps most importantly, if there are good reasons to maintain contact with the homeland, obviously knowing the language is of great instrumental value. In earlier times, possibilities for contact with the homeland (in case of immigrants) were limited. Travel was expensive and impractical, and communication by phone was costly enough that it tended to be limited to occasional calls. Most of the American Norwegians studied by Haugen (1953) and referred to in Chapter 2, would never see Norway again, their only contact being letters and the occasional visitor (e.g. a priest). A modern-day Norwegian immigrant to the US, on the other hand, will probably travel home for family visits nearly every year and maintain almost daily contact with family

and friends back home through social media. With the advent of cheap air travel and cheap or free means of communication through cell phone and Internet, modern immigrants often have an identity called 'transnational' (Vertovec, 2001), which signals that more than before they may feel they belong to both places. An inevitable consequence of extensive contacts with the monolingual homeland is that there will be people in one's network who are monolingual HL speakers (Ritter, 2014, p. 274), necessitating maintenance of the HL for the bilingual. This could mean that immigrant languages stand a much better chance of survival in the future than they had until the late twentieth century.

Some of the other factors mentioned above as promoting the maintenance of Turkish are less clear-cut if applied to bilingual settings in general. Physical segregation of the group, availability of media and school lessons in the ethnic language, and the existence of ethnic organizations do not necessarily promote language maintenance. They have been argued to do so in the case of Turks in Western Europe, but one can easily imagine that they could promote shift too. Segregation may go together with discrimination and the wish of parents to provide their children with a better life: not transmitting the ethnic group language may be one way to stop their being discriminated against. Media and school lessons may be unpopular or ignored, especially if the shift has progressed quite a bit and people feel anxious about their proficiency in the HL, or if there is no role for the HL in formal domains of life, which would, among other things, require literacy in the HL. There are many other factors that can be important in a particular setting, and not play a role at all in others.

### *The primacy of everyday interaction*

Generalizing, we can say that understanding the multitude of social, demographic, economic and cultural factors that influence the outcome of the battle for maintenance cannot be done without an investigation of how they affect everyday language choices of actual bilingual speakers in frequently occurring communicative settings. The most frequently occurring setting is, for most people, everyday informal speech with friends, family and coworkers. Many if not most of the factors can ultimately be related to their impact on whether speakers choose the HL in a given situation or not. How exactly they influence these choices may differ from setting to setting. If the number of domains or situations in which speakers choose the HL is high, the language is maintained. If it is low, it will die, unless the trend can be reversed.

#### 3.3.4 Investigating language choice

Several factors play a role in the choice for a particular language.



*Interlocutor effects*

Studies of language choice in minority communities show that choices are rarely clear-cut. Only when the conditions are right for a stable diglossia pattern, with fixed roles for each language that every community member accepts as the normal state of things, and with no need to contest the status quo, will it be possible to construct a clear decision tree: in Situation X, the language chosen is A; in Situation Y, one chooses B. Rubin (1962) described a situation like this for the everyday choices between Guaraní and Spanish in Paraguay (but see Gynan, 1998 and Estigarribia, 2015 for evidence that this situation does not apply anymore, if it ever did). On the other hand, while such strict, almost prescribed, choice patterns may not often be the reality on the ground, there are usually recurrent patterns that can be identified. These patterns are the basis for people's ability to adhere to the communicative norms of their community and to say something about language choice when asked about it in questionnaires. One consistent finding is that the most important factor determining which language you will use in any given communicative situation is who your interlocutor is. Other factors that exert some influence include the topic of conversation and characteristics of the physical setting. If your work involves only the majority language, it will often be easier to talk about work in that language. As for setting, if your minority is stigmatized, you may not want to speak its language in a public setting, such as a store or a city office. However, it seems that who you talk to is what determines your choices most.

The interlocutor effect is partially a matter of habit. If you talk to a person often, conversation with that person becomes a routine. Routines have fixed qualities, and in this case one of those qualities is the language in which you interact. If a family has always spoken its immigrant language at home, say Spanish in a fictional Mexican-origin family in an American city, and one of the sons brings in a girlfriend who doesn't speak Spanish, it may be difficult at first for the family members to speak to each other in English. They may want to do so out of politeness, so that the guest can understand what is being said, but it is very difficult to break the habit. Yet, this setting provides one of the easiest language choice scenarios involving an interlocutor effect: if the person being talked to speaks only one of your languages, obviously you choose that language for communicating. However, in bilingual communities, there are many communicative settings in which all interlocutors know both languages. Information on language choice in such situations is less clear-cut. For one thing, this is because the choice patterns themselves are messy, with speakers opting for codeswitching, a mixed style of speech in which both languages are used. However, when language shift is underway, this is usually visible in the language choices made with different types of interlocutors. Generally, the pattern is that children use the majority language

with each other, and only use the minority language with parents, other people of the parent generation, and in cases of advanced shift, only with the grandparents. There aren't too many community-wide studies of such choice patterns, but some examples are Gal (1979) for language choice in Hungarian-speaking peasant communities at the extreme Eastern edge of Austria (in the region called Burgenland), an indigenous minority language, and Li Wei (1994), a similar study of the Chinese immigrant community in Newcastle-upon-Tyne in the United Kingdom.

### *Generations*

A common way of documenting ongoing language shift is to compare the language use of older and younger generations in a community. A common pattern is that telltale indicators, especially language choice patterns and proficiency measurements, differ considerably across the generations, for instance showing that an HL is well maintained and in active use by the grandparent generation, but is little used, and often not mastered very well, by the children's generation. In extreme cases, these two generations cannot talk to each other very easily, grandparents not speaking the majority language well and the grandchildren not really able to sustain a conversation in the HL. However, other cases, such as Turkish in the Netherlands, are less extreme. In Table 3.1 we find roughly the following pattern (Backus, 1996; Eversteijn, 2011):

**Table 3.1** Generational shift in Turkish Dutch communities

Interacting with ...	Grandparents	Parents	Siblings	Friends	Colleagues
Generation 1, born in Turkey, came to the Netherlands as adults	T	T	T	T	T/D
Generation 1A, born in Turkey, came to the Netherlands as young children	T	T	T	T/D	D/(T)
Generation 2, born in the Netherlands as children of generation 1	T	T	T/D	T/D	D
Generation 3, born in the Netherlands as children of generations 1A or 2	T	T/D	T/D	T/D	D

Differences between generations are important for determining whether particular features of the HLs are due to ongoing change, attrition or what is often called 'imperfect acquisition'. To discuss this issue, we need to know what the language use of the older generation, which acts as input for the younger generation, looks like. We will return to this issue in Chapter 6.

A typical phenomenon in immigrant families that show a large discrepancy between language competences of children and parents is in brokering (Orellana & D'warte, 2010). This occurs when children must do a lot of translation work

for their parents and other elders. They will accompany their parents to all kinds of official business, including visits to the doctor and schools, to ease the communication between those institutions and the parents who have little knowledge of the language used in these official settings. Through experience, children often become essential mediators in this way between the community and the mainstream (cf. Garcia, 2009).

### *Family language use*

Considering the importance of the home environment for language maintenance, it is not surprising that studies of language choice pay particular attention to what goes on inside families, leading to the emerging field of 'family language policy' (cf. Deprez, 1996; Spolsky, 2009; Kopeliovich, 2010; Ritter, 2014; Schüpbach, 2009; Schwartz & Verschik, 2013, and Pauwels, 2005; all of these are of course indebted to the early language choice research referenced above). Deprez (1996) is a large-scale investigation of language choice practices in immigrant families in Paris (of Arabic, Creole and Portuguese backgrounds). Typical of the family language policy paradigm is more explicit attention for how parents attempt to control how their children speak. Lanza (1997) describes parental strategies such as refusing to understand the child if it does not speak in the HL (see also De Houwer & Bornstein, 2016; King, 2016).

While the sociolinguistic literature mostly focused on the association of languages with particular life domains and particular interlocutors, family language policy studies combine this older tradition with the interactional focus of micro-sociolinguistics, investigating actual conversations in HL speaking families. Most studies in the field adopt a qualitative or mixed methods approach, aiming to understand how the parents feel about the importance of maintenance (i.e. what their 'language ideology' is), how they try to implement their 'language policy', and to what degree interactions in the family in various settings (e.g. at the dinner table) conform to that policy. One finding is that within the same community there will usually be considerable differences as to which families aim to keep the HL alive (Schwartz & Verschik, 2013, p. 9). Some feel it is important to learn the language because otherwise the children will not really be able to claim the ethnic identity. The other extreme is the attitude that the HL is useless in the further life of the children, so maintenance of the HL is either not on the agenda at all or made secondary compared to becoming a full member of the host society.

This qualitative perspective complements the generalizing picture of the aggregated macro perspective with information on variation within a community and on the direct effect of communicative motivations in actual conversations on linguistic choices. Ritter (2014), for example, shows how first-generation Russian immigrants in Germany switch to Russian in conversations with their

children for pragmatic reasons and to German for others. The second generation has different patterns, but again, switching patterns are associated with particular pragmatic features. This links up with the literature on functional motivations for codeswitching, to be reviewed in the next chapter, which has generally turned up similar findings.

It is not always easy to work out exactly how to model the complex relationship between the larger sociolinguistic conditioning of language choice and the particularities of individual conversations, a tension that goes back to the main division in sociolinguistics, between macro and micro approaches, or between variationist and interactional sociolinguistics. In the case of language choice research, both poles are concerned with the relationship between a speaker's identity and language choice, but the empirical data and theoretical positions are very different. Macro approaches typically go for the bigger picture and identify the factors that lead to maintenance and shift, while micro approaches typically address the subtle issues involved in conversational structure. Identity is not always involved in language choice, as sometimes it is simply the better availability of words in the one language than in the other that effectuates a 'language choice'. We will return to the motivations behind language choice below.

### *Networks*

Clear though the impact of generational differences and family routines may be, issues of language choice often involve many factors at the same time and focusing only on the demographic backgrounds of speaker and interlocutor will obscure their impact. For that reason, it is useful to also look at finer-grained pictures of bilingual reality. It is much the same in sociolinguistics in general, where originally the concept of 'speech community' was used to account for the existence of large-scale linguistic distinctions, such as 'sociolects' and 'varieties'. It is indeed useful to talk about, say, 'the Dutch speech community' when making general statements about the linguistic situation in the Netherlands; but this aggregate level is too coarse when examining how people speak in one socio-cultural group (like artists) or in a particular kind of discourse context (like a staff meeting). Sociolinguistic researchers often study linguistic behavior in such smaller groups. Two kinds of social groupings have been studied in detail: social networks (Milroy & Li Wei, 1995; Lanza & Svendsen, 2007) and communities of practice. Both were developed to zoom in on specific subgroups within a speech community. Roughly, social network studies have focused on language use particularly in small and tight-knit groups that have certain permanence, while communities of practice are more ephemeral sometimes and, in any case, pertain only to part of someone's life. A community of practice, that is, is a group that emerges out of a joint activity, such as a job or a hobby. A person is normally a member of numerous communities

of practice, each of them linked to one aspect of his or her life. Because of the increasing compartmentalization of our lives, communities of practice typically represent a modern phenomenon.

Language choice questionnaires are typical data used in social network analysis. If you can draw up a complete picture of someone's social network, and you can link each node with a favored language (i.e. the language used in communication with the person connecting with that node), then you have a complete picture of that person's language choice. If you have these data for a representative sample of speakers, and preferably from different age groups as well, it should be clear in what parts of the network the HL remains strongest. What social network analysis makes possible above and beyond simple conclusions about a language's vitality is with what kind of network ties a language is maintained best, or worst. It may make a difference, for example, whether a network is dense and multiplex: if most of your contacts also interact with each other, and you see these people in many different contexts (at home, at work, on the football field, etc.), chances are that the language choice for the interactions within this network is relatively stable. Once a person has a social network in which there are few nodes that favor the HL, however, it is hard to stave off language death. It will leave scattered individuals who may know the language, but since they have no means to use it on a sustained basis, it is impossible to build up the usage-based level of comfort with the language to be able to use it fluently.

### *Communities of practice*

The study of language use in communities of practice has had considerable impact on sociolinguistics in the past two decades, as it has shown that speakers adapt their language use to their interlocutors and to the setting in which the communication takes place (Eckert, 2008). This is not the place to review all that work, but its take-home message is that questionnaire data about language choice may miss the subtle differences in behavior between different parts of one's life.

One study in which this has been analyzed in detail is Ayduk et al. (2000), who documented a range of stylistic choices, including the language of communication, in two groups of Latino gangs in a Northern California town. All participants in the study spoke Spanish and English, though with different degrees of fluency. The degree to which they used either language was to a large extent determined by the culture of the gang to which they belonged. Simplifying, one gang was associated with Spanish, the other one with English. Whether in any conversation a particular gang member used English or Spanish was at least partially dependent on the degree to which gang membership needed to be symbolized. In HL studies the concept community of practice has been used to explain bilingual language use patterns in the Japanese community in Australia (Oriyama, 2017). Oriyama argues

that Japanese is particularly used in community-based Japanese language schools or classes. She suggests that such different schooling options are likely to have differential impacts on the formation of students' perception of their identities.

It is clear that if HL use is restricted functionally to certain domains, it is also restricted formally to the registers typical for those domains. As a result, proficiency in those registers may diminish, despite the fact that overall vitality if the language may remain high. In our work on Immigrant Turkish, we have documented that the immigrant population tends to form its subordinate clauses in a way that betrays influence from Dutch. Most Turkish subordinate clauses are structurally very different from Dutch ones, as they are non-finite, but some subordinate constructions, used relatively sparingly, do resemble the canonical Dutch finite structure more. Several studies (Schaufeli, 1991; Treffers-Daller, 2011; Yilmaz, 2013, and especially Onar Valk, 2015), have found that in the immigrant community, use of these formerly rare, but Dutch-like structures has shot up, edging out the inherited canonical Turkish structures. Further details will be given in Chapter 4; here, we just wish to note one aspect of this phenomenon, germane to the discussion in this section.

As was noted at the beginning of the chapter, Turks in Holland tend to not use Turkish in any of the domains associated with formal language use (school, work, public administration). Thus, they don't use formal registers of Turkish very much or at all. Subordination is more frequent in formal registers. Non-use of formal registers seems to correlate with non-use of Turkish subordination structures. Instead, when a subordinate clause needs to be formed, it is formed the Dutch way, since Dutch is the language used in formal domains, and hence Dutch subordination structures are well entrenched in their grammatical competence. (See Rothman, 2007; Kupisch & Rothman, 2018 for differences in language use if HL is also a school language and for effects of register in general and hypothesis of MICD (missing input competence divergence)).

### 3.3.5 Ways of influencing language choice

States can implement language policies to facilitate or hinder maintenance of minority languages, but it is important to realize that it is what speakers do that has the more immediate impact on the fate of the HL. This is not to deny, of course, that policies can influence what people do; we just want to emphasize that whether maintenance or shift results depends on people's actions. Clyne (2003, p. 68) refers to this speaker prerogative as the 'right to maintain and the right to shift'. However, the activist stance is never far away in the study of language shift, because ultimately most sociolinguists sees shift as a bad thing, since it confirms the inequality in status. There are models for how to turn state-of-the-art information on the

vitality of a HL into policy proposals (Grin, 2003; Lo Bianco, 2008). Often, such proposals aim to raise the status of the language, which in turn would induce parents to transmit the HL to their children. Efforts to reverse language shift (Fishman, 1991, 2001b) may be aimed at many things, but crucial is the phase in which families and neighborhoods are encouraged to start transmitting the HL again to the youngest generation. Community activists often focus on transmission. Claiming a place for the language in the educational system, for example, is one way in which the status of the language can be raised, and perceived usefulness is important in nudging parents towards using the language with their children. See further discussion in Chapter 11.

### 3.4 Shift

If the social factors reviewed above are stacked against a minority language, speakers will often opt for the majority language in their everyday choices. Language shift is technically the result of speakers choosing not to speak the language anymore, usually in the form of not transmitting it to their children. This is usually done out of a negative attitude towards the language, and this negative attitude usually has its roots in the low social position of the group. Often, speakers blame their language for holding them back, and, in many cases, they may well be right, in the sense that not speaking the majority language smoothly and without an accent limits your socio-economic opportunities.

#### 3.4.1 Shift and acculturation

Shift is particularly prevalent in those sectors of society that face limited opportunities in other ways as well. Canagarajah (2008) shows there is an association between low level of education and negative views on Tamil as a minority language, as spoken by immigrant communities in the UK and North America. In the most blatantly racist circumstances, children have been forced to grow up part or most of the time away from their families, so that culturally and linguistically they can be acculturated into the mainstream. This has had disastrous consequences for quite a few Aboriginal languages such as Kayardild in Australia (Evans, 2010), or Native American languages in the Americas.

At the same time, the literature on minority languages shows that languages can survive despite being dominated by another, more powerful language, and the outcome is not preordained. Languages can survive indefinitely in a subordinate role, as in diglossia, and social relations can even be reversed so that the language gains in prestige again. However, once a language lacks prestige vis-à-vis

a mighty second language, drastically limited input in the HL may eventually lead to its demise.

While language maintenance and shift are usually seen as self-contained empirical topics, they can profitably be seen as part of the larger phenomenon of cultural maintenance and shift. Acculturation is the outcome of the battle between the adoption of host culture norms, values, beliefs and behaviors and the retention of another set of these things from the home culture. Of course, language has its own dynamics, which justify a separate to some extent. Language shift, for example, is completed when the language is dead, while it is unclear whether one can ever say that a culture has gone. In any case, cultural maintenance often survives language shift, though that doesn't mean the original culture is maintained completely.

### 3.4.2 When shift reaches its endpoint

Undeniably, languages may die. The outcome of the struggle is all too often that the minority language disappears at some point. Many languages that used to be spoken are now dead, and almost every year the media will pick up on the passing away of the last native speaker of a language.

Nancy Dorian was among the first to describe in detail what happens to a language when it is what she calls 'dying', in her case a variety of Scottish Gaelic spoken in three isolated fishing villages on the Scottish east coast, far away from other Gaelic-speaking communities. In such cases, the language is not really spoken by anybody anymore as an ordinary vehicle of everyday communication, save perhaps for a few elderly people. The term 'heritage' in these cases is associated with a past that is irretrievably gone.

East Sutherland is in the Scottish Highlands, north of Inverness. Until the Clearances of the 19th century, most of the people living there spoke Scottish Gaelic as their mother tongue, and many will have been more or less monolingual. The Clearances changed all that, mostly by forcibly removing those people, but one side effect was that a few small fishing villages on the East Coast saw their population increase considerably, creating Gaelic speaking centers far from the main Gaelic speaking parts of Scotland on the western coast and on the Hebrides.

The combined forces of isolation from other Gaelic speakers and increasing integration into the English-speaking world around the villages, basically the forces of modernization, led to a fairly rapid decrease of Gaelic usage in the mid-20th century. Dorian's work, based in fieldwork done in the nineteen seventies, describes this process in detail, and focuses both on the sociolinguistic pressures that stimulated the use of English and discouraged the use of Gaelic, and on the grammatical and lexical implications this had on Gaelic as it was dying. It introduced the concept of the 'semi-speaker' (see Chapter 1): people who knew Gaelic



to some degree but were not really fluent. They had learned the language in childhood but had never really used it as their main language for everyday communication. Actually, Dorian distinguished between various gradations of semi-speaker status. The diagnostics of a semi-speaker were essentially non-use of the language, or at least extremely limited use of it, and a host of grammatical deviations from the grammar of a 'full speaker', most of the time simplifications. The concept of 'semi-speaker' and its definition are remarkably similar to what nowadays is called 'heritage speaker'. Together with Annette Schmidt's (1985) description of similar changes in the Dyirbal speech community in Australia, Dorian's work stands as one of the major early landmarks in the study of language death.

However, one or two generations before the situation described above, the East Sutherland Gaelic community had been using Gaelic as its main language. At that point, the language must not have looked like a piece of heritage culture at all. Yet, English was present in the community, and people were bilingual. The situation described by Dorian is the outcome of the process of language shift as it evolved in the ensuing years. As more and more people spoke more and more English, and started raising their kids primarily in English, usage and command of Gaelic went down, and gradually it turned from a vital vehicle of communication to a relic of the past, known by some because they are old or because they are particularly interested in local culture and history, just like they may be interested in the history of the local fishing industry, as just another item of heritage.

There is no transitional moment, however, and for this reason it is useful to treat what is often referred to as 'minority languages', especially in the literature on maintenance and shift, and as 'HLLs' in much of the modern literature on the linguistic impact of multilingualism on ethnic group languages, as more or less synonymous concepts. In all cases, the language of interest is spoken by a community that has its roots in immigration or encroachment by a powerful neighbor or invader, and that is in some position of vulnerability due to lesser social status. Given the prominence of the term 'minority language' in the literature on maintenance and shift, we have retained it in the current chapter, and in general when referring to literature from this domain.

This way it becomes possible to carry out systematic comparisons between languages caught up in different stages of the continuum between life and death. It is to be understood, though, that death is not inevitable. While the demise of Welsh, for example, has often been predicted, the language is still spoken widely. Basque has more speakers now than it had several decades ago. Similarly, while immigrant languages often cease to be spoken after three generations, Turkish seems to be maintained quite well by the immigrant populations in Western Europe. Situations involving language maintenance are found across a huge range

of specific circumstances, with wildly differing degrees of how well the language involved is maintained.

Shift situations produce varieties of the ancestral language that show signs of wear and tear. Parts of the vocabulary are not known anymore, the grammar is simplified, whole discourse registers fall into disuse, and in all domains there is contact-induced change because the other language influences the HL. These changes are usually seen in a negative light, as language loss, though a more neutral perspective, that they are signs simply of language change, is also possible. In addition, what is lost will often be replaced by something that is added, usually in the form of borrowing from the other language. Something is said to be lost from a language if older speakers have it, or if previous generations had it, while younger speakers don't. Since immigrants come from elsewhere, many immigrant languages have a non-immigrant counterpart variety which can then be used as a baseline for comparison. Loss within a speaker, i.e. if someone had a feature and now has lost it, is referred as 'language attrition'. Contact-induced change, loss and borrowing, and attrition will be discussed in Chapter 6.

### 3.4.3 Power versus solidarity

To explain maintenance and shift, we need to look at the economic, cultural and political circumstances in which a bilingual community resides. The power dynamics shaped by these factors generally decide about the fate of a language. For immigrants, one aspect of life they need to master, and a project they often eagerly embark on, is learning the language of the society they have moved to and making sure that their children learn that language well. If shift is the end result, that language of the wider society is of course the language that the immigrant community shifts to. Turkish immigrants in Holland, for example, learn Dutch. Similarly, and more insidiously, indigenous minority groups often face the pressure of a more dominant majority and have equally good reasons to know the language of that majority, despite the fact that such minorities are in their own country, where they have resided for centuries or longer, and in a way have had multilingualism imposed on them. Typical cases include languages such as Quechua, Nahuatl, Xhosa, Dyrirbal, Yakut, Breton, Frisian and Sorbian. All of these are languages that were once used in all domains of life but have come to face stiff competition from a more dominant language, in most cases the language of people that once came to their territory as invaders or conquerors. There is a certain political injustice in that, but essentially the situation is the same as for immigrants: in order to get ahead in life, it is good, perhaps essential, to know the majority language. It is the language of power.

All-important in these developments are language attitudes. If speakers are afraid to speak the language, its decline is hard to stop. After all, a language has ceased to exist once nobody speaks it anymore. It is for this reason that stigmatization is very damaging to the survival chances of a language. The negative attitudes that come with stigmatization from the outside are easily internalized by the speakers of the stigmatized language, and this may well make them reluctant to use the language at all. They don't want to be noticed publicly speaking a particular language if others will frown upon it, and they will try to keep their children from having to experience the same humiliation.

In theory, of course it is perfectly possible for any group to just add that societally important language to their repertoire and maintain the ancestral language regardless. Many groups have done so, and this leads to active bilingualism, sometimes with both languages allocated to relatively stable domains of usage. However, the basic asymmetry in power between the languages in question often causes minority populations to feel their ancestral language is only a hindrance, and in any case serves little practical use. As a result, they may abandon the language, and language shift results.

One may wonder, if the pressures to use the more dominant language are so intense, why any minority would ever maintain its group language. The reasons are twofold. First, the intensity of the pressure varies across groups and across time and is not always so great and so insidious that socio-economic, or even physical, harm would come from not giving in to it. In such cases, the drive to change one's habits is not so strong. Second, often the minority language serves the important function of marking solidarity within the minority group. It is a marker of heritage, a badge of one's identity. By choosing the minority language when one also could have chosen the majority language, one is effectively conveying the message 'you and me, we both belong to this group'.

That is not to say that if identification with the language is strong enough this will insure language maintenance. There is always the countervailing force of the attraction of the other language. Also, limited proficiency in the HL may get in the way: speakers of an HL may not want to put up with the feeling of inadequacy that perhaps inevitably comes with not speaking the language of your own ethnic group well enough.

The issues discussed in this chapter go some way to explaining how a formerly healthy language, used in all domains of life, becomes an HL with domain restrictions, and often attracting negative attitudes from its speakers.

### 3.4.4 Language shift and ethnolects

Research on sociolinguistic variation has demonstrated that language use is often determined by background characteristics, and broad features such as social class, gender and age have often been identified as exerting this kind of influence. Another feature that plays this role sometimes is ethnicity. Strictly speaking, this is outside the scope of this book, since such studies are about ethnically determined variation in the majority language, in the language an ethnic group has switched to. In many of these cases, the HL will be long gone. However, ethnic varieties, or 'ethnolects', sometimes arise as a by-product of HL loss. In such cases, the shift stops just short of total linguistic assimilation as also shown in Chapter 2. Textbook cases are Irish English and Andean Spanish, language varieties that arose out of a colonial situation. These two examples also illustrate the life cycle involved: Irish English arose more than a century ago, and by now few speakers of Irish English are themselves involved in a language shift from Irish to English. It certainly arose because of language shift, though, and as Irish speakers were adopting English as their main language, they transferred grammatical features of Irish to English, perhaps as a result of ordinary L1 to L2 transfer, such as is often observed in especially early stages of SLA. At the same time, it must have started serving as a badge of Irish identity at some point, since otherwise it is not easy to understand why the ensuing century didn't witness a growing approximation towards British English. Andean Spanish, on the other hand, is still in development, as it is the by-product of a currently on-going shift from Quechua and Aymara to Spanish.

### 3.5 When language choice is not clear-cut

Codeswitching complicates the picture of maintenance and shift, since that picture presupposes the reality of separate languages. If language choice gradually shifts so that in fewer and fewer communicative situations the HL is chosen, it is a simple question of numbers. But in empirical reality, the drama of shift plays out in an arena in which language choice often involves the use of both languages at the same time. In many communities, abundant codeswitching is the normal way of speaking.

There have been many studies of codeswitching in a great number of bilingual settings (Poplack, 1980; Myers-Scotton, 1993). We know the phenomenon is extremely widespread, and yet we know very little about the degree to which it is representative for how bilinguals communicate. This is ultimately because people do not find it easy to answer questions about the extent to which they code-switch, other than saying that they 'use both languages mixed together' or

‘avoid mixing the languages.’ Because of this, data collection must proceed through more time-intensive techniques, and for the most part this has been the recording of spontaneous conversations in which bilinguals use both languages. Collectively, the many studies that have used this data gathering technique have uncovered all kinds of interesting things about how languages get mixed, but because it is such a time-intensive way of collecting data, researchers have focused on catching people on tape of whom they knew that they would code-switch. In some cases, other manipulations were made in the design of the study to maximize the extent of codeswitching that would be obtained. In any case, it would be too expensive simply to first select a representative sample of random people in the bilingual community, and then record enough samples of their language use in a sufficiently representative range of communicative situations with a representative sample of interlocutors they regularly interact with. Yet, this is what would be needed to answer the simple question: how often is the language choice in a given situation not the minority language and not the majority language, but a combination of both?

From the evidence accumulated we can conclude though that at least some people in most bilingual communities code-switch. Usually, if they practice this kind of speech, they will do it in informal in-group conversation. If the HL is still the language most often used in the community, codeswitching tends to be asymmetrical, the HL clearly functioning as a kind of base language. If the majority language is known at least as well, codeswitching also tends to become more symmetrical, with a larger role for both languages. Once languages alternate frequently, it becomes hard to answer language choice questions with A or B. The fragment below, featuring two speakers Ş and E, comes from unpublished data by Derya Demirçay, and it exemplifies mixed speech by Dutch Turks: the italicized bits are Dutch. The fragment is relatively typical of the data, and it is neither in Dutch, nor in Turkish. It is in both.

Ş: *Sarma is moeilijk hoor kolay değil pişirmesi ney.*

‘Stuffed vine leaves **is really difficult**, they are not easy to make and so on.’

E: *Annem gewoon ateşe koyuyo pişiyö öyle.*

‘My mom **just** cooks it over the fire like that.’

Ş: *Ja mijn moeder doet echt ehm mijn moeder doet dat verschilt per pan wat zij gebruikt.*

‘Yes, my mother does it really, eh my mother does it, depends on the pan she uses.’

E: *Ow ja wij doen het in.*

‘Oh yes, we do it in.’

Ş: *Of düdüklüde yapıyo of gewoon pan. Annem doet meestal met düdüklü, omdat ze dat lekker vindt lekkerder vindt.*

‘She cooks it **either** in the pressure cooker **or in a regular pan**. My mother **does it generally with the** pressure cooker, *because she likes it that way, thinks it’s tastier.*’

Not all communities exhibit this kind of sheer unbridled mixing, and not everybody in a community that exhibits it does. To stay with Immigrant Turkish communities for a bit longer, Keim and Cindark (2002) compared three different networks of young second generation Turkish immigrants in the German city of Mannheim, and found that while one group code-switched profusely, more or less along the lines of the fragment above, other groups didn’t engage in any mixing. In one case this concerned a network of people who placed pride in both their Turkish heritage and their German identity: for them, the best way of honoring their Turkish heritage and their German identity is to speak both languages as best as they could. In their ideology, speaking both languages well means keeping each language free of influence from the other language.

### 3.6 Summary

This chapter explored a number of social aspects of HL. We hope to have demonstrated the importance of attending to both social and linguistic factors when studying language contact. The broad categories that need to be distinguished are language maintenance and language shift, which each have important consequences for language contact phenomena in HLs. The most eye-catching aspect of language use in many bilingual communities is when people do not choose to use just one of the languages in a communicative setting but use both, in the communicative mode called codeswitching.



# Bilingual language use

## 4.1 Introduction

The previous chapter introduced the social factors that determine the status of a language as an HL. These factors set in motion all kinds of developments in how the language is used and what it looks like lexically and structurally. Pronunciation and intonation may change, new words appear while others disappear, words get combined in new ways, the grammatical structure starts to show deviations from what the formerly ‘non-heritage’ incarnation of the same language looked like, and communicative styles may increasingly look different from how the language was used in the past. The current chapter will look at how and when people use their HLs and the dominant language together or apart in everyday life. Central questions include the following: how do the HL and the dominant language of the country take different positions in the lives of heritage speakers? What happens when you can understand and speak both languages? How do you decide when to use which language, and when do you mix them? Which mixing phenomena are typical, how do they affect the HL, and how do particular social situations lead to particular linguistic outcomes?

Section 4.2 will briefly introduce various aspects of bilingual language use. The primary focus will be on codeswitching. The challenge is to explain why codeswitching exists in the first place, what forms it takes, and what factors determine this. Section 4.3 focuses on how bilingualism affects long-term developments in the HL. It reviews the social and pragmatic functions of codeswitching. Section 4.4 reconnects the characteristics of codeswitching and language change with the social context in which HLs function as described in the previous chapter. To what extent do ideologies on language use determine what elements of language are and are not switched? It also dips into the concept of ‘*linguaging*’, the notion that the way we use language carries social meaning, as it shows how we wish to portray ourselves: language choice is part of this identity creation.

The take-home message will be that bilingual speakers take from each language what they need in order to convey the message they want to share, but that ideological tension almost inevitably creeps in at some point because associations



with each language are seldom completely neutral, and communication is always sensitive to those associations.



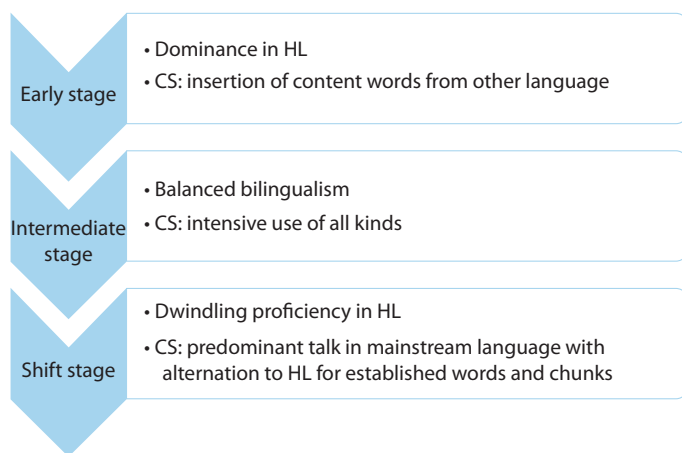
#### Main goals of this chapter:

- To understand why bilinguals mix languages.
- To understand how language mixing evolves.
- To understand how people respond to language mixing.

## 4.2 Codeswitching and borrowing

When you start tackling the literature on codeswitching, it will not be long until you encounter references to the difference between codeswitching and borrowing, and the difficulties involved in making the distinction (see e.g. the discussion in Poplack, 1980 on this issue and the summary treatment in Poplack, 2017). It is perhaps helpful to think of the two phenomena as different sides of the same coin: codeswitching is the mixing of elements (words and structures) from two (or more) languages at any one point in time, during an act of communication (the synchronic side of the coin); borrowing is the gradual acceptance of the use of such foreign words and structures as normal conventional elements of the language (this would be the diachronic side of the coin). For example, when speakers of any language in the world start using the English word *e-reader* in their everyday discourse, this can be described as an act of synchronic codeswitching. At the same time, however, every usage of the word contributes to its gradual diachronic incorporation as a conventional element, a loanword, within the language's stock of vocabulary items. This section will discuss and illustrate this distinction with a focus on the characteristics of bilingual speech.

#### 4.2.1 Early stages: Just foreign content words



**Figure 4.1** Stages in the status and use of a language in a minority community

The life course of an HL glides seamlessly from a pre-contact stage in which it is a healthy language used for everyday discourse to a language that is not well known anymore by its speakers and is certainly not used much. At the end of this cycle, most HL speakers employ the societally dominant language as their main means of communication, also amongst each other. By then, the HL will, for many of its speakers, not be the psycholinguistically dominant language anymore either. This is a gradual development that speakers might be hardly aware of as it is happening, but for the sake of presentation we can break it down into stages, for instance the three presented in Figure 4.1. This is an abstract idealization because not every HL goes through exactly this development, and more subtle distinctions are possible. However, to some extent the development holds for most HLs that have been studied to date.

Probably the most striking and frequent phenomenon is that HLs tend to incorporate vocabulary from the other language. As soon as a speech community becomes bilingual, i.e. long before the community language in question starts changing its structural make up, words from the other language start to appear in everyday discourse, in the phenomenon known as ‘codeswitching’. As we will see shortly, there are many types of codeswitching, but its beginnings usually just involve the simple insertion of content words. Myers-Scotton (1993) analyzes insertional codeswitching from the perspective of the Matrix Language Frame model. This model has been further developed in many subsequent publications. At this point, the language being spoken is still clearly the HL: most of the words, all of the grammar and all of the functional elements (articles, case markers, plural

affixes, tense inflection, etc.) are from that language: the only foreign presence is the occasional content word. Most of the time, these words are almost inescapable as they refer to concepts only available in the other language. As anyone who has spent time in a foreign country knows, all languages have seemingly unique words that are hard to translate, precisely because they refer to concepts that are more or less unknown in your own culture. As a result, your language does not have a word for it. If then you need to refer to that concept, the easiest thing to do is to just use that foreign word. The result is the simplest type of CS: the insertion of a single foreign content word. Many such words end up as established loanwords, and, in that case, they are known as cultural borrowings.

Every language has such cultural borrowings, and many of them have spread across the world. Many of the global food staples go by names that betray where they originally came from: *tomato* and *chocolate* both come from Nahuatl words, for example. The same holds for geographical features, such as Norwegian *fjord*. The most obvious modern equivalent is the rapidly spreading English-origin vocabulary of digitalization-related phenomena, such as *spreadsheet*, *computer* and *download*. However, it is important to understand that this is just the latest incarnation of a very common phenomenon: the joint spread of cultural phenomena and objects and the words that denote them. Few Dutch speakers are aware of the Latin origin of the arch-Dutch word *kaas* ‘cheese’, but it is the Romans who brought the technology of cheese-making to the Germanic peoples who spoke a language that would later develop into Dutch. They did the same with some other concepts named by words we would now consider textbook examples of basic vocabulary, such as ‘street’ and ‘kitchen’. Not all such cultural loanwords survive the passage of time. Around the time the soccer World Cup was held in South Africa in 2010, many languages took in the word *vuvuzela*, to denote the horn that fans brought with them to the stadium to make noise. But at the time of writing this book, the word seems to have more or less dropped out of use and is presumably already fading from memory.

Clearly, the initial inroads from the other language arise out of lexical need. As Chapter 3 has illustrated, people become bilingual because they experience a need for two languages in everyday life. Even if each language has its own domains of usage, it is inevitable that one needs to talk about things in one language that happened in the parts of one’s life experienced through the other language. Just as inevitably, this involves the need to refer to concepts that only the other language has words for, and this lexical need leads to the use of the relatively predictable foreign words in insertional CS, and quickly these words become relatively established in the HL as cultural loanwords.

The following box contains some of the Dutch words that were encountered in the speech of first-generation Turkish immigrants in The Netherlands, who belong

to one of two groups. One is the group of early labor migrants, known at the time as ‘guest workers’ (from the German *Gastarbeiter*), who moved from Turkey to Western Europe in the nineteen sixties to take factory jobs and, such was the plan, to return home again after ten years or so. They did not have much opportunity, or reason, to learn Dutch, but certain Dutch words were nevertheless needed, as they were the unique labels for relevant Dutch concepts. The other group consists of the spouses of resident Dutch Turks; these arrived in The Netherlands after the wedding, i.e. as adults. Since they come to Holland to raise a family, and presumably to live there for a long, long time, they tend to pick up Dutch much quicker than the old labor migrant generation did. The Dutch words appearing in their Turkish are not as restricted to the obvious lexical gap-filling type, but in general these HL-dominant speakers mainly import words from Dutch that *add* something to their Turkish. The box below contains some iconic examples.

*Ziekenfonds* – During the seventies and eighties, Dutch health insurance was organized in such a way that the unemployed and people with a relatively low income were eligible for free health care. Obviously, this was important for the labor migrants with their unskilled and sometimes dangerous labor. The insurance that was offered to these classes of people went by the name of *ziekenfonds* (literally ‘sick people’s fund’). The word was one of the first Dutch words to be attested as an insertional code-switch in Turkish data (Boeschoten & Verhoeven, 1985)

*Terras* – A culturally significant place in Dutch public life is the outdoor section of a café or restaurant. It is where, weather permitting, friends gather for a drink (whether coffee or alcohol) to chat and relax in the sun. The general word for these outdoor sections is ‘terras’ (terrace), and it features in reports of one’s day, as in ‘*guess who I met at the terras?*’ or ‘*we went to the terras after work*’. In the codeswitching data of Turkish immigrants who had just arrived in the country a couple years earlier as newlyweds, we encountered this word a lot, inside Turkish clauses. The speakers had obviously grasped that the word stood for a cultural experience for which there was no adequate Turkish equivalent, at least not one that was precise enough.

*Hemelvaart* – The first-generation immigrants who arrive as adults upon getting married to a Dutch Turk participate in Dutch public life much more than the original generation of labor migrants. As young parents, they are involved in their children’s school careers, and as a result, Dutch school vocabulary enters their Turkish. The word *Hemelvaart*, ‘Ascension Day’, carries considerable cultural significance, not because the religious background of the term is very significant but because it’s a holiday. As Ascension Day is always on a Thursday, it’s always the start of a four-day weekend sometime in May. As such, it shows up prominently in the school calendar. We found insertion of the word in Turkish utterances quite often.

The insertion of cultural loanwords has the effect of adding words with very specific and culturally significant meaning to the lexicon (while certain native words may start to disappear for lack of usefulness, for example agricultural words if the

labor migrants came from rural areas). However, it doesn't affect the grammatical structure, such as basic word order or the way verbs get inflected for tense or person, nor does it affect the stock of basic vocabulary, such as words for common objects, for family members, or for basic actions (e.g. doing, giving, or walking). At this stage, proficiency in the other language tends to be low, which means the HL is still used for pretty much all domains of life, which in turn means that speakers keep practicing the HL every day of their lives, constantly reinforcing their proficiency. However, with the more complex types of CS that occur later, cross-linguistic impact becomes more dramatic, and it is at this stage that the HL truly starts turning into a 'HL' in the Unesco definition of language under threat. This is the second stage of Figure 4.1.

#### 4.2.2 Intermediate stages: Diversified codeswitching patterns

With bilingualism becoming more intense, the types of CS start to diversify. More and more content words get inserted, and they are no longer just those words that convey cultural concepts that were unknown to the HL speakers before contact. They also become more complicated structurally, as what gets inserted into the HL grammatical frame is often more than just a simple content word, for instance a short expression (such as 'that's cool') or the combination of an adjective and a noun (such as 'public swimming pool'). In addition, other forms of codeswitching appear, including the use of discourse markers such as 'you know' from the other language, and alternation between complete utterances in each language.

##### *Continued use of inserted content words*

As a direct continuation of the early contact effects, the amount of insertional codeswitching tends to increase as the type of bilingualism starts to shift towards greater balance between the languages, itself a direct result of the increased use of the other language. Insertion is when a chunk of one language is 'inserted' into a clause from the other language. As speakers have a greater vocabulary in that language, there are now many more words that they may want to use while speaking the HL. Fewer of these fill lexical gaps in the narrow sense. Foreign words may also be preferred only because they convey the intended meaning better than their HL equivalent, perhaps because they carry different connotations. For example, some foreign words may just be activated quicker than their native equivalents, because the speaker uses them more often overall. This would hold for words that belong to semantic domains normally spoken about in the other language. In yet other cases, foreign words may get used simply because the speaker enjoys codeswitching. At this stage, therefore, the use of words and expressions from the other language can be explained by a number of factors other than semantic specificity, including

belonging to topics that favor the other language, the speaker's good proficiency in the other language, as well as considerable frequency of use of that language in daily life.

Because of this complex causality, it is often hard or impossible to provide a clear source of motivation exactly why a word from the other language was inserted. You can try and do it for the Dutch words in the following two Turkish utterances as shown in (1) and (2). Most likely, you will be able to come up with plausible explanations, but without further information you will not have the possibility to prove that your account is correct.

- (1) *Coğ-u student-ler böyle ya, ooh, mesela bu sene afstuderen*  
 many-POSS.3SG student-PL such INT oh for.example this year graduate  
*yap-ar-sa iyi baan-ı var*  
 do-AOR-COND.3SG good job-POSS there.is  
 '[for] most of the **students**, it's like this, you know, if they **graduate** this year,  
 for example, they'll have a good **job**'
- (2) *iyi o zaman adam-dan roddelen yap-ar*  
 good then man-ABL gossip do-AOR-3SG  
 'there's a lot of good **gossip** about him then'

One might argue, for instance, that the words *student* and *afstuderen* in (1) belong to a domain, studying at college, that for these speakers was a 'Dutch' semantic domain: they were living in Holland, were thinking about their education and about college, knew people who were going to college, talked often about student life, and almost all of this will have been in Dutch. Small wonder, then, that the words involved would be well entrenched in their mental lexicon. It is perfectly possible that the speakers felt that the Turkish equivalents of these Dutch words did not cover the same semantic load exactly, since their meaning has to be interpreted against the background of the whole Dutch college system. A similar story holds for *baan* 'job': the speakers were entering the job market at the time and the Dutch word for 'job' will have been used often, both by them and to them, in their everyday life. Finding an explanation for the use of the verb *roddelen* 'to gossip' is harder: any attempt to link it to the speakers' everyday life would eventually have to argue that the concept is referred to more commonly in Dutch than in Turkish, a hypothesis that is impossible to prove without very specific information about that everyday life. Ultimately, what all explanations come down to is the argument that frequent use of a Dutch word in the speaker's everyday life will increase the chances of that word ending up as an insertional code-switch in the speaker's Turkish. Frequency of use itself is of course related to semantic usefulness.

Our Turkish-Dutch data are full of insertional code-switches (Backus, 1996; Eversteijn, 2011, and Demirçay, 2017 are dissertation-length reports on three different datasets). Most Turkish utterances have one or more Dutch words in them, and most of those do not obviously fill a lexical gap in Turkish. For most, all we can say is that they are not the most basic of words, and that quite a few, but not all by far, belong to semantic domains that are associated with parts of life in which Dutch is spoken (school, stores, public domain, etc.). Similar considerations hold for any other corpus of bilingual speech, from any other place in the world (see for example Zabrodska, 2009 on Estonian nouns appearing in the Russian utterances of Russian speakers living in Estonia). It is a challenge for future research to arrive at more systematic analyses than the somewhat ad-hoc explanations given above. Published explanations of individual loanword selection often have a plausible ring to them, but it is not so obvious what would falsify such analyses, let alone how we could predict which words are going to be code-switched and borrowed.

When a language becomes an HL, it is a natural development that its lexicon takes in many words from the socially dominant language. The ultimate reason for all this borrowing is that the speakers experience a large part of their life through the majority language, and this has its effects on the degree to which words from that other language get entrenched into the speakers' mental lexicon. As a result of higher entrenchment, these words are accessed easier in speech, and will often beat their HL equivalent to the line, as it were. This shows up as the kind of codeswitching illustrated above. Keeping those words out requires some sort of purist effort, and as we will see, some communities do attempt this.

### *Insertion of chunks*

In addition to a higher quantity of foreign insertions, the insertions also start to be of a different nature at this stage of bilingualism and HL development. They are often longer chunks, such as compound nouns, combinations of an adjective and a noun or of a verb and an object noun, or full prepositional or adverbial phrases. As such, they may contain functional morphemes from the source language, such as plural suffixes or infinitive markers. We will illustrate this again with examples from the same Turkish-Dutch data we use throughout this section, but similar examples have been documented for most language pairs for which data come from people whose bilingualism can be considered more or less balanced.

In the first example the adjective and noun are both Dutch. Together they form a conventional collocation in Dutch. Therefore, it is likely that the speaker inserted them together into the postpositional phrase anchored with the Turkish postposition *-nan* 'with'.

- (3) *ondan sonra lauw water-nan yika-yınca ...*  
 then after lukewarm water-with wash-while  
 ‘and then, when you wash it with **lukewarm water ...**’

The next example shows an inserted verb-object collocation: the Dutch collocation for ‘know a language’ combines the noun *taal* ‘language’ and the verb *beheersen* ‘to master’/ ‘to know’. Note that it is just the co-occurrence of two content words in the collocation that is preserved; all grammatical inflection is in Turkish.

- (4) *bir sürü taal-lar-ı beheersen yap-ıyor-ken* (Backus, 1992)  
 one range language-PL-ACC know do-PROG.3SG-while  
 ‘while he **knows** a lot of **languages**’

Finally, Example (5) features three Dutch elements: the inserted word *vriendin* ‘girlfriend’, the inserted prepositional phrase *van mijn begeleider* ‘of my supervisor’, and a full Dutch clause with which the utterance ends. The latter is an example of alternation, to which we will turn in the next subsection.

- (5) *bir sene beraber tanış-ıyor-du-k ya, o-nun vriendin-i van mijn*  
 one year together meet-PROG-PST-1P INT 3SG-GEN friend-POSS of my  
*begeleider, die had een vriend in Turkije.*  
 supervisor DEM had a boyfriend in Turkey  
 ‘One year they met each other, this **friend of my supervisor, she had a**  
**boyfriend in Turkey.**’

The final example contains an inserted adverbial phrase. Note that it contains quite a bit of Dutch morphology and syntax: the Dutch negation, the past participial morphology in *opgemaakt* ‘made up’ (a circumfix *ge-V-t*), and the word order also follows a Dutch pattern. This kind of insertion stretches the limits of what we could still accept as instantiating ‘insertion’.

- (6) *Ora-ya git-ti-m düğün-e helemaal niet op-ge-maak-t of zo*  
 there-DAT go-PST-1SG wedding-DAT totally not up-PRTC-make-PRTC or so  
 ‘I went there, to a wedding, **not wearing any make up or anything.**’

Before turning to alternation, the logical endpoint of a development from single word insertion to ever larger inserted chunks, we need to discuss one other type of language mixing that often co-occurs with the kinds of codeswitching discussed so far.

### *Foreign discourse markers*

Not all inserted lexical items are content words. Another very typical kind of code-switch is the use of foreign pragmatic particles or discourse markers. Unlike content words, these are not directly involved in conveying the content of



a message but rather demonstrate something about the speaker's attitude and how the utterance should be interpreted. Because they allow the speaker to comment on the ongoing conversation, Maschler (1998) refers to them as elements of 'metalinguage'; Matras (1998) has suggested the cover term 'utterance modifiers.' He helpfully describes them as gesture-like words that help to direct and shape conversation. Typical examples are words meaning 'probably', 'that is' or 'you know', conversational formulas such as 'thank you', 'hello' and the like, and conjunctions such as 'and', 'but' or 'because'. Borrowing of such items into HLs is widely attested in a variety of contact situations. Some examples include: Salmons (1990) about English discourse markers in Texas German; Maschler (1998) on Hebrew discourse markers in Israeli English; Matras (1998) on borrowed discourse markers from various languages in different varieties of Romani; and De Rooij (1996) on French discourse markers in Swahili as spoken in the Shaba region of the Democratic Republic of Congo (then Zaire). To cite one more example, Turkic languages have very few conjunctions (those it has were all borrowed). Many Turkic languages have undergone considerable influence from Russian, especially during the Soviet era, and many of them now have quite a range of Russian conjunctions and discourse particles (Wertheim, 2003 provides an overview). However, our Turkish-Dutch codeswitching data do not show abundant use of Dutch conjunctions, only of discourse markers such as the affirmative particle *ja* 'yes' and the tag question *weet je?* 'you know?'. Apparently, some contact situations give rise to the borrowing of conjunctions more than others; this observation has been accounted for both in terms of typology and of pragmatics.

Contrary to content words, insertion of these markers is grammatically easy. Discourse markers are only loosely connected to the utterance and conjunctions are positioned in between two utterances. Matras (1998) explains that they help to diminish the cognitive load that bilinguals may experience. It may be easier on the mind to have only one set of such gesture-like words. Auer (1999) has suggested the term 'fusion' for just such situations.

### *Alternation*

As bilinguals become more at home in the other language, an even more striking type of CS also soon makes its appearance: actual switching to the other language. This is the type that gave the phenomenon its name, as the speaker actually moves ('switches') from one linguistic system (a 'code') to the other. This subtype is now often referred to as 'alternation'. Usually, the switch is at the boundary between one clause and the next; often those clauses go together in connected speech. In written form, they would often be complex clauses, the main clause being in one language and the subordinate clause in the other. Or, as in the first example below, from Turkish-Dutch codeswitching again, the first part is in one language and the

other part in the other, and the two parts are separated by a conjunction that can be in either language. The Dutch and Turkish clauses in the examples below are integrated to various degrees. One could argue that the ‘then’ clause that goes with a conditional ‘if’ clause forms a tighter unit with that other clause than the two coordinated clauses in the first example, or the afterthought-like addition of *kan ik meelachen* ‘so I can laugh too’ in the final example.

*bunlar oynamaya kalkınca sen de kalkman lazım onlarla en hoe moet je dan op de rest letten?*

‘when they get up to dance, you must get up with them as well, **and then how can you keep an eye on the rest?**’

*Ja, birde onlar orda veya vast of ze werken langer*

‘yeah, and then they have a permanent job or they’ve been working there **for a longer time**’

*... birde weet je wat de probleem is? Birisi burda ne yaparsa geriyanlısı da peşine gidiyor.*

‘and do you know what the problem is? Whenever someone does something here, the conservatives come after her’

*Ja, ondan sonra bana söyledi ki je moet Nederlands gaan volgen.*

‘yeah, and then she said to me: ‘you’ve got to go to Dutch class’

*zamanında oynarsan, antrenmana gelirsen, dan moete ge toch’s zaterdag w el in de basis staan he?*

‘if you play in your own time, and go to practice, **then surely you should be in the starting line-up on Saturday, shouldn’t you?**’

*helede migrain-i varsa zou ik niet aanraden ...*

especially migraine-POSS there-is-if

‘especially if she has **migraine, I wouldn’t recommend ...**’

*ben konuştuğunuzu biliyorum, yüziüme söyle, kan ik meelach-en*

‘I know what you guys are saying, say it to my face, [so] **I can laugh too**’

In the codeswitching literature, examples of alternation are most often studied from a pragmatic perspective, the overriding research question being ‘why did the speaker code-switch at this point?’ This question was discussed in Chapter 3, and it will also play a role in Section 4.4 in this chapter.

Like insertion, alternation is found in pretty much every community in which there are bilinguals who are reasonably proficient in both languages and in which attitudes against language mixing are not so strong that they effectively keep people from engaging in it. The Turkish-Dutch examples used in this chapter resemble

what has been found in most other bilingual settings that have been described in the literature. However, settings do differ in the intensity of codeswitching: in some communities, separation of the languages is more pronounced than in others.

#### 4.2.3 Shift stage: Development towards HL status in the narrow sense

In Chapter 3 we saw that minority languages often give way to the other language in a process referred to as ‘language shift’. Shift basically entails that the community that once spoke the HL in most in-group communication now mostly chooses the socially dominant majority language even in this domain most intimately connected with the home culture. Codeswitching patterns sometimes reflect this shift, as the majority language increasingly comes to dominate not just language choice patterns but also the appearance of bilingual speech. In our Turkish-Dutch codeswitching data we see this in the speech of some members of the second and third generations, who use mostly Dutch interspersed with Turkish utterances, conversational chunks and the occasional word. That is, codeswitching is mostly alternational, and the stretches in Dutch tend to be longer.

Other data from third-generation in-group speech show incessant mixing, in which few utterances are monolingual. To convey this degree of bilinguality, we include the following extended fragment (the supplementary materials include a longer excerpt of the recorded conversation these few speech turns are taken from). It would be hard to say whether this conversation, between two teenage girls, is in Dutch or in Turkish. It is in both, though by and large most clauses can be assigned to either Turkish or Dutch.

##### Fragment of Dutch-Turkish codeswitching

- E *Öyle işte. Ja sowieso. Zij zo ja canım acayip sarma çekti falan hatta eve bırakıyodum. Ik zo söyleseydin aşağıda vardı. Zij zo valla de.*  
 ‘So that’s it. **Yes anyway.** **She was like yes** I really feel like stuffed vine leaves and so, I was actually dropping her off at home. **I was like** if you had told me we had some downstairs. **She was like** really?’
- Ş *Hadden jullie het gemaakt?*  
 ‘Had you made some?’
- E *Wij hadden sowieso aşağıda vardı tencerede vardı ya.*  
 ‘**We had some anyway,** there was some in the pan downstairs.’
- Ş *Ha iyi yapmıştın.*  
 ‘Oh good, you had made some.’
- E *Zij zo ja söyleseydin şapardım enzo.*  
 ‘**She was like yes,** if you had told me i would have **and so forth.**’
- Ş *O da mı biliyo sarmayı neyi?*  
 ‘She also knows about stuffed vine leaves and stuff?’

- E *Ja man. Zij zo kesin bana öğret dit dat. Ik zo öğretiriz.*  
 ‘Yeah man. She was like you definitely have to teach me and so on. I was like I’ll teach you.’
- Ş *Sarma is moeilijk hoor kolay değil pişirmesi ney.*  
 ‘Stuffed vine leaves is really difficult, they are not easy to make and so on.’

### *Connection between codeswitching patterns and language maintenance or shift*

Once people are know both languages well enough that they can practice alternational CS, their proficiency in the other language (the language that is not the HL) is good enough for it to be a feasible choice as the means of communication in many or all domains of life. This means that their bilingualism will be subject to the push and pull factors that determine the ultimate outcome of maintenance or shift, the factors reviewed in Chapter 3. Roughly, one can say that the more often the other language gets activated, the higher the chance that the first language will be affected. This impact can take various forms. One is the appearance of ‘direct’ contact effects: the use of words and structures from the other language. This chapter has thus far illustrated the transfer of lexical items. Chapter 7 will take a closer look at the transfer of grammatical structure. A second form of impact takes is more ‘indirect’: because of the decreased use of the HL a certain rustiness sets in. Lack of practice leads to lessened skill; linguists refer to this as ‘attrition’. Basically, it means that HL speakers experience increased difficulty in retrieving the right HL words and in following the HL grammatical conventions. Finally, impact simply shows in decreased use of the language, as a logical corollary of the increasing use of the majority language. As the number of times speakers choose to use the HL in any given communicative setting approaches zero, the language dies.

### *The HL after the shift*

In this chapter, many phenomena were illustrated with data from the Turkish immigrant community in the Netherlands. Dutch Turkish is a relatively healthy and vibrant HL. As long as there is extensive CS, the HL is still extensively used, and extensive use means extensive practice, guaranteeing a certain degree of proficiency. Many HL studies, on the other hand, focus on languages that have basically dropped out of normal usage. One may wonder whether the ‘CS stage’ is a logically necessary stage in between. In other words, we come to a question that has often been asked with thinly veiled anxiety: is CS a signpost on the road to shift?

After language shift has been completed, by definition the HL no longer exists as a living language. However, as we described in Chapter 3, language shift does not always go to completion, and in any case the process is often a long drawn out one. The further the shift towards the other language progresses, the more

the original group language becomes a moribund language, while at the earlier stages it is more like the vital minority languages often studied in the literature on language maintenance and shift. One might think that the majority language does not need much discussion in a book on HLLs, but on the other hand the HLL often leaves its mark on the language shifted to. This influence is often the only linguistic legacy left behind by the HLL once it really has gone extinct.

As the mirror image of what happens at the very early stages of becoming an HLL, cultural vocabulary describing traditional customs, food etc. may survive the shift. Some of the relevant concepts, like ethnic food dishes, may make their way into the general majority culture, while other words will only be relevant to the culture of the minority community. The reason that this is possible is that language shift does not have to entail a culture shift. To be sure, minorities formerly speaking an HLL may well melt into the mainstream completely (popularly referred to as ‘assimilation’ or ‘integration’), but often they do not, at least not 100%. For instance, after a shift from Yiddish to English in the USA or to Russian in many parts of the former USSR, some Yiddish lexical items were retained, particularly certain descriptive and expressive words and expressions. Here are some examples: *gefilte fish* ‘stuffed fish’ in USA (Jewish) varieties of English from Yiddish *gefilte fiš*; Russian *бeбexи* *bebexi* ‘things, stuff, possessions’ < Yiddish *bebexes* (in turn borrowed from Polish), *халoймeс/халeймeс* *xalojmes/xalejmes* ‘sandcastles, pointless dreams’ etc. (See also Levine, 2000 and Verschik, 2014).

In our Turkish-Dutch data, we can see how Turkish words enter Dutch through codeswitching: once there is a lot of alternation, there is also insertion of Turkish words into Dutch. Many of these Turkish words have the strong cultural associations mentioned above. In the examples below, the word *köy* ‘village’ has connotations of the Turkish countryside, perceived as conservative, culturally old-fashioned and backward, that its Dutch semantic equivalent *dorp* doesn’t have. In the second example, we see the Turkish word *karışık* ‘mixed’ inserted into a Dutch clause; *karışık* is often invoked among Turks to describe their unique immigrant identity.

*maar dat is weer köy, he ...*

‘but that’s the countryside again, isn’t it?’

*hier ben je ook geen Nederlander, je bent gewoon karışık*

‘here, you’re not Dutch either, you’re just mixed’

But not all Turkish insertions have such clear cultural connotations. Presumably, the words we saw above are better candidates for retention should Turkish stop functioning as an active HLL in this community than a common word such as *açşam* ‘evening’, which we see inserted into Dutch in the following example.

*wat ga je akşam doen?*  
 ‘what are you doing **tonight?**’

Common in such situations is also the retention of utterance modifiers or discourse markers from the HL. Consider for example Yiddish *davke* ‘exactly, namely’ in varieties of US Jewish English (whose speakers often do not speak Yiddish anymore) or *bekitser* ‘in short’ in Jewish Russian (Anna Verschik, p.c.). The most likely explanation is that such words are well suited to conveying a Yiddish flavor to the conversation while not requiring the speaker to actually speak Yiddish. This is no different from what makes many people around the world use the English discourse markers *awesome* and *like* in their native language: they feel it conveys a modern flavor to their speech.

Finally, personal names and place names also often stay behind when the language is gone. This is even true if the group that once spoke the language from which the name derives is totally gone. Newcomers may keep the old toponyms, substitute them, or re-analyze them. In Baltic linguistics, the study of hydronyms (names of rivers, lakes etc) is a well-established tradition. Judging by the spread of Baltic-origin hydronyms, it is possible to make assumptions about what areas were once populated by Baltic tribes, for instance. Many hydronyms with Baltic etymology have been established in the territory of modern Ukraine and possibly further to the East, reaching the Ural Mountains. All over the world, evidence such as this suggests a past in which particular languages were once thriving, and then became HLs as their speakers came to be bilingual, and then at some point afterwards, language shift became just a distant memory.

### 4.3 Language change

To understand why bilingualism often leads to language shift, or at least to a position somewhere along the path to shift, it is important to consider how codeswitching affects language. If a community starts mixing in words from another language into its native language, that language is changing, at least lexically. Across bilingual individuals and across communities, there is enormous variation in the degree to which the languages get affected, but very often bilingualism leads to one or both of the languages being spoken in ways that differ from how monolinguals, if there are any, speak them. Awareness of change and difference affects attitudes towards the language, and attitudes affect behavior. The link between bilingualism and language change will be explored further in Chapter 7; here we will take a look at the particularly intimate relationship between *insertional codeswitching* and one type of language change, namely *lexical borrowing*.

As more and more speakers of the HL use a particular word from the other language in their daily HL speech, that word becomes more and more part and parcel of the HL. At some imperceptible point, it will have become an established word in the HL. As an established word with foreign etymology, it is a loanword. Exactly when a word has been used often enough as an insertional code-switch to have reached this status is impossible to say. As a rule of thumb, though, we can assume that the word needs to be used by a large part of the speech community and needs to be, if not the default choice than at least a common choice whenever the concept it stands for needs to be verbalized. There is no need to be more precise, since the path towards the status of established loanword is a continuum and any decision to establish a cut-off point, beyond which the word is accepted as a loanword, is essentially arbitrary.

Importantly, being established as a loanword doesn't mean that the word is not recognized anymore as a foreign body, so to speak. As long as there is bilingualism, it is likely that most or all foreign words are widely recognized as foreign-origin. Whether this matters or not depends on socio-political circumstances. This is because there is a tight link between the societal issues discussed in the previous chapter and the social context of communication that is the central theme of the present chapter.

#### 4.4 Codeswitching in its social context

Here we will consider the broader context of codeswitching, raising a set of further issues, starting whether codeswitching represents a third language or not. We then present a few cases studies and turn to the concept of 'linguaging'.

##### 4.4.1 Does codeswitching represent a third language?

Language use that is characterized by extensive CS often reflects just one of the registers in the repertoire of speakers. Grosjean (2001) has suggested that bilinguals move up and down a continuum of bilinguality, meaning that they may be in a relatively bilingual or in a relatively monolingual mode. This is meant to be understood in terms of activation: while both languages are always activated to some extent, a bilingual mode activates them both to a high degree, while when speakers are in a monolingual mode the language not being spoken is relatively dormant. An example of a monolingual mode is the use of the dominant language in a workspace where no one else knows the HL; an example of a bilingual mode is a conversation with peers who are all fluent in the HL and the dominant language.

In the field of HLs, there has been a great deal of attention for the social meaning conveyed through language choice and codeswitching. The simplest model has been the division into a 'we' code and a 'they' code (Gumperz, 1982), pitting the two languages of a bilingual against each other. One, the HL is the language of the group ('we'), and therefore of solidarity, while the other language, the one that is powerful in society, is the language of the outside ('they'), but also the language that provides access to economic opportunities. Based on investigating linguistic choices made in various countries around the world, Gumperz and associates suggested that by merely using the 'we' code, speakers conveyed an appeal to solidarity with their interlocutors, emphasizing their shared ethnicity or group membership. Use of the other language, the 'they' code, on the other hand, conveyed an appeal to power, emphasizing participation in the socio-economic fabric of the wider society.

It is not so easy, then, to establish what the repertoire is of a typical HL speaking community. Minimally, it includes (at least) two languages. If there is little codeswitching, and each language has its proscribed domains of use, there is not much more to it. However, since there is often extensive codeswitching as well, it is important to include this in the repertoire description, and this is where it gets hard. One possibility is that there are three 'languages', again with their fixed domains of use: the majority language for communication outside the group, the HL for communication with monolinguals in that language (e.g. first-generation immigrants or people 'back home') and a mixture of the two for in-group communication with bilingual peers. However, there are two problems. One is that there are many different ways of mixing the languages (as documented in Section 4.2), and it is far from obvious that they can all be lumped together as reflecting one 'variety', a 'bilingual lect' (Backus, 1996). The other is that while we have no trouble recognizing that the majority language and the HL have some degree of stability, allowing us to recognize speech as for instance 'Dutch' or 'Turkish' in the first place, we generally don't know whether the mixture has similar stability. There are, for example, no dictionaries or grammars of a mixed Turkish-Dutch bilingual lect. However, for some bilingual communities, a number of fairly stable mixtures have been postulated. Sarhimaa (1999), for example, situated five varieties on a continuum to describe the repertoire of the Karelian speech community in Northern Russia, close to the Finnish border. At the end points are varieties of monolingual Karelian and monolingual Russian, and in between are Karelian-dominant mixing, balanced mixing, and Russian-dominant mixing. However, in most described cases, there seems to be simply a lot of relatively ad-hoc codeswitching between the languages involved, with individual switches being determined by pragmatic motivations of the speaker and various other factors related to the proficiencies of speaker and hearer and conventions in the community. Perhaps then, it is best



to describe the repertoire of a bilingual community as simply involving two languages. While language choice conventions may specify the use of just one of the languages in some communicative domains, they may also specify the use of both languages in other domains. Under particular circumstances, the diachronic result of extensive codeswitching may be the solidification of mixing, and in such cases a mixed language may be born.

It was mentioned that individual cases of codeswitching may serve pragmatic functions (this type of codeswitching is sometimes called “meaningful juxtaposition”). The rest of this section will look at some of the ways in which it does this. By and large, people code-switch for two reasons. One has been described in detail earlier in this chapter: the word, chunk or expression is easier said in the other language. This can be for semantic reasons (it conveys the concept best) or because it is activated easier. The second reason is pragmatic: the switch serves to draw the attention of the interlocutor to a specific communicative function the switch indexes. Sometimes this is due to the sheer shock value of the change in language: the interlocutor pays more attention to what one is saying because something unexpected happens, in the same way as when one raises one’s voice, laughs when saying something, or suddenly starts to whisper. However, if codeswitching is frequent, its potential to draw attention through unexpectedness is obviously limited. The other source of indexicality, however, will virtually always be present: each language tends to symbolize a set of values. This is where the terms ‘solidarity’ and ‘power’ come in.

To return to Immigrant Turkish, Keim and Cindark (2002) compared three different networks of young second generation Turkish immigrants in the German city of Mannheim, and found that only one group code-switched profusely, a group that emphatically saw itself as bicultural. Other groups didn’t engage in any mixing. In one case this concerned a network of people who placed pride in both their Turkish heritage and their German identity: for them, the best way of honoring their Turkish and German identities is to speak both languages as best they could. In their ideology, speaking a language well means keeping it free of influence from the other language.

By and large, the evidence we have seems to suggest that codeswitching is a natural outcome if speakers know two languages, but that it will be constrained if there are social factors that work against it. If the community, or some parts of it, frowns on language mixing, intensive codeswitching as illustrated above would constitute an act of serious rebellion. As a result, the literature also contains descriptions of communicative practice in bilingual communities in which codeswitching is conspicuously absent, or just rare. Poplack, Sankoff and Miller (1988) note that French speakers in Canada often flag the English words they are about to use, sort of as an apology. Treffers-Daller (1994) notes that over the years the incidence of codeswitching between French and Dutch in Brussels has gone

down. In both of these cases, linguistic tensions may well have something to do with these developments. An interesting case is described by Aikhenvald (2002). The Northern Arawak language Tariana, spoken in the Vaupés region near the border between Brazil and Colombia, is heavily influenced by the unrelated local lingua franca Eastern Tucanoan (Aikhenvald, 2002). This influence is only found in the grammar, though. There are very few loanwords, and there is almost no codeswitching. The Tariana maintain a purist language ideology that effectively stops people from ever engaging in the use of foreign words.

#### 4.4.2 Linguaging

The overarching perspective in this chapter has been that whenever we speak, we do more than just convey the referential information that is contained in words and the grammatical patterns in which they are ordered. Communication is much more than that, and while this has always been recognized by linguists in the sense that there is a field called ‘pragmatics’, the full extent of communicative richness is not always taken to be part of what linguistics should study. Pragmatics, as traditionally understood, is about sentence meaning. While lexical semantics deals with the meaning of words, and truth conditional semantics about the literal meaning of sentences, a theory of pragmatics explains how sentences receive their meaning in the context. This meaning often contains much more than what is strictly inferable from the words and structures. For this reason, it is hard to put an upper limit of what still belongs to pragmatics, and thus to linguistics. In fact, many sociolinguists reject the notion that there should be a limit and refer to ‘social meaning’ instead. As we communicate, we convey social meaning constantly. Some have used the word ‘linguaging’ (Jorgensen, 2008) for this, to convey the idea that ‘language is a verb’, which echoes the earlier turn in anthropology in which culture was seen as something people perform: ‘culture is a verb’. In terms of language, ‘linguaging’ can be seen as roughly synonymous with ‘language use’, the topic of this chapter. The essential point is, terminology aside, that anything we say carries social meaning: it says something about how we wish to portray ourselves, how we define the communicative situation, and what we want to achieve beyond conveying the purely referential information.

It may be useful to point out that this is still a somewhat limited, ‘lingua-centric’ view of what we do while communicating, as the verbal message contained in an utterance co-occurs with numerous other nonverbal signs, such as communicative markers (hesitations, affirmations, question markers, etc.), sounds, gestures, head movements etc. (Clark, 1996). In HL studies, we investigate what happens to the language, but there may well be parallel developments in these other modes of communication.

## 4.5 Summary

This chapter has focused on bilingual language use. We have tried to explore why codeswitching exists in the first place, what forms it takes, and what factors determine this. How does codeswitching come to be part of the bilingual repertoire? We also discussed the issue of how language mixing evolves over time. How does bilingualism affect long-term developments in the HL? We then turn to the social and pragmatic functions of codeswitching, reconnecting the characteristics of codeswitching and language change with the social context in which HLs function as described in Chapter 3. To what extent do ideologies on language use determine what elements of language are and are not switched? Finally, we touch upon the concept of 'linguaging'. This involves the notion that the way we use language carries social meaning. It shows how we wish to portray ourselves: language choice is part of this identity creation.

# Methods for collecting heritage language data

## 5.1 Introduction

As with all languages, there are various ways of collecting data on HLs. Three chief traditions inform current practice, namely collecting spontaneous language data, collecting questionnaires, and data elicitation. The first method of collecting spontaneous data is inherited from the study of language contact and minority languages. Many studies are based on recorded samples of spontaneous conversation, sometimes in the form of interviews or parent child interactions. Such data are immensely valuable as a window on how the language is used in everyday life and what its structure and vocabulary look like. It can only be used, however, for HLs which are actually still in use. In addition, it is a time-intensive way of collecting data, which makes it hard to obtain enough data to make generalization to the community level feasible. The second tradition of using questionnaires comes from sociolinguistics and is especially for gaining information on attitudes and on use patterns. Surveys make it possible to reach enough participants to allow general conclusions but have the disadvantage that they rely on metalinguistic knowledge. People can only answer questions about issues they are aware of, and in settings that involve linguistic tension or anxiety, it is not always obvious that people are truthful in their answers. Interviews go some way towards alleviating those problems. The third methodological strand of eliciting specific data is informed by traditions in the study of second language acquisition and elicits data in more or less controlled settings. This ranges from pencil-and-paper judgment tasks to laboratory-based psycholinguistic experiments. Potentially, these methods can solve the problems of both other methods, but in practice there are many problems to overcome in moving the informal practice of speaking HLs at home and on the street into the sterile conditions of the laboratory.

Together, these different ways of collecting data make it possible to triangulate findings and state conclusions with greater confidence. On the other hand, data do not always converge (for discussion see Nagy, 2015), and in this sense the use of different methodologies to study the same HL setting can also expose the full complexity of language use and language proficiency in ways that would be impossible if only one method were used.

As the chief goal of the chapter is to equip researchers with a firm basis on which to choose the methods that best fit their research questions, the chapter will start with a brief discussion of the issue of validity in the choice of methods: we will discuss the concepts of internal, external validity and ecological validity in Section 5.2. This will help evaluate the various methods that will be discussed in the main section that follows in Section 5.3. Section 5.3 will include a historical overview of the various methods that have been employed, with examples taken from various representative studies. The emphasis will not be on what these studies found but on how the method employed allowed particular kinds of conclusions. We will also discuss ways to handle recordings, and how issues of transcription and annotation can be approached. The third main part of the chapter (Section 5.4) describes and evaluates the main methods, while the final Section (5.5) briefly discusses how to choose the appropriate method for the research you want to engage in. The most central concern throughout will be the degree of external and ecological validity that is needed, as it is not always easy to find a balance between the richness of the output afforded by recordings and the control that experiments afford over what that output consists of.

**Main goals of this chapter**

To briefly present criteria for the evaluation of the various methods used.

To survey the different methods used to gather HL data.

To learn how to choose the appropriate method in a given setting.

To become acquainted with some practical aspects of recording and transcribing data.

## 5.2 Criteria for evaluating a particular method: Validity

Ordinarily, research methods are evaluated in terms of their validity: how valid is a given method for achieving the stated purpose of the research? Validity is an enormous issue in discussions of research methods in the behavioral sciences, and we cannot even begin to discuss it fully here. We will just briefly introduce three measures of validity.

**Internal validity** is the extent to which you can say that no other variables than the one being studied was responsible for causing the obtained results. Do you actually measure what you think you are measuring? For instance, imagine you study an HL that has SOV word order and is dominated by a majority language that has SVO order. Your hypothesis may be that speakers increasingly resort to SVO order. In that case, you may want to investigate word order in naturally produced conversation or design an experiment in which HL speakers have to produce sentences. With both methods, your hypothesis is that they will produce

more instances of SVO than a control group consisting of monolingual speakers of the HL. In both cases, however, you will have an internal validity problem to contend with. Basically, if a speaker of an HL uses a grammatical structure that was not present in the language before contact, it is difficult to say whether that usage was the result of direct influence from the other language at the moment of speaking (i.e. a case of ‘interference’). The reason is a problem of internal validity: the cause of the selection of that structure may indeed be interference, but it may also be that this is the structure this particular speaker always uses. In the latter case (‘imperfect acquisition’), interference happened in the past, to this very speaker or to others from whom he picked up the language (for example his parents), or to ancestors further down in history. In the case of the experiment, let’s say you give the participants simple sentences to imitate (we will discuss such research later in this chapter; essentially, this design would give you maximum control over what participants produce, which maximizes comparability). If participants produce SVO sentences when repeating stimulus items with that order, they may do this because their HL grammar now has SVO as its basic order, but they may also simply remember the stimulus sentence correctly and produce it verbatim from working memory. If you then conclude that SVO is produced a lot by the HL speakers because of influence from the other language, you face a problem of internal validity: the SVO responses may not be caused by their HL grammar containing SVO, but simply by their ability to repeat the stimulus item. We will discuss a case like this later in this chapter, in which the problem was tackled by having the crucial structure embedded in a longer stretch of talk, too big to contain in working memory.

Another common problem involving internal validity occurs when a speaker rejects a particular linguistic example presented by the linguist in an acceptability task. The rejection may be attributed to the ungrammaticality of the grammatical structure of the stimulus item, while in fact the participant just found the situation described in the example to be unlikely or culturally inappropriate, rather than linguistically ill-formed. We will see that internal validity is a tricky problem in studying heritage speakers; it is especially salient for the phenomenon of ‘imperfect acquisition’, alluded to above and to be discussed in detail in Chapter 8.

While internal validity is about the degree to which the data actually justify the conclusions you want to draw from them, **external validity** is about the degree to which the data allow for generalizing the conclusions beyond the participants in and the circumstances of the study. This type of validity is very important in behavioural research, as it always aims to make statements about human beings in general or at least about a meaningful subset. Consider the basic linguistic task of language documentation. If you want to describe a small endangered language, and you work, for instance because of limited funding or access, with a single

speaker who has been living away from the speech community, there is a problem with external validity. This speaker may have developed all kinds of special characteristics in his or her language use that are different from those of the people in the community, such as lexical and grammatical influence from a foreign language and loss of vocabulary items that are not relevant to his or her current living situation.

Another example of a problem with external validity is the common problem in the study of minority languages that it is hard to determine the numbers of people required to say, afterwards, with some confidence that the data represent the community. In our studies of codeswitching in the Dutch Turkish community, for example, so far perhaps a total of fifty people have been studied, out of a total population that over the years must have numbered half a million. In addition, participants were often selected because they had been observed to practice codeswitching in their everyday speech. The sample of fifty people, selected non-randomly, recorded in settings that facilitate codeswitching, and recorded for only an extremely small proportion of their everyday informal interaction, at one particular moment in time, does not justify generalizations about how Dutch Turks speak. Yet, such concerns are rarely voiced. On the other hand, a consistent finding in HL research is that heritage speakers differ quite a bit from one another, which tends to limit researchers' appetite for making sweeping generalizations.

Finally, there is a third type of validity which is called **ecological validity**: if the description of everyday speech is your research target, can you be sure that the type of language data you have collected are exactly like the naturally occurring language use that you were after? Might it not be some artificial result of your data collection procedure? In general, this problem seems to be absent if your method is to use recordings of naturally occurring conversation. However, people may behave differently when they know they are being recorded, since being recorded is not part of normal everyday interaction. All other methods definitely have to contend with the issue of ecological validity, as they all distort reality in some way. To continue with the fieldwork example, there is an ecological validity problem with the common practice to ask translations of isolated sentences of the type 'Mary built a house'. These sentences often contain lexical third person referents for both subject and direct object and lack all kinds of discourse markers and further contextual support. They have internal validity (indeed they represent the structures used to say 'Mary built a house.') but they are lacking in ecological validity, since ordinary speech will contain few sentences with two lexical referents, they will often be in the first and second person, and they will come with all kinds of contextual support, including discourse markers (e.g. 'uhm, yeah, she built that house in, oh, 1975, or something, I think'). In addition, of course, the practice of translating arbitrary sentences is not part of everyday behavior.

All three kinds of validity will be seen to raise important issues when evaluating the use of the different methods that are regularly turned to in HL research.

### 5.3 Overview of methods used

In one of the key early works in sociolinguistics, Labov (1972a) listed *texts*, *elicitations*, *intuitions*, *experiments*, and *observations* as the key methods used in linguistics, and the same holds true for the study of HL to a large extent. Observations of spoken data have been the prime source of information on how HLs are used, while survey data have typically been used to gather data on language choice. Written data have not been used much historically in HL research (but see Section 5.3.2 for some references). Thanks to the increased occurrence of informal writing, now that modern communication technology has freed up writing on the Internet and in text messaging from the pressure to reproduce the standard norms that traditionally applies to writing interest in written documents has increased. Finally, experiments have become more popular in HL research recently.

#### 5.3.1 Spoken data

Much of the linguistic work on HLs has focused first on the basic task of description. Interest in minority languages often derives from one of two circumstances: either the language is only poorly described if at all, or the language is undergoing remarkable changes. The first circumstance speaks to what is the key task of linguistics of documenting the linguistic diversity in the world (Evans & Levinson, 2009). The second, likewise, makes HLs interesting cases for the attempt to account for a basic linguistic fact: languages are dynamic. As linguists, we need to understand why languages change all the time and how they do so; HLs provide interesting material.

Accounting for diversity and change requires a comparative database that contains accurate descriptions of as many languages as possible. Now if you were given the task to describe formal British English, you would be able to rely on a lot of findings that are already out there: descriptions by others (in published grammars), analyses of particular constructions, dictionaries, and easy access to native speakers. None of these advantages await the student of a hitherto understudied HL.

It stands to reason that if you need to find out how people in a given community speak, you just go there, listen, and somehow note down what you hear. Speech is very fast, so you cannot transcribe it as it unfurls. Therefore, you will need to record that speech, and base your descriptions on the transcripts. The question



is: what do you record? Sociolinguistics, the field that has dealt with this question most extensively, has had to deal with the Observer's Paradox (Labov, 1972a): you may want to know how people speak, but the knowledge that they are being recorded will alter their way of speaking. Obviously, this raises ethical issues to which we will get back in the final section, but suffice it to say that you cannot record people without their consent (or in the case of young children, the explicit consent of their parents), not even in the interest of scientific investigation. This is a problem, however, because of the reasonable sociolinguistic tenet that if we want to investigate language in any setting, it should be the most basic communicative setting in which it is used: ordinary informal face-to-face communication between friends and family. We will see later that there may be good reasons to also look at other settings: more formal language use, written language, and conversation with strangers. By and large, however, everyday language use is what we aim to describe and account for.

Thus, the recordings on which we are going to base our descriptions and analyses need to be of informal conversation between people who are used to conversing with one another. Various further questions now present themselves: which participants to select, what to ask them to do, how to organize the recording, and how to handle the data. How they are handled depends on the specific research question, and thus specifics will differ across studies. Here we will discuss these issues assuming the most basic question of all, as it has informed many HL studies, past and present: how do people in the community speak?

### *Participant selection*

External validity requires that a sufficiently large and random sample of people from the community is selected for participation in the study. Typical sources of variation in a community need to be taken into account, so you would want people of different age groups, of both genders, and of various social groupings (whichever seem to be important in the community under question, e.g. class, race, ethnicity, religion, etc.). However, these concerns immediately confront the researcher with a severe feasibility problem: you can't possibly record as many people as you would need to draw confident conclusions about the community in general. In practice, you will have to settle for a less comprehensive picture, selecting whatever types of speakers seem most relevant. In cases of endangered languages, there is no problem of this kind as researchers will work with the only speakers still alive; in less dramatic cases, it means selecting participants that allow you to get the data that allow you to investigate your research question. If that question is the general one about how people speak in the community under study, you'll have to select representatives of whatever subgroup seems relevant, in whatever numbers you can manage, and just be careful not to state your generalizations too ambitiously.

In a typical study of codeswitching in an HL, Backus (1996) selected 21 participants from the immigrant Turkish community in the Netherlands on the basis of their willingness to be recorded. At the time, much was still unknown about possible codeswitching patterns in general, let alone for this community. Therefore, any member of the community would do as long as he or she code-switched in everyday speech. However, since that couldn't really be known beforehand, Backus started out recording any speaker who was willing to participate. It turned out they all code-switched, though some more than others. In terms of participant selection, we could say that this was an exploratory study: if in principle any participant can be useful or useless given the purposes of the study, it makes sense to select participants randomly. It turned out that the participants who were first generation immigrants did not code-switch much, while members of the second generation as well as a so-called 'intermediate generation' (those who immigrated as kids) practiced it a lot. This finding was used in later codeswitching studies to direct participant selection towards those groups who were likely to engage in it. Note, however, that this bypasses a problem of external validity: the numbers were way too low to be considered evidence that all first generation members code-switch only minimally and all members of the other two generations do it a lot.

### *Research design and procedure*

Issues of design seem relatively straightforward if the data to be collected are recordings: you get people to converse, and you record the result. However, there are some options to consider, and the choices you make have consequences for the data you will get, and hence for what you can conclude on the basis of them. Here are some basic questions to consider:

- Do you want people to talk in groups or in dyads?
- Do you want them to talk freely or about particular topics?
- Do you tell them beforehand what the study is about?
- Are you going to be present during the recording?
- If so, are you going to participate?
- For how long do you want them to talk?
- Are you going to record them only once or several times?
- If the latter applies, are you going to vary the circumstances (interlocutors, topics, settings)?
- And do you just want to record the speech, or do you want to videotape the conversation?

If it is vernacular speech you are after, a case can be made for having people talk in groups and in duos, since these are the configurations common in everyday life. However, transcription gets more difficult the more speakers are present, since in

groups people interrupt each other more, talk more often at the same time, and start little side conversations among subsets of the group. For that reason, you may want to compromise ecological validity somewhat and keep the groups small, like up to four or five people. If you want to be able to compare the various recordings, it may be useful to have all participants talk about the same topics. It will make it possible to check to what extent people make the same lexical and grammatical choices, for example. However, this constitutes a more serious compromise of ecological validity, as people do not normally engage in conversations in which they are told what to talk about. If the conversation is framed as an interview setting, this kind of control would be more natural, since in one-on-one interviews the interviewee will naturally accept that the interviewer determines the flow of conversation. Obviously however, an interview setting comes with a huge ecological validity problem of itself if the goal of the research is to describe the language use of everyday conversation: we don't go around being interviewed all day. The combined problems caused by the Observer's Paradox and the ethical objections against recording people without consent make it impossible to avoid problems with ecological validity entirely: it is just something to live with, and to deal with in the best way possible given the specific research questions you have. The bigger the problem with ecological validity, though, the harder you have to work to show that your research question is a relevant one.

Many studies aimed at describing lexical and grammatical characteristics of minority languages, including issues of codeswitching and language mixing, have used recorded conversational data but differ from each other in the ways in which they have dealt with the abovementioned questions (Table 5.1).

Considerations regarding external and ecological validity (i.e. whether generalizations are warranted) are rarely the topic of explicit discussion. Note, however, that it is not obvious that findings from these studies can be interpreted as representative for the entire community. The few studies that document variation within an HL community testify that generalization may be an issue that needs further study (Bentahila & Davies, 1995; Jacobson, 1977; Backus, 1996; Li Wei, 1994). Keim and Cindark (2002) examined differences in codeswitching styles in the Turkish immigrant community in Mannheim and found at least three different styles of speech among members of the second generation: one that was mostly informal urban German, one that alternated between monolingual standard-like styles of both languages and one that was mostly mixed. The styles were used by distinct groups, with distinct identities. Clearly, if only one of the groups had been studied, generalizing its communicative style to the level of the community would have been incorrect. It would have assumed external validity where there was none.

**Table 5.1** Some earlier studies and the data collection methods used

Author	Date	Methods used
Haugen	1953	Written sources, census surveys, informal observations, recordings.
Lehtinen	1966	One-on-one recordings with one individual in various locations.
Pfaff	1979	Use of a corpus of spontaneous conversational recordings of fluent bilinguals made by a community member.
Hlavac	2003	Interviews with 100 members of the Croatian immigrant community in Melbourne (one on one or duo or group?)
Halmari	1997	Longitudinal case study of three Finnish-Americans covering sixteen years, since arrival in the U.S (6.8, 7, and 11.6 years old at the time of immigration).
Smits	1996	Free conversations, translation tests, and acceptability tests (using stimuli collected in research 23 years previously) with older heritage speakers of Dutch in Iowa.
Doğruoz	2007	One-on-one interviews and group conversations with 43 bilingual Turkish-Dutch speakers between the ages of 17–45, both first and second generation. In addition, acceptability ratings of conventional and unconventional constructions.
Lleó	2016	Spanish-German and Spanish-Catalan child interactions, available through the HZSK (Hamburger Zentrum fuer Sprachkorpora).

### *Data handling: Storage, transcription and annotation*

Once recordings have been made, they need to be prepared for analysis. To make sure you can work on the data to your heart's content, they should be recorded with equipment that allows secure storage on computers and in databases. As recording techniques have improved and become more accessible, the standards for audio and video data have gone up, with the trade-off that good recording requires a considerable amount of memory space.

The first step in data analysis is transcription of the data. The task may seem straightforward, but as soon as you start transcribing you will realize that decisions have to be made that once again depend on the specific research questions you have, and thus also on your theoretical outlook. Here are some of the more important questions to decide on:

- Do you need to transcribe everything or is your research question so specific that you can get away with transcribing just the passages that hold relevant material?
- Do you need to transcribe the speech in phonetic detail or can you get away with an 'orthographic' transcription, i.e. writing the words the way they would be written in a written text?

- How much of the conversational structure should you preserve in your transcript, i.e. do you need to record all hesitations, pauses, cases of overlapping speech, cases where speakers raise their voice or whisper, laughter, etc.?
- In case of videotaped communication, in what detail do you want to transcribe non-verbal communication, such as gestures, body movements, gaze direction, etc.?

A different kind of question is what software to use for transcription. If you are going to be the sole user of the data, you might think that you can just open a new Word file or whatever, and start transcribing. However, here too it is important to think beyond your immediate research concerns. First of all, it is increasingly important that data are stored in an accessible format, also to be accountable. Scientific work has to be replicable: another researcher who does not agree with your data analysis should in theory be able to redo your analysis, and therefore have access to the data. Storage in a larger database is the most secure way to do this, since the database is professionally managed and comes with useful guidelines you can follow, keeping you from reinventing some wheels.

Second, transcribing in a format specially developed for transcription will help you down the road. Such software packages help you organize the transcriptions in ways that have proven to be useful, force you to fill in crucial metadata information you might otherwise forget about (such as social and language background characteristics of the speakers, circumstances of the recording, topics in the conversations), and most importantly perhaps, they make subsequent analysis easier because they come with in-built search functions and aids for quantitative (and sometimes statistical) analysis of the data. A handy intermediate choice is using Excel, since it comes with many ways to do quantitative analysis. It is very important, while gathering, storing, and analyzing your data, to use a systematic and detailed metadata schema.

A number of database storage and analysis programs have been developed over the years. There is much happening in this field, so we can only give some pointers, citing liberally from the websites of these programs.

The *CLAN* program was designed and written by Leonid Spektor at Carnegie Mellon University, Pittsburgh, USA. CLAN stands for Computerized Language ANalysis. The program CLAN is designed specifically to analyze data transcribed in the CHAT format, a format frequently used in the Child Language Data Exchange System (CHILDES) Project, as well as other database systems. These are now being added, and include the AphasiaBank and CABank. These are part of the overall TalkBank system <[childes.talkbank.org/clan](http://childes.talkbank.org/clan)>.

The program *Toolbox* is a data management and analysis tool especially designed for field linguists and developed by the Summer Institute of Linguistics

International. It is particularly useful for lexical data, and for parsing and glossing recorded and transcribed texts, but it can be used to manage virtually any kind of data. Toolbox is a text-oriented database management system. The underlying software package offers full user flexibility in the design of any type of database, but for ease of use, the Toolbox package includes readymade database definitions for typical dictionary and text corpora <[www.sil.org/computing/toolbox](http://www.sil.org/computing/toolbox)>.

*Praat* is a computer program for analyzing, synthesizing, and manipulating speech, and creates high-quality visual representations. The Praat program was developed by Paul Boersma and David Weenink of the Institute of Phonetics Sciences of the University of Amsterdam, the Netherlands <[www.praat.org](http://www.praat.org)>.

*ELAN* is a professional tool, developed at the Max Planck Institute for Psycholinguistics in Nijmegen in the Netherlands for creating complex annotations of video and audio resources. An ELAN user can add an unlimited number of annotations to audio and/or video streams. Annotations can be sentences, words or glosses, comments, translation or a note on any feature observed in the recording. Annotations can be created on multiple layers, called tiers. These can be hierarchically interconnected. An annotation can either be time-aligned to the recording or it can refer to other existing annotations. All annotations are in Unicode and transcriptions are stored in an XML format <[www.tla.mpi.nl](http://www.tla.mpi.nl)>.

### 5.3.2 Written documents

Given the widespread consensus that the target of observation in linguistics, especially in sociolinguistics, is informal everyday vernacular in-group speech, it is understandable that most data in HL studies are drawn from spoken rather than written data, in addition to experimental data. However, the attitude towards the usefulness of written data is changing, and this is related to the fact that in recent times, there has been an enormous increase in informal writing. When people write emails, text messages and posts on chat forums, they don't feel as bound to the rules of the standard language as they do when writing at work or at school, i.e. producing essays, formal letters, reports, memos, etc. The result is that informal digital writing represents another form of vernacular, and this makes it relevant for HL studies as well. Below, we will first take a brief look at the use of written data in older studies, and then explore some of the possibilities offered by digital written data.

#### *'Ethnic' newspapers and other written documents*

For several HL communities there are older published texts which potentially document the gradual development of the language. A case in point is Texas German. In Texas there were German-language newspapers from the 19th century

until 1957 (Boas, 2009). However, these have not been a widely used source for HL research, with the exception of some lexical studies such as Gilbert (1965). The reason is probably their supposed lack of ecological validity: the newspapers probably presented a more 'standard' German than the varieties that were actually spoken. Nevertheless, such publications were often a good source for studying the loanwords from the other language that had penetrated the lexicon of the HL. Loanwords, loan translations and borrowed constructions often make it into the written standard once they are in general use and not the target of overt purism, and hence show up in such written sources. Modern corpus linguistic research makes use of this, for instance, in the study of Anglicisms in the world's languages, e.g. Zenner, Speelman & Geeraerts (2013).

Another source that has been available to older HL studies consists of people's autobiographical documents, such as diaries and personal letters. The same problem with ecological validity holds for these documents as well, but to a lesser extent, since private writing is less subject to normative pressure. Klatter and Kroon (1997) used private documents of Dutch-origin immigrants in New Zealand to study the development of Dutch as an HL in this migration context, showing changes over time in use of tenses and verb inflections. In addition to their linguistic usefulness, of course, both ethnic newspapers and private documents of community members provide a wealth of information on the cultural and anthropological characteristics of the HL community, including people's attitudes to the HL and to the dominant language.

### *Chat sites and social media*

In the digital age, the playing field has changed concerning the use of written data in sociolinguistics. Thanks to the widespread opportunities for informal writing, social media and chat forums provide great opportunities for vernacular data collection. Where writing used to be confined largely to formal contexts, with the exception of shopping lists, postcards, short personal notes and the like, nowadays people spend much more time writing informal text messages to each other and contributing posts to threads on discussion forums. It is far from clear to what extent such writing is still sensitive to some of the normativity that guides more formal writing, but what is clear is that there are remarkable similarities with informal spoken conversation, concerning linguistic choices at the levels of words, grammar and discourse style.

Particularly attractive as sites for data collection are chat forums, tweets, and other social media targeted at HL speaking communities. Dorleijn and Nortier (2009) and Androutsopoulos (2013) provide state of the art overviews of codeswitching studies in computer mediated communication.

### *Observations, e.g. of linguistic landscapes*

Perhaps as part of a general trend to widen the focus of linguistics beyond its traditional form-based domains of sounds, words and grammar, sociolinguistics increasingly looks at the general communicative ecology within which language use takes place. Partially this shows in the increased attention to the multimodality of communication, prompting researchers, for instance, to transcribe non-verbal aspects of communication in addition to what is said. However, the trend is also visible in the more anthropological approach to language choice issues taken in *linguistic landscape* studies. In such studies, concrete visual manifestations of the HL in the public domain (signs on stores and restaurant windows, street name signs) are documented, providing valuable information about the degrees to which, and the ways in which HLs are used.

### 5.3.3 Survey data and questionnaires

Survey data are often used to collect data regarding language choice. Most studies have participants fill out questionnaires in which they are asked about which language they speak to whom and in what circumstances. The largest quantitative reach is provided by census data in countries that have them. Sometimes existing official census data provide useful information, but not all countries collect such data and their usefulness to sociolinguistic investigation is necessarily limited.

There is often a language question embedded in such national surveys, asking which language people speak, for example, at home. Not all countries administer a census, though, and then survey research is the only way in which language choice data can be gathered from large numbers of speakers. As is common throughout Humanities and the Social Sciences, there is a negative correlation between the size of the population sample and the theoretical depth of the analysis the data allows. Questionnaires usually focus on common settings, differentiating between language chosen with particular interlocutors (parents, siblings, other family members, friends, etcetera) or particular situations (at home, on the street, in school, at work, etcetera), through particular media (face-to-face conversation, broadcast and print media, Internet, etc.). Eversteijn's (2011) study of language use in the Turkish immigrant community in the Netherlands is a rare example of a study that asked such questions in combined form, which is useful because language choice with a particular conversation partner may differ across conversational settings (e.g. 'What language do you use with your siblings when you're in a store on vacation in Turkey?'). For many such situations, participants had little trouble saying which language they choose, especially if the language abilities of interlocutors necessitated either Turkish or Dutch. However, people are generally hard pressed to give a full account of their language choice behavior through



responses to multiple choice questions, for at least three reasons. First, as holds for a lot of human behavior, the reasons people have for the language choices they make are sometimes very subtle, and not readily accessible to consciousness. Second, choices may vary across conversations, so that the same conversational setting may not always yield the same language use. Third, bilinguals may code-switch back and forth between the languages, and participants are rarely able to say more about this than that they indeed use both languages.

The qualitative study of language choice in relatively few actual conversational settings and the quantitative study of reported language choices in many different settings nicely complement each other. Codeswitching has proven difficult to ask questions about in surveys, most likely because people are less aware of when and how they practice it. It is easy to answer 'both' to a question like 'which language do you use with your best friend?', but it is not so easy for the average participant to provide further insight into when exactly language A is used and when language B, let alone what conditions switching from one into the other.

In addition, it is never certain that self-reports yield truthful answers, or whether the answers people have given in full honesty actually match the facts. For example, Cavanaugh (2013) found that speakers of the Bergamo dialect of Italian think that their dialect is nearly gone, and claim they virtually never speak it, while usage data show that it is used extensively. At the other end of the continuum, ethnographic research and conversational analysis allow detailed investigation of the factors influencing particular episodes of language choice, in a small number of actually observed conversations. Generalization will always be an issue with this approach, however, since there is no foolproof way of knowing whether the episode observed was typical or not.

Many community studies of HLs have employed a questionnaire of some type. A good example is Irizarri van Suchtelen (2016), who carried out a web-survey in the Chilean former refugee community in the Netherlands. There was a number of multiple choice questions about language use and language attitudes, but informants were free to add qualitative comments and give examples of language mixing (which they did). Another example of a successful and balanced questionnaire can be found in Nagy, Chocieł & Hofmann (2014)

#### 5.3.4 Experimental data

The problems concerning the external validity of spontaneous speech data have led to a gradual increase in experimental methods in research on language contact. If it is so difficult to gather enough data to obtain sufficient numbers of instances of particular words or structures, it may be useful to elicit these elements. This can be done in many different ways. The basic principle is that

a carefully constructed experimental design allows you to test a hypothesis. In one of our studies about HLs in Holland, we wanted to know whether Dutch Turkish has changed its way of forming subordinated clauses, away from the Turkish non-finite and pre-verbal type (meaning the subordinate clause is not inflected for tense and person and is positioned before the matrix verb) to the Dutch finite and post-verbal type. Corpus data (Doğruöz & Backus, 2009; Onar Valk, 2015) suggest this development, as Dutch Turks use subordinate clauses with an inflected verb more often than a control group in Turkey, and they more often placed these after the matrix verb. However, it is not so easy to say that this confirms the hypothesis. It may be that by sheer coincidence the lexical and pragmatic conditions which promote the use of finite and/or post-verbal subordinate clauses arose more often in the conversational data of the bilingual speakers. Or perhaps there was a subtle priming effect from one speaker who happened to use these structures a lot, triggering more use of these structures by the other speakers as well. If data are based on relatively few conversations, this can easily distort the overall quantitative analysis. The biggest problem, however, is that there is no way of knowing whether the results found for the participants, even if they suggest the expected change in Dutch Turkish, can actually be generalized to the wider community, since the number of speakers, numbers of hours of speech, and numbers of different conversational settings sampled is too low. The conclusions pertain to the group of participants in the particular setting in which they were studied. Generalization beyond these specific speakers and settings to the speech of the community in general can be suggested but is not warranted: essentially there would not be any evidence for it.

A single experimental study cannot remedy all of this, but studies can be set up to tackle these problems one by one. Onar Valk (2015) is an example. To test whether the generalizations suggested by the corpus findings could be confirmed for a much larger group of speakers, an experiment was set up in which speech was elicited from participants that was bound to contain a lot of subordinate clauses. Participants were given stretches of talk and then had to repeat them in a test setting. Each stretch contained two to four sentences, some of them containing subordinate clauses. The stretches were too long to fully commit to working memory, so the idea was that speakers would use their productive grammatical competence to form the repetitions, including the subordinate clauses. This set-up allowed the researcher to investigate the speech of more than 200 participants, both from the immigrant community and from a control group of monolinguals in Turkey, which significantly improves external validity compared to what can be done using recordings of spontaneous natural conversation. However, since repeating stretches of talk in a laboratory setting is not exactly daily communicative routine, ecological validity is compromised, as in all experimental work.

The results were encouraging. Differences between immigrant and control group were the same as was found with conversational data: Dutch Turks formed subordinate clauses significantly more often with the Dutch-style features (finite and post-verbal) than Turks in Turkey. The converging evidence suggests that both the external validity problem associated with a small database of naturalistic recordings and the ecological validity problem associated with the experimental distortion were minimal in practice, at least in this case. Note that we would not have known that if the researcher had only used one of the two methods: if she had only used naturalistic recordings of a small number of people we would not have known how generalizable the findings were; if she had only used the experimental study we would not have known to what degree the behavior in the lab setting corresponded to behavior in the wild.

In addition to the elicited imitation (also called 'sentence repetition') technique employed in Onar Valk's study, many other techniques have been used in HL studies. The most common one is some form of elicited production, in which participants are asked to narrate a particular story, often on the basis of a video clip or picture book they have been shown right before the recording. Quite famous are the so-called 'Frog Story' narratives, based on a picture book (Mayer, 1969) that is well suited to eliciting narratives that all but force the participant to use relatively complex tense marking and narrative structure (Polinsky, 2011). It is not a familiar story but one very frequently used in child language and language contact research. Elicitations like these are akin to the recording of natural conversations, since they elicit relatively free speech from the participant, but the speech is experimentally controlled in several ways. First of all, there is no dialogue involved, and in addition the speech that is produced is tightly controlled.

A variant of this involves no visual stimuli but simply asks participants to narrate a familiar story. Toribio (2004) asked Spanish speakers of Hispanic background in the US to narrate the story of Little Red Riding Hood in various modes (they were either instructed to speak monolingual Spanish or to use both languages if possible, leading to codeswitching). Clyne (1967) elicited comparable speech samples from speakers of Australian German by having them look at a small set of iconic photographs with particular relevance to the immigrant community.

Another version of this paradigm uses videos as elicitation material. This was applied in Moro (2016) and Irizarri van Suchtelen (2016). Irizarri van Suchtelen designed some of his own clips; other clips and videos were taken from existing collections. Applying a technique widely employed in typological work, short videos depicting just one action or event (such as acts of giving) were produced which were shown to participants with the task to say what is depicted or acted. Such videos can be manipulated to elicit particular grammatical structures. The advantage of this approach is that it can be used with many different languages, ultimately yielding a

comparative database. Obviously, if the research focus shifts over the years to other kinds of constructions, new videos will have to be designed. Depending on the cultural background of the participants, there may be some unease when an adult has to describe clips that look childish or culturally inappropriate. The video of the Sandman which second language researchers often use to elicit progressive aspect (Los & Starren, 2013), for example, was excluded from their collection because participants in the original homeland of the HL could not relate to the video.

Gullberg, Indefrey, and Muysken (2009) list several other techniques that are often used in psycholinguistic research and that can fruitfully be used in research on HLs. We will discuss one more study here, one that used priming to elicit particular constructions. Priming can be used to measure the influence that a dominant language may have on an HL. Kootstra and Şahin (2018) confronted Papiamentu speakers in Aruba and the Netherlands with a series of pictures depicting acts of transfer, and the subjects were asked to describe what they saw in Papiamentu. However, before each picture the subjects were also exposed to a recorded Dutch sentence involving either the Double Object construction (with the order Indirect Object – Direct Object, as in ‘gave Mary a book’) or the Prepositional Object Construction (with the reverse order, as in ‘gave a book to Mary’). The study showed that speakers were influenced (‘primed’) in their Papiamentu picture descriptions by the particular Dutch pattern they had heard. Without going into details, this indicates that the prepositional pattern, originally marginal in Papiamentu, is gaining ground in the language, i.e. it is better entrenched in the mental grammars of the individual speakers. This is just another way of demonstrating contact-induced change.

Finally, the internet offers many new possibilities for experiments which can be done at a distance, through programs such as the Amazon Mechanical Turk, a website where people can perform tasks, such as participating in an experiment, often for a small payment.

### 5.3.5 Judgment tasks

The discipline of linguistics has long made use of judgment tasks, in which participants rate stimulus items. The method used to be associated mostly with generative linguistics, for which it has always been of prime importance to know whether a particular grammatical structure is grammatical (‘well-formed’) or not (‘ill-formed’). Asking about grammatical structure directly would require the use of meta-language that doesn’t come naturally to speakers without extensive training (‘noun’, ‘direct object’, ‘inflected verb’), so the practice is to use sentences that exhibit the pivotal structure. The question asked of participants in such studies is often a binary one: is this sentence grammatical or not?

Such binary grammaticality judgment tasks are one variant of a family of tasks which all rely on the notion of acceptability, or conventionality. They ask participants to rate the degree to which the stimulus item asked about is in general use, or conventional, normal, acceptable. As acceptability is a matter of degree, participants are often asked to judge the items on a scale (often in the form of a five- or seven-point Likert scale), rather than forcing a binary choice between 'acceptable' and 'not acceptable'. An alternative scale is used in *Magnitude Estimation* tasks, in which participants construct their own scale.

In minority language studies, the use of such tasks has long been rare, but the last decade has seen an upsurge. The historical non-use of acceptability tasks is probably related to the sociolinguistic nature of much of the work on these languages, and sociolinguistics has long distrusted judgment tasks, perceiving them as distinguishing between right and wrong uses of language. Sociolinguistics tends to reject such terms, as what is 'incorrect' to some is just the normal way of speaking to others (usually the non-elite). However, when acceptability tasks are framed more in terms of what participants perceive to be common or rare in their own language or in the language they hear around them, they are eminently suited to studies of HLs. Specifically, they make it possible to tackle the problem of data scarcity. If a structure doesn't occur often in naturalistic recordings, it is hard to say anything about it. If, for instance, you suspect that a particular HL is replacing its preference for a double object construction with a prepositional structure, and you want to investigate this using spontaneous conversation, you are dependent on speakers using enough utterances in which they potentially have the choice between the two structures. A judgment task is one way around that problem. You can ask a large pool of suitable participants to look at a sufficient number of items, representing both structural options, and ask them to judge how common they are.

Important here is the difference between spoken and written tasks and well as the reliance on time constraints. For instance, Kupisch (e.g. 2013, 2014) and colleagues have used acceptability judgment tasks when comparing second language learners and heritage speakers. They used both written and spoken tasks to make sure heritage speakers are not disadvantaged (they could all read but it may be suspected that they were less proficient readers than the control groups of majority language speakers and L2ers, given lack of occasions to practise written language). Moreover, it was suspected that second language learners may have an advantage over the heritage speakers because they are used to the kind of task (written + acceptability judgments) and because they can access explicit or metalinguistic knowledge, while heritage speakers are less likely to have such knowledge. This is why the tasks were done under time pressure (limited response time).

As usual, there are various issues to deal with, though. Obviously, there are problems with ecological validity, since we are not asked in ordinary life to give

these kinds of judgments. More specific to judgment tasks, however, is how to deal with potential problems of internal validity. We would like to conclude on the basis of judgments whether or not a particular structure is common in the HL. However, since the evidence is indirect, it is not immediately obvious whether the judgments reflect actual frequency. Should evidence from judgment tasks from a random and sufficiently large sample of participants converge, however, with that from naturalistic recordings from a small group of people from the same population, the conclusions drawn will be more robust.

Perhaps the most difficult aspect of the design is to decide what to ask the participants to do exactly. In most cases, the research question will be which option is used most when members of the community need to convey something that can conventionally be voiced through either option. Therefore, a logical option could be to ask participants to rate the frequency with which the structure is used in the language they hear spoken around them day by day, by others and by themselves. However, this question can be asked in various ways:

- How often do you use this sentence?
- How often do you hear this sentence?
- How common is this sentence?

A separate issue is that the structures have to be given in the form of actual utterances, since otherwise people would have to be given abstract grammatical patterns to judge. This would require the technical linguistic terminology most people have no knowledge of. However, since judgments of those abstract patterns is what we are really after, not judgments of the actual sentences, we need to make sure that participants do not end up judging the semantics of the stimulus items. Let's say you want to know about the preference for a double object construction or a prepositional construction and you give participants the following two sentences to judge:

- I gave him the keys
- The conductor gave the memorial plaque to the second violin player

It is likely that the first sentence would be rated as more common, but the effect would presumably be due to the more common words used in that item (the personal pronoun, the everyday word 'the keys', and the fact that keys are often given by one person to another). It would, therefore, be erroneous to conclude that the Double Object Construction is more common than the prepositional construction. This effect can be countered in two ways. One is to make sure the stimulus items, at least the 'critical' ones (in order to keep participants from focusing on the construction in focus too much, there will probably be a number of 'filler' items as well), do not differ from each other as much as the ones in the

hypothetical example above, or have the same kind of variation distributed among the stimulus items for both variants. The other measure one can take is to instruct the participants as carefully as possible not to focus on the meaning of the utterances. However, the fact that you cannot use technical meta-linguistic vocabulary remains a formidable stumbling block.

The study by Onar Valk (2015) mentioned above also used judgment tasks, to further see whether evidence would converge regarding subordinate clauses in Dutch Turkish. Recall that analyses of natural conversational speech as well as of repetitions in an elicited imitation task suggested that Dutch Turks were much more likely than monolinguals in Turkey to use subordination structures that resemble the equivalent structures in Dutch (building finite and post-verbal subordinate clauses), rather than the structures considered more typical for Turkish (non-finite and pre-verbal), which were indeed used more by the monolingual control group in both kinds of data. Onar Valk also administered a judgment task to similar groups of participants, in which they were asked to what extent they heard the sentences included as stimulus items in the language use of those around them. The results confirmed the other findings, at least partially. Dutch Turks accepted the Dutch-style subordinate clauses significantly more often than monolinguals. However, they also accepted the other structures, the ones considered conventional for Turkish, to the same extent. Combined, these results suggest that in the immigration context, i.e. as an HL, Turkish has indeed changed its way of forming subordinate clauses. The change has propagated to some extent: what used to be a pattern used with lesser frequency (if the monolingual data indeed allow us to extrapolate that its characteristics are similar to those of the variety of Turkish brought into the immigration contexts decades ago – a problem of external validity) is now the numerically dominant pattern. However, the high judgment ratings for the conventional Turkish structures suggest that there is a certain disconnect between use and knowledge: these structures are still mastered well, but not used a lot. Whether the less frequent use will mean less entrenched knowledge after some time has passed is hard to say, and impossible to prove without some form of longitudinal study. On the other hand, the similarity in scores across various types of subordination brings to mind an observation by Polinsky (2006, p. 196) regarding grammaticality judgments from HL speakers:

“Unlike full language speakers, speakers of a reduced language cannot be accurately tested for acceptability judgments. If asked “Can you say...?” or “Is the following correct?” speakers usually accept what they are offered, unless some very basic principle of grammar is violated. Likewise, these speakers’ decisions on forced choice seem almost random.”

Basic insecurity about the conventions would also bring out the findings Onar Valk (2015) reported. This is not the place to further interpret these findings, but you should take away from this that it may be useful to combine two or more methods in the same study, since they may shed light on different aspects of the phenomenon. Sometimes HL speakers perform better in their judgments than in their production (for instance van Osch & Sleeman, 2018). In addition to the Onar Valk (2015) study, Hopp and Putnam (2015) also make use of acceptability studies when working with a moribund HL German community.

A more sophisticated approach to judgment tasks is Magnitude Estimation, widely used in psychology and introduced into psycholinguistics and sociolinguistics by Bard, Robertson, & Sorace (1996). Participants in a magnitude estimation experiment judge items using their own rating scale. In a standard example, often used in the instruction to participants, a line of certain length is shown and participants are asked to give this line a number for its length. Then a next line is shown, either longer or shorter than the previous one, and participants score that line as well. The difference in scores reflects the participant's estimate of the magnitude of the difference between the two lines. Similarly, if a correctness or acceptability score is given to a particular linguistic example, the next example will be judged in relation to it, and given a score that reflects the degree to which that new example is better or worse, or judged to be more or less frequent.

#### 5.4 Summary and conclusion: Which method to choose?

The great advantage of all experimental techniques, including judgment tasks, is that specific linguistic patterns can be targeted.

Given the great diversity of research questions, the different ways in which the three kinds of validity can be weighed, and the dependence on available resources, there is no single best method to recommend. Each researcher will have to choose the most optimal method, or combination of methods, to serve his or her purposes. In choosing your research method you should answer the following questions for yourself:

##### *What am I going to study?*

It makes a difference whether your study has wide scope, aiming to explore a whole language or variety, or a much more specific scope, for instance the degree to which HL speakers use Construction A or Construction B. Consider questions such as the following:



- Does your study aim to provide a general description of a language variety, perhaps compared to a baseline?
- Does it need to describe the sociolinguistics of the community, such as the language choice patterns it displays?
- Does it need to document variation within the variety?
- Does it target particular linguistic phenomena (whether syntactic, lexical, phonological or discursive in nature)?
- If so, are these phenomena rare or frequent in every day speech?
- Can they easily be elicited, for example with visual stimuli?

### *Who am I going to study?*

It also makes a difference whether participants for your study are easy to find, and whether your research questions make it important that they have particular background characteristics. Again, the optimal choice depends on your specific research questions. In general, questions regarding the participants include:

- Do the speakers you will be working with have to speak the language on an everyday basis?
- Should they have a good command of the HL, or are you actually focusing on a range of proficiency levels?
- Should they be bilingual in the HL and the majority language?
- Should they be literate in the HL?
- Should you impose particular demographic criteria on participant selection, such as an age range, a gender requirement, membership of a particular generation, a social grouping or an ethnic group, etc.?
- Should participants be able to participate in experiments, i.e. do they have to have some experience in test taking, and can they move at all (e.g. to a university lab)?
- Are there any ethical concerns one should be aware of?

### *How much data do I need?*

Finally, it is almost never the case that you can do everything you would like to do. Time is limited, participants cannot be taxed too much, and doing research costs money. Research questions determine what kind of design your study will have, but within those limits, there are still choices to be made, and they require consideration of questions such as the following:

- How much time and funding do you have at your disposal?
- If a corpus of spoken language material is needed, how much time can you allocate to data collection, transcription and annotation?
- Do you have access to support staff, such as assistants who can help with transcription?
- Do you have funding to pay for participants' time investment and travel costs, or for a token of gratitude?
- How much time do you need to find participants?
- Are participants all in the same place or are they dispersed over several localities? If you need baseline data for an immigrant group, are you able to travel to the place where you can get those data (i.e. the homeland)?
- How easy will it be to find participants?
- If you need experimental data, do you have access to the requisite lab facilities? This is not much of an issue if participants carry out a paper-and-pencil task, but if reaction times need to be measured, or audiovisual responses need to be recorded accurately, you may need access to a lab.

Altogether, it is often best to combine different methods, if at all possible. As we have shown above, often this will allow you to say something about different aspects of a phenomenon, provide more robust conclusions, or to show that things are more complicated than it seemed before you did your study. Those kinds of surprises drive science forward, so a falsified hypothesis or an unexpected result can be a good thing.



# Studying variability in heritage language speaker populations and the base line

## 6.1 Introduction

Most researchers who study heritage speakers are interested in linguistic variability and change, as was already discussed in Chapter 3. Language variability and change can be studied from different angles. The two perspectives explored most often in heritage language studies are the proficiency frame and the language change frame or perspective. In the proficiency frame, the question is to what extent factors of the acquisition scenario such as age of onset of the language(s), the quantity and quality of the input, and competition from other language(s) affect language proficiency. In the language change perspective, the central question is to what extent we should interpret variation and change as an effect of language contact. Both approaches investigate which aspects of language are susceptible to change in the heritage setting, but this chapter takes the proficiency perspective as its starting point.

In order to establish change or variation one needs to have a baseline comparison. For example, we can only argue that age of onset of acquisition is a factor in determining language proficiency of HL speakers if we have a baseline group who is similar in as many aspects as possible to the HL speakers except for this factor. If we want to argue that there is contact-induced language change we need information on the state of the language before the alleged change took place and we need to exclude the option of non-contact induced language change as much as we can.

This chapter is concerned with two central questions connected to finding a suitable baseline. The first is a WHO-question: Who should we compare heritage speakers to, in short: who provides the baseline in comparison to which we can establish that change has occurred? The second is a HOW-question: How can we compare heritage speakers to other speakers of the same language in a precise way? While in the previous chapter we discussed research methodologies in general, comparing different groups of speakers of the HL is a key methodological issue and therefore this aspect of research design is discussed in more detail here. Empirical results obtained then follow in Chapter 7 and theoretical models to account for the findings in Chapter 8.

The first part of this chapter is devoted to the WHO-question. Section 6.2 introduces six groups of possible baseline speakers. Sections 6.3–6.6 zoom in on specific aspects of the heritage acquisition scenario that need to be taken into account when considering the baseline. The first set of factors concern the timing, the quality and the quantity of the input and effects of differences in these factors on HL acquisition (6.3). The second set concerns language use (or ‘output’ and ‘production’; 6.4.) The third set concerns the social embedding of the language (6.5) and the fourth set of factors concerns stylization, register variation, dialect variation and social variation (6.6). Finally, Section 6.7 discusses how to assess linguistic profiles in different learner groups to deal with the inter-speaker variation produced by the factors discussed in Sections 6.3–6.6, so as to make a meaningful comparison between speakers possible. Section 6.8 concludes the chapter. Kupisch (2018) considers four aspects of bilingual acquisition, including cross-linguistic influence, the role of age of onset of bilingual input, input and language dominance, distance to the homeland. Cross-linguistic influence will be considered in Chapter 7, while the other three aspects will be discussed below.

**Main goals of this chapter:**

To give an overview of possible baseline groups.

To give an overview of individual background factors that need to be taken into account when comparing heritage speakers to other learner groups and language users.

To give an overview of tests to assess linguistic profiles including proficiency, language attitude and language use.

## 6.2 Establishing the baseline and the problem of monolingual bias

The language change and variation perspective on the study of HL hinges on finding some kind of change over time between two varieties of a language. A key issue is that of the baseline, the state of the language before the change took place. What would the language have been like if it had not developed in a heritage setting? The acquisition frame asks if the acquisition outcome would have been different in a different acquisition scenario. These questions are closely related and therefore we discuss the selection of baselines for the two approaches together. This section discusses six possible ways to do comparative research with a baseline and the advantages and disadvantages of these approaches. A central problem is what may be called the ‘monolingual bias’ (Ortega, 2014, 2016). It is often difficult to compare bilingual speakers of any kind to monolinguals, since monolingual processing inherently has different properties from bilingual processing, the diffusion over domains of usage of a language (e.g. home language versus work language)

may crucially influence language performance, and grammars of bilinguals may be different due to differences in qualitative and quantitative input.

### 6.2.1 Standard language grammar

One possibility is to simply compare the language of heritage speakers to the constructions described in grammars of the standard language of the country of origin. For well described languages this option is quick and cheap. The risk, however, is that the linguistic descriptions may not reflect actual speech, that they may not capture individual variation present in homeland speakers, nor is it accurate if the speakers acquired a non-standard vernacular. Consequently supposed deviations from the hypothetical standard do not reflect what really happened.

### 6.2.2 Exchange students and other recently arrived native speakers

Non-heritage native speakers often form the baseline when researchers are interested in change due to the heritage context. A cost-efficient and practical baseline – if available – is the use of exchange students from the home country. For example, the speech of Chinese Americans may be compared to the speech of exchange students from China in the US. The speech of the exchange students will most likely be more natural than the speech described in grammar books. Heritage speakers and exchange students can undergo the same tests. This allows for comparability, and since groups of individuals are tested variation within the non-heritage native speaker group can be better captured. However, exchange students may not match the heritage speakers well for three reasons. They may be different with regard to their general level of education and schooling in the language; they may have different regional backgrounds than the heritage speakers; and it is possible that their language shows some effects of contact-induced change as well.

### 6.2.3 Transnational research design

A third possibility is to adopt a transnational research design: data gathering takes place both in the country of origin and in the one where the HL speakers live. A possible advantage of this method is that heritage speakers and homeland speakers can be matched regionally and socially. Like with exchange students as a control group, this design allows the researcher to test heritage speakers and homeland speakers in the same way, so comparability of the data is ensured. Because multiple individuals can be tested it also allows the researcher to capture variation in homeland speakers. Aalberse et al. (in press) compared two generations of Wenzhounese

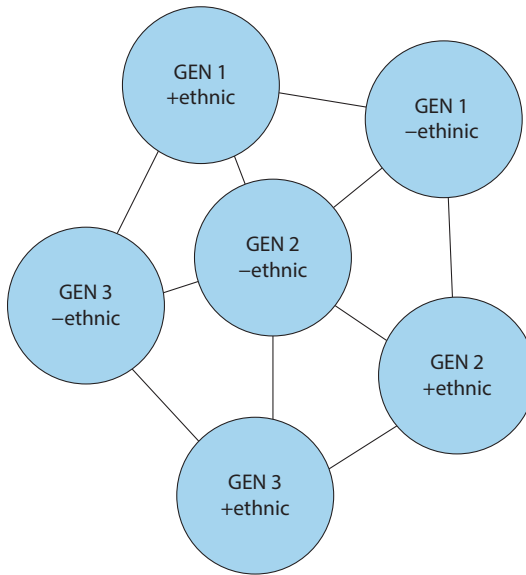
speakers in China and in the Netherlands; the difference they observed between the two generations in the Netherlands was also attested in China.

A first caveat is that test material may be interpreted differently by HL speakers and homeland speakers because of socio-cultural differences. Information on avoiding or minimizing cultural bias in designing test material can be found on the website for experimental field linguistics <<https://experimentalfieldlinguistics.wordpress.com>>. Our own research experiences underline the need for culturally appropriate material. Bob Borges (p.c.) reports that participants in Surinam were puzzled by a video they had to describe because it featured the sandman, who was utterly unfamiliar to them. Similarly, Wing Chee Chau (p.c.) noticed reluctance on the part of elderly Chinese participants when they were asked to describe cartoons.

Second, if migration took place some time ago, data gathered at the present moment in the country of origin may not reflect the language spoken at the point of departure well. For example, the Cantonese spoken in Hong Kong today is much more influenced by English than the Cantonese of 60 years ago (Li, 1999). When heritage speakers return to the country of their ancestors, they are sometimes told that they sound archaic (cf. Chau, 2011; Irizarri van Suchtelen, 2016). While differences between heritage speakers and home country speakers will often be due to cross-linguistic influence from the dominant language, it may also be the result of heritage speakers not having participated in changes in the language as spoken in the homeland (cf. Aalberse & Moro, 2014). For instance, a standard language may have become more dominant in the homeland in the intervening years since migration, making the comparison less neat.

Third, differences in literacy level between the heritage speakers and the homeland speakers may cause differences in how they speak (cf. Rothman, 2007; Pires & Rothman, 2009; Kupisch & Rothman, 2018; Bayram et al. 2017). Fourth, heritage speakers may have a longer history that sets them apart from homeland speakers. For example, many Moluccans lived in military barracks before they arrived in the Netherlands from Indonesia. These barracks were often situated on Java and contact with Javanese changed the Ambon Malay language of the Moluccans before they arrived in the Netherlands (affecting for example pronoun use). The Malay spoken in the barracks was often referred to as Tangsi Malay (literally barracks Malay), suggesting it had diverged noticeably from the original Ambon variety. Tangsi Malay is not documented and therefore a comparison between Heritage Ambon Malay and the language spoken on Ambon is complicated: the effect of the barracks period is not clear (Adelaar & Prentice, 1996; Bos, 1977; Jonker, 2009; Moro, 2016; Tahitu, 1989). Finally, transnational research may simply be difficult to organize or too expensive.

## 6.2.4 Vary subject populations



**Figure 6.1** Connections between different subject populations

A fourth possibility is to only study the HL in the bilingual context, but to distinguish between sub-groups. The chosen design will partially be determined by the specific characteristics of the community, such as how many generations ago bilingualism started.

The illustration in Figure 6.1 reflects the search for connections between different subject populations. Dimensions for the variation could be generation (first/second/third), proficiency (high proficiency/low proficiency), and length of residence (long term residents/recent arrivals) or ethnic orientation (cf. Nagy, Chocie, & Hoffman, 2014). Questions can be: to what extent does the score on one variable predict the other? Do low proficient speakers have low word per minute rate, as in the work of Polinsky (2008b)? Does a low word per minute rate predict aspects of grammar? Does generation predict grammatical patterns? And do scores on one aspect of grammar predict other grammatical features? The advantages of this approach are that it can tease apart effects on different dimensions and that it is precise. The disadvantages are practical: to make this approach work many respondents are needed and much background information on the respondents is required. An example of a study that takes all this into account is Moro (2016).

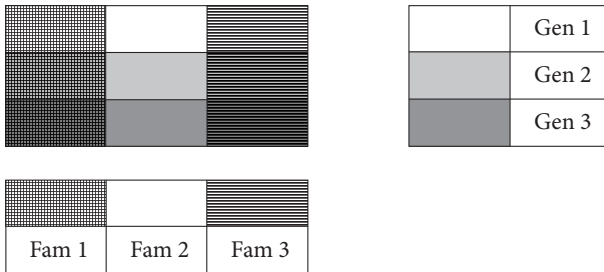
The options described in 6.2.3 and 6.2.4 are by far the most common in contact linguistics in general. On the other hand, there is also much work that does



not include a baseline at all, either because of methodological choices that were made in older days or because there is no baseline population at all. The modern HL literature injects the field with new methodological rigor in this respect, and its lessons should be taken on board where it is possible.

### 6.2.5 Cross-generational family studies

A fifth possibility is to compare differences within and between families by testing different heritage families as illustrated below in Figure 6.2.

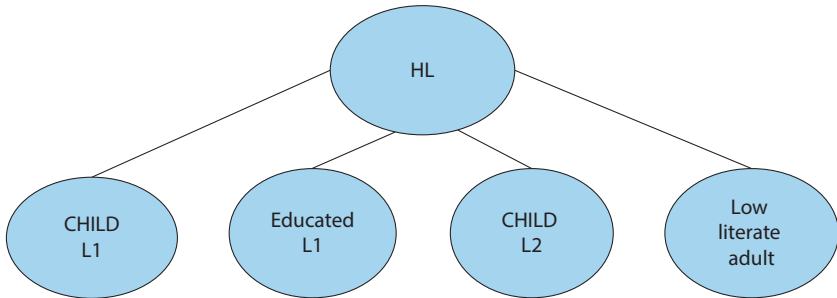


**Figure 6.2** Cross-generational family studies, illustrated here with three generations. The oldest generation is reflected by the lightest colour, the youngest generation by the darkest colour. Family membership is symbolized by different patterns, illustrated here for three families. The question in family research is to what extent output in each of nine cells is decided by family membership (different shades) and to what extent by generation (relative darkness)

This approach helps tease apart the effects of being from a certain generation from those of being from a certain family. HL communities are sometimes very small, and this promotes the crystallization of family-based conventions. Chen and Shirai (2010) and Chen (2012) show that Chinese HL families sometimes have their own style of aspect marking. In their studies, the number of aspect markers children use could to a large extent be predicted by aspect use of the parents. If both family style and generation can affect language output, a comparison between and within families can show at what levels speakers from the same generation line up and on what level speakers from the same families do. Social and geographical background of speakers will also be better matched when they are from the same families. Laleko (2010, p. 249) illustrates how the type of input received by heritage speakers from their parents (and more generally from the heritage community) may not be analogous to the input available to non-heritage native speakers. Testing the output of the parents of heritage speakers will enable the researcher to tease apart changes that result from the heritage acquisition scenario on the one hand and changes induced by the input of the first generation on the other hand (cf. Pires

& Rothman, 2009; Domínguez, Hicks, & Slabakova, in press). Problems with this approach are practical. Members of the family may be scattered over the country or even over the world; (grand) parents may have returned to the home country and not all members may want to participate in the research.

### 6.2.6 Multiple baselines



**Figure 6.3** Multiple baselines that can be used

Since so many factors play a role in the output of HL speakers, it is also possible to compare them with different groups with the same language combination and see in which domains they match those groups. Figure 6.3 list some possible comparisons, in which ‘Child L1’ refers to monolingual first language acquisition, ‘Educated L1’ to adult L1 speakers who are educated and literate in the L1, ‘Child L2’ to acquisition of a second language after the age of 4 but before puberty (see Blom, Polišenská, & Weerman, 2007 and Schwartz, 2006), and ‘Low literate adults’ to monolingual adults who have received little or no schooling. By comparing these different groups, one may be better able to disentangle effects from age, input and schooling.

Pires and Rothman (2009), for example, compared heritage speakers with adult and child homeland speakers and with speakers with low and high educational backgrounds. One could also compare heritage speakers of Spanish with English as L1 to child L2 learners of Spanish with English as L1 in order to investigate the effects of age of onset and of limited input. Ortega and Byrnes (2010) suggest that new statistical methods enable us to depart from group comparisons altogether and focus on individual variation instead. Instead of creating different groups a priori, let’s say heritage speakers versus non-heritage speakers, we can analyse the data of all speakers and investigate which factors cause speakers to be similar or dissimilar. It may be the case that the distinction heritage versus non heritage is not the crucial variable that explains variation between speakers.

### 6.2.7 Bilingual baselines

Finally, one can compare bilinguals to other bilinguals. Rather than comparing bilingual heritage speakers to monolingual native speakers, Kupisch (2013) compares them to bilinguals who have the same languages but live in a place where what is the HL for the first group is the majority language. The idea behind this comparison is that if bilingual minority language speakers perform differently from monolingual native speakers this can be for two reasons: different knowledge of the language (a *representational* difference) or different processing capacities (a difference in *production*) (see Chapter 9 for more on processing). Bilingual speakers have a larger body of knowledge than monolinguals and this may slow down their processing; in addition, their speech may show effects from interference. By comparing speakers from bi-national families in two different settings, for example speakers from French-German families in France and in Germany, one can better distinguish the general effects of bilingualism (its processing effects) and differences in proficiency. For example, in both cases speakers will have a larger body of knowledge than monolingual speakers of French and both store vocabulary and rules from the same two languages and may both encounter competition from both languages. Yet, speakers from these groups may differ in the proficiency they have in each language. The speakers of bi-national French-German families in France will on average have a higher command of French than the speakers of bi-national German-French families who live in Germany. The idea is that it is fairer to speakers of heritage French in Germany to compare them to German-French bilinguals in France, because unlike monolingual native speakers of French, they share similar processing and interference challenges.

The bilingual baseline approach is valuable in studies that adopt the proficiency frame, because it enables a fairer judgment of factors that are indeed proficiency related: results will be relatively unpolluted by processing effects. From the perspective of the language change frame, the approach would need to include assessment of monolingual native speakers, because in both cases we would want to know whether their language use is different from that of monolinguals. Although processing constraints and preferences are different from proficiency as causes of variable production, both affect language use and thus in turn they can both potentially affect the direction of language change.

### 6.2.8 Summary

We looked at seven possible types of baselines and ways to control for undesired variation. The discussion was divided into single and more varied baselines, in which various groups are compared to the HL speakers and to one another.

The single baselines include (1) grammar books, (2) exchange students and (3) homeland speakers. The varied baseline includes looking at different kinds of immigrant speakers: different generations and different ages in (4), members of the same families across generations in (5), various monolingual baselines including illiterate speakers and speakers of different dialects in (6), and different bilingual speakers of the same language pair in different social circumstances in (7). The selection of baselines and the demarcation of the research group are difficult because many factors cause individual variation both in monolingual and bilingual groups. In the next section we will discuss factors that lead to individual variation.

### 6.3 Factors in individual variation in the acquisition perspective: Timing, quality and quantity of the input

A number of factors related to the timing, the quality and the quantity of the input may account for individual variation. We will discuss them briefly one by one.

#### *Age of onset of bilingualism*

The age of onset of the second language potentially affects proficiency in both the first language(s) and the second language. Montrul (2008) finds support for this in a large number of studies on HL acquisition and second language acquisition. Further support comes from studies on child L2 acquisition, and on international adoptees. Montrul proposes a hierarchy, from international adoptees > simultaneous bilinguals > sequential bilinguals > child L2 > adult L2. Speakers low on the hierarchy have relatively low L1 retention and high L2 success, whereas speakers on the right have solid proficiency in the L1 and are less likely to reach high proficiency in the L2. Echoing the often-proposed critical period for L2 acquisition, Montrul proposes a critical period for language loss. The earlier exposure to the L1 stops, the more severe language loss will be. After puberty, any language loss should be limited primarily to “usage effects”, such as lexical retrieval and hesitations.

An important distinction is that between simultaneous and sequential bilingualism (the latter sometimes subdivided in early sequential bilinguals and child L2 learners). The general prediction following from Montrul’s work is that sequential bilinguals, who speak only the HL in the first years of their lives, acquire and retain the HL to a larger extent than simultaneous bilinguals, whose HL is in competition with the dominant language from birth on. However, Kupisch (2013) finds that simultaneous bilinguals do not necessarily end up with more divergent grammars than sequential bilinguals. Kupisch (2018), surveying the results from different studies, likewise finds little evidence in support of age of onset effects.

We may conclude that if there is a difference between the two groups of early bilingual speakers it will be a probabilistic difference. Sequential bilinguals may be more likely to reach monolingual-like performance in the HL, but simultaneous bilinguals may reach the same result. Age effects for the level of L1 retention are only reported for speakers who became bilinguals in childhood. Past puberty, studies do not find age effects (see Hyltenstam et al., 2009 for an overview). Schmid (2002), for example, investigates participants who became bilingual during the age range 14–36, and finds that L1 maintenance did not vary as a function of age. Instead, it was correlated with other factors, in her case the attitudes caused by ethnic persecution.

Related to the difference between simultaneous and sequential bilinguals are the age of arrival, the parental native languages, and whether the household one grows up in is monolingual or multilingual. Kupisch (2018) also stresses the importance of the nature of the input in determining the outcomes of HL acquisition. We discuss the factors determining the nature of the input below.

### *Time spent in the heritage country during childhood*

The quantity and nature of the input of the HL is also affected by the time spent, if any, in the heritage country, for example during vacations. Lein, Kupisch, and van de Weijer (2016) compared voice onset times in French-German bilinguals in Germany and France. All speakers are simultaneous bilinguals, but they do not perform the same way. The speakers in Germany were more target-like in their weaker language than the speakers in France. Lein et al. offer an explanation in terms of complexity to explain this difference, but also consider the quantity of exposure to the HL during childhood. The speakers from Germany, on average, spent more time in France during their childhood than the speakers from France spent in Germany. The speakers from France, by contrast, spent much more time in Germany during adulthood (since they moved there), than the speakers from Germany spent in France. Lein et al. (2016) suggest that this may mean that time spent in the heritage country during adulthood cannot compensate for the smaller amount of time spent in the heritage country during childhood. That interpretation fits with observations in Flores & Rato (2016) who find that age of emigration is the best predictor for having a monolingual-like accent while prolonged stays back in the heritage country in adulthood, has no significant effect on perceived global accent. Lein et al. (2016, pp. 744–745) remark that if more exposure in childhood is indeed the relevant explanation that this does not “speak directly to the “earlier-is-better” view (e.g. Abrahamsson & Hyltenstam, 2009), because all speakers had been exposed to both languages from birth, but indirectly in the sense that relatively more input in the HL is more beneficial early in life than it is later in life.

### *Mono- versus multilingual households*

In some cases, heritage speakers are genuinely sequential bilingual speakers because they moved to the host country a few years after they were born and both parents have the same L1, ensuring that the societal language was really acquired later than the home language. However, there are also speakers who are classified as sequential bilinguals because both of their parents speak the HL, but who are also simultaneous bilinguals to an extent, because, having been born in the host country, they were exposed to the societally dominant language outside the home from birth. This exposure may have been of low intensity before school, especially if they have not attended day care, and therefore classification as sequential bilinguals may be more or less accurate.

In some studies, the opposition between multi- and monolingual households is the central factor (for example Irizarri van Suchtelen, 2016; Flores et al., 2017). Irizarri van Suchtelen found that speakers from mixed households were more innovative, and Flores et al. (2017) found children with more exposure to the HL at home (one language households) to show faster acquisition of the home language than children from mixed households. However, she also observed that speakers from mixed households eventually caught up. They only show a delay in comparison to children from one language households.

In other studies, the age of arrival in the host country is the main factor, but in most cases later arrival implies sequential bilingualism and early arrival simultaneous bilingualism. An example is Yeni-Komshian, Flege and Liu (2000), who examined pronunciation proficiency in Korean-English bilinguals in both languages. The pronunciation of L1 Korean of participants who arrived between the ages of 1 and 7 was distinctly accented, while those with ages of arrival between 12 and 23 were rated the same as monolinguals. Participants with ages of arrival between 1 and 9 pronounced English better than Korean, whereas the opposite was observed for participants with ages of arrival between 12 and 23. Flores and Rato (2016) similarly found age of emigration to be the only significant predictor for non-native accent.

In yet other studies the precise age of onset of the second language is used rather than the binary distinction between simultaneous and sequential bilingualism. For example, Bylund and Jarvis (2011) investigate event conceptualization in Spanish-Swedish bilinguals in Sweden. They find that the earlier speakers were exposed to Swedish, the more likely they were to construe events with an explicit endpoint in Spanish, making Spanish converge with Swedish conceptualization strategies.

### *Parental native language*

It is also important to take the quality of the input heritage speakers receive into account. One factor that potentially affects this quality is the manner in which the parents acquired the language. Whereas some parents speak an HL with their children that they also spoke with their own parents, this is by no means always the case. In Chapter 3 maintenance and shift were discussed in detail, and we saw how a minority language may be transmitted less and less to the next generations. However, patterns may be more complex. For example, Chinese parents who spoke a Wu-dialect with their own parents may speak Cantonese with their partner and Mandarin (cf. Li Wei & Zhu Hua, 2010) with their children, and Moluccan parents who spoke ‘Bahasa Tanah’ (language of the land, e.g. indigenous language) with their own parents, may speak the Moluccan lingua franca Ambon Malay with their children because this is the language shared by members of the community. Akan has become the community language of Ghanaian migrants in the Netherlands. Even those who would not have spoken Akan in Ghana, for example people from the North, will speak Akan, for example to project a Ghanaian identity (Margot van den Berg, p.c.). If parents are not native speakers of the language they use at home, this may affect the acquisition process of their children.

The complicated relationship between a parent’s first language and the HL they transmit to their children is reflected in the following quote by Mrs. Tse, a Chinese woman in Scotland, taken from Hancock (2014, p. 68):

*Hakka is our first language, English because of the children, Cantonese is used at the Chinese school, and we speak Putonghua to the kitchen staff. My daughter is going to dance classes and the teacher only speaks Putonghua, so I have to teach her that language.*

### *Parental language strategy and modes of speech*

Parents select various language choice strategies, as summarized in Lanza (1997) and Montrul (2008, p. 101). Some parents for example refuse to answer their children if they don’t use HL or they repeat an utterance made in the dominant language in the HL. Strategies like this increase the chance that the child will become an active bilingual rather than just an overhearer. Active use has a positive effect on proficiency. Indeed, in many countries there are websites that give advice to parents in bilingual families on how to deal with the complexities of language choice, school and educational demands, the linguistic behaviour between the spouses, and the preferences of children. Up until, 2010, the publisher Multilingual Matters published the *Bilingual Family Newsletter*, which is publicly available now. Parental language strategies may be influenced by what De Houwer (1999)

calls ‘impact belief’, the belief on the part of the parents that they may have some control over their children’s language use and acquisition.

### *Parental language use/language mode patterns outside the family*

Research on post-puberty attrition suggests that if immigrant parents get to use the L1 at work, this has a protective effect on the L1 (Schmid, 2007; Schmid & Dusseldorp, 2010). Schmid (2007) interprets this (limited) effect as an effect of language mode. When speakers habitually use their L1 in a monolingual mode, other languages are largely (though never completely) deactivated, making cross linguistic interference less likely. Therefore, the quality of the parental input could be affected by the language modes in which the parent(s) use the language outside the family. Kupisch (2018) analyzes the evidence that real or perceived distance to the homeland plays a role in determining the nature of the HL; this distance very likely affects parental language use strategies.

### *Caretaker background*

Parents are not always the main source of language input for children. Some Chinese heritage children in the Netherlands, for example, spent much time in Dutch guest families (Geense & Tsui, 2001, p. 94). Children also go to (bilingual) day care centers, they may grow up with their grandparents (in the heritage country or the country of the parents) and one of the parents may have a different language background. Nannies and babysitters also affect the type of input children receive. Montrul & Sánchez-Walker (2013) investigated the predictive value of different factors on Differential Object Marking (in which different types of objects, e.g. animate versus inanimate ones, are marked differently) in different groups of Spanish heritage speakers, distinguishing between speakers that tended to omit the Spanish object marker *a* (as in *yo veo la mujer* ‘I see the woman’) and speakers who did not (*yo veo a la mujer*). With regards to caretaker background they found that of the Spanish heritage speakers who performed like monolingual natives on a test, 83% had home care in Spanish. The children who tended to drop object marking, were much more likely to have received English-medium day care (42%). Li Wei (1994) found similar results for Chinese heritage speakers.

### *Sibling birth order*

Proficiency in the HL can differ greatly between siblings; this applies particularly when the community-wide shift is just starting. In many cases the oldest child speaks the HL most native-like (Shin, 2002). Once the oldest child enters school,



the dominant language may get introduced at home, inducing a shift in the quantity of input. Because of increased use of codeswitching there may be qualitative differences in the input as well. The higher the number of older siblings a child has, the larger the chance that (s)he will use the dominant language in daily speech. Sibling order thus affects language use, and therefore language input, and therefore language proficiency. Bridges and Hoff (2014) found that there was a significant difference between heritage speakers with and without older siblings.

#### 6.4 Speaker characteristics, language use and language output

There are also a number of factors relating to speaker characteristics and language use that may account for the individual variation we find in HLs. We will discuss them briefly one by one.

##### *Language use patterns*

Bohman et al. (2010) show that although language input is important for the first steps in acquisition, actual use of the language is an important predictor for ultimate attainment. In order to make compare speakers it is therefore important to know about their language use patterns. Collecting this information, however, is not an easy task. In research on bilingual acquisition parents are often asked to keep language diaries for their children, or their language use may be extensively observed and recorded. These options are not always available, however. Amount of exposure can affect the rate of acquisition (e.g., Gathercole & Thomas, 2009; Jia & Aaronson, 2003; Oller & Eilers, 2002; Montrul & Potowski, 2007; but see Gutiérrez-Clellen & Kreiter, 2003), while type of exposure can affect the rate of acquisition and/or the ultimate attainment level (e.g., Jia, Aaronson, & Wu, 2002; Place & Hoff, 2011; Scheele, Leseman, & Mayo, 2010; La Morgia, 2011). Unsworth (2013) introduces the notion of cumulative length of exposure, a measure intended to capture the sum of bilingual children's language exposure over time. She shows that for some domains of language both current amount of exposure and cumulative length of exposure were found to be significant predictors in predicting bilingual success. In a similar line Putnam and Sanchez (2013) argue that the key factor leading to changes and ultimate decay of an L1 heritage grammar is the degree of activating and processing of their L1 throughout the course of a heritage speaker's lifetime.

Gollan et al. (2015) report that the sheer number of different HL speakers with whom children interact in their youth affects their proficiency, independently of how much they use it. The authors hypothesize that this is because different speakers use different words, thereby increasing the vocabulary of the heritage speakers.

In addition, it may produce a contextual diversity effect: representations of words become more robust when they are used in more contexts. Taken together, both higher frequency of words in the input and higher variety in the contexts in which they are encountered makes them subsequently easier to retrieve. Speaking the language with different people may provide this greater variety in contexts (cf. Adelman, Brown, & Quesada, 2006).

### *Domains of use*

Some HLs are used only at home, whereas others are community languages spoken in shops, at work, in school, in church, during community events, or on the street, and used in the media (radio, TV, newspaper, Facebook). In immigrant contexts, second and further generation speakers often become dominant in the language of the country because this is the language used in school. However, there are many other sociolinguistic settings. The range of domains in which the HL is used, as sketched in Chapter 3, and thus the number of interlocutors the language is used with may be a factor that influences language attainment (Place & Hoff, 2011; Gollan et al., 2015), as is described in more detail in 6.5. If a language is used as a community language this may positively affect its use in other domains, such as in the media, further increasing rate and diversity of input. Knowing speakers who only know the HL (such as grandparents) also increases the chance of speaking like homeland speakers.

### *Language aptitude*

Bylund et al. (2010) and Bylund and Ramírez-Galán (2014) investigated the role of language aptitude in L1 retention in, respectively, prepubescent and post-puberty migrants. Bylund, Abrahamsson, and Hylstenstam (2010, p. 447) loosely describe language aptitude ‘as an innate, relatively fixed, talent to acquire and process language structure.’ Language aptitude can be measured, for example with the LLAMA Language Aptitude Test (Meara, 2005) or the Swansea Language Aptitude Test (Meara, Milton, & Lorenzo-Dus, 2003). Bylund et al. (2010) find that language aptitude positively correlates with high scores on a grammaticality judgment test for prepubescent bilinguals. While low language aptitude speakers have low GTJ scores if they became bilingual before puberty and contact with the L1 is limited, high language aptitude offers protection against L1 loss in this group, despite low L1 exposure. The quantity of exposure is thus a crucial factor for the low language aptitude group but not for those with high language aptitude scores. Language aptitude has no effects on the level of L1 retention for speakers who became bilingual after puberty.

## 6.5 Social embedding in the multilingual speech community and the larger society

The social embedding of bilingualism and the HL likewise may affect the input in the language and use of the language. We discuss some aspects of this social embedding below.

### *Schooling and literacy*

Many HLs are learnt in informal settings only. Various studies report that whether or not speakers have access to the HL in a school environment affects their linguistic outcomes. Pires and Rothman (2009) compare the use of inflected infinitives in heritage speakers of European Portuguese and of Brazilian Portuguese in the U.S. In European Portuguese the inflected infinitive is used in formal and informal speech but in Brazilian Portuguese it is used in formal registers only. Pires and Rothman (2009) find that European Portuguese heritage speakers use the inflected infinitive whereas the Brazilian heritage speakers do not. Apparently, there is nothing about the heritage acquisition scenario itself that hinders the acquisition of this feature; rather it is about the nature of the input. Heritage speakers who never use formal registers obviously cannot acquire features that only belong to that register. In this case, like monolingual speakers with low educational levels they do not acquire the inflected infinitive. Pires and Rothman (2009) refer to this outcome as ‘Missing input competence divergence’ which basically states that what is not in the input cannot be learned. While some aspects of language will be completely absent in informal registers, there are also distributional differences between grammatical constructions in the informal and formal register. Torres Cacoulios (2000) relates the increased use of progressives in heritage speakers in the US to the fact that heritage speakers use the language in informal registers only: informal registers show higher rates of progressive use and thus the restriction to the informal registers accelerates change. Note however that the overextension of progressives is widely attested in heritage speakers, the informal register may explain this partly, but other factors will probably also play a role; see Brown & Putnam (2015) and Moro (2016).

The role of schooling in the acquisition of the formal register or of features that are prevalent in formal registers is quite direct, but schooling can also affect parts of language that are readily available in the informal context. Kupisch and Rothman (2018) compare Italian and French heritage speakers on five aspects of language that are not necessarily connected to the formal register, but which still do seem to be affected by schooling. They investigate the same variables in two different groups, namely gender assignment, gender agreement, adjective

placement, article use with generic DP's and VOT in Italian and French heritage speakers in Germany. None of these features are limited to the formal domain. The French heritage speakers are schooled in their HL at the *Lycée Français*, whereas the Italian heritage speakers are schooled in German and only attend weekly classes in Italian from one teacher. They find that the French heritage speakers outperform the Italian heritage speakers in all five domains and hypothesize the difference may partly be related to schooling in the HL. Schooling increases access to a more standard-like variety of the language and this may increase the variety of interlocutors one uses the language with as well as the sheer amount of use. As we saw, these are factors that stimulate proficiency.

Although daily schooling in the HL such as in the *Lycée Français* is the most effective factor, weekly classes can be effective too, at least temporarily. Bylund and Díaz (2012) investigated two groups of 12th graders with Spanish as L1 and Swedish as L2. One group still attended weekly Spanish classes whereas the other group no longer did so because of conflicting class schedules. The two groups did not differ with regard to the total number of years of HL class attendance, age of arrival in Sweden, length of residence, or amount of L1 contact. The group with weekly Spanish classes did better on a grammaticality judgment test and a cloze test. The study suggests that HL classes can protect against attrition.

The idea that the level of schooling in the HLs correlates with relative vulnerability of the L1 is further supported by work by Bayram et al. (2017). They investigate four groups of Turkish bilinguals in Germany, who differed in degree of literacy. The authors argue that all speakers have the same mental representation of the Turkish passive, but that the level of L1 literacy had a positive effect on monolingual-like production of passives. This suggests that intergenerational transmission may be affected by literacy.

Classes, literacy and schooling in the HL thus all affect the quantity and the quality of language input and use. They may also affect metalinguistic awareness. Research by Tarone & Bigelow (2005) and Tarone, Bigelow, & Hansen (2009) has shown that language intake in illiterates is different than in literate speakers. Heritage speakers are usually literate in the dominant language, and in that sense unlike Tarone's monolingual participants, for instance because the HL is not a written language or because it played no role in the education of the heritage speaker. As literacy affects (meta)linguistic awareness, heritage speakers may process their language differently from heritage speakers who are literate in their HL (cf. Oller & Eilers, 2002; Benmamoun, Montrul & Polinsky, 2013a).

### *Language prestige and language ideology*

As noted in Chapter 3, language prestige plays an important role as it affects the degree to which a language is used (see also the special issue on this topic; Kasstan, Auer, & Salmons, 2018). Carreira (2004) found a difference in linguistic output between communities in which there were large numbers of Spanish speakers, a high density of foreign-born Latinos, and Spanish enjoyed commercial, social, and professional status, such as Miami, Florida, and communities such as San Antonio, Texas in which the Latino population is predominantly U.S. born, and Spanish overall has lower status. For a high prestige HL there may be fewer barriers for using it, and it will be more likely that there is support for the language in terms of available media, books, films etc., leading to more use and literacy and thus a boost for proficiency. Kupisch and Rothman (2018) contrasted a number of studies on five similar constructions and found that French speakers were closer to the monolingual norm than Italian speakers. Language ideology differs across these two languages. French has a strong focus on the national norm, whereas the Italian speech community is more open to (dialectal) variation. For more on the role of prestige and ideology, see Kasstan, Auer, & Salmons (2018).

### *Settlement patterns and immigrant networks*

Settlement arrangement can influence the extent to which HL speakers may interact with each other, and thereby affect the extent to which the language is used, as discussed in Chapter 3. Whereas some immigrants live together in particular sections of towns and enjoy mainstays of heritage culture such as community shops, other immigrants may be rather isolated. For example, in the Netherlands Chinese migrants who arrived before the 1990's very often live outside the larger cities (Extra et al., 2002; Geense & Tsui, 2001). Almost every town in the Netherlands is inhabited by at least one Chinese family, often the owners of a family restaurant. Chau (2011) reports that the Amsterdam-based heritage speakers of Cantonese in her study watched Cantonese soaps and listened to Cantopop because they could share their experiences with their Cantonese speaking peers living in the neighbourhood. The Cantonese speakers in the smaller town of Venlo, on the other hand, did not have many Chinese peers, and they preferred Dutch soaps and music that was not Chinese. The ethnic presence in their social network thus influenced their media choices, which in turn affected language input and language use, and therefore language proficiency. Li Wei (1994, p. 182) shows that those British born Chinese who were members of the True Jesus Church had a higher proficiency in Chinese and mixed Chinese and English more than others. Reports on language use of Ambon Malay immigrants in the Netherlands show a relation

between living in the *wijken* (neighbourhoods for Moluccans only) and migrants who lived in Dutch neighbourhoods (Veenman, 1994, 2001; Gijsberts & Dagevos, 2005; Moro, 2016). Finally, Domínguez (2013) shows that the finding that Cubans in Miami use more null and post-verbal subjects than Cuban homeland speakers may be motivated by the contact with non-Caribbean variants of Spanish that comes with living in multicultural Miami.

### *Superdiversity*

The notion of superdiversity was developed by Vertovec (2007, 2010) to refer to the fact that in many contemporary urban settings many different linguistic and cultural backgrounds come together, and therefore many different communicative and behavioural patterns co-exist and interact. As a result, an HL speaker may be confronted with many competing norms for how the HL is spoken. Chileans living outside Chile, for example, may be exposed to the conventions of Chilean Spanish, but also those of Spain and other Spanish-speaking countries such as Argentina or Mexico. In addition, all of these norms may be shifting over time. As Hancock's (2014) quote in 6.3.3 of a Chinese first-generation immigrant to Scotland shows, norms can be multi-layered. Wong Filmore (2011) illustrates a case of heritage Korean where a norm is imposed that nobody speaks at home.

### *Additive versus subtractive bilingualism*

The terms additive and subtractive bilingualism were coined by Lambert (1975, 1977, 1981) to contrast two frames in which bilingual situations can be experienced. The difference is caused by the value that is attached to bilingualism by the dominant society. In the additive frame the acquisition of another language is seen as an addition, as something extra and of worth. In contrast, the subtractive frame pictures a 'form of bilingualism experienced by ethnolinguistic minority groups who, because of national educational policies and social pressures of various sorts, feel forced to put aside or subtract out their ethnic languages for a more necessary and prestigious national language' (Lambert, 1981, p. 12).

Some studies suggest that subtractive bilingualism negatively affects proficiency in both the first language (the HL) and the dominant language (Wong Filmore, 1991). Many reports on the rapid disappearance of HLs come from the United States and one possible reason for this may be that US attitudes to bilingualism often favour the subtractive bilingualism frame. Many heritage speakers in the US feel a pressure to shift to English completely and this negatively affects HL proficiency.

*Social class, gender, age, geographic background, register*

As also pointed out in Chapter 3, when people identify with a particular gender or age group this may affect the way they speak. Some researchers investigate whether and how heritage speakers use the HL for identity work (see for example Nagy, 2015). Lynch (2014) notes that Spanish heritage speakers in the US often lack a speech community. Other groups of heritage speakers such as Norwegian speakers in the US (see Bondi Johanessen, 2018) have always formed a tight knit network. One effect of being in a tight-knit community is that you can develop special conventions for making use of both languages. Li (2011) reports examples of 'translanguaging' in which Chinese heritage speakers refer to themselves as 'white collar dogs' in Chinese. Since they usually work in white collar jobs this label is partially appropriate but it receives extra meaning because the words 'white collar dog' is pronounced as *bai ling gou* (白領狗), which sounds like 'bilingual.' We follow Li (2011) in interpreting translanguaging as a specific kind of bilingual language use (or 'play' in this case), different from code-mixing. Whereas code-mixing combines lexical units (or larger) from two or more languages into bilingual utterances or discourse, translanguaging conveys social meanings associated with different languages in a multilingual setting, often combining resources in playful ways. Translanguaging is languaging (see Chapter 3) with resources from more than one language. It integrates aspects of both languages into new multi-layered elements.

One characteristic of HLs often noted by researchers is the absence of register or stylistic variation. HLs are sometimes derogatorily referred to as 'kitchen varieties' (Polinsky & Kagan, 2007) that lack the repertoire to convey social signals, limited to only the informal everyday vernacular. Manosuthikit (2013) shows, however, that heritage Burmese speakers have developed a new system of terms of address combining elements from Burmese and English. Manosuthikit (2013) reports, for example, on a brother and a sister who use 'bitch' as a gender-neutral intimate term of address for each other. The development is motivated by Burmese, the lexical selection by English. Note that the original feminine association of 'bitch' in English plays no role. This practice could lead to a new emergent style, similar to the playful bilingual style of Chinese-English bilinguals Li (2011) describes.

Independent of the question of whether heritage speakers in general show stylistic variation, it is very important to take stylistic variation into account in analysing the speech of the first generation. We know that languages that function as the dominant language are affected by aspects such as age and gender and social class. Because heritage speakers tend to use the language with a very limited number of interlocutors, the social characteristics of the interlocutors have a large impact on how the heritage speaker speaks and this in turn has consequences for the selected baseline. For example, when a heritage speaker uses the language only

with his or her mother who had a lower-class background in the country of origin, then preferably the baseline variety is also that of female lower class speakers. We know from sociolinguistic work that both class and gender and the interaction between class and gender affect which syntactic structures are being used.

Cheshire (1999) reports on the use of double negation in British English and shows that whereas middle class boys and girls do not use double negation at all, lower class girls use it 20% of the time and lower-class boys 60% of the time. Now let's go back to the heritage speaker who only speaks the language with his mother. He may copy the 20% double negation of his mother. If he is compared to age matched boys it may look like he is losing double negation and if he is compared to middle class peers, it may look like double negation is an innovation. Taking into account the social background of the language input providers, including SES and gender, will help interpret any findings.

## 6.6 Identity work, style shift, variation, and change

So far we have discussed variation in terms of the opportunities HL speakers have to increase their proficiency. As discussed in previous chapters, however, language use is also about identity work. Speakers may show variation because of social reasons. Like all languages, HLs may be subject to stylistic variation. Speakers may speak differently in different registers or at different stages of their lives. They may show accommodation to the researcher, they may or may not command a more formal style of their HL, and their HL may or may not have undergone stylization in the interplay between majority and minority ethnicities.

Style shift can be a type of age-grading. Age-grading refers to changes across the life span of an individual that are typical for many, in different generations. Rickford and Price (2013) studied the language use of two African American women in their teens and in their thirties. Whereas in their teens their language contained many instances of copula omission and absence of third person *-s*, this was no longer the case in their thirties. This change from vernacular African American English to Standard English is typical for many speakers. Teenagers will use vernacular characteristics and will decrease their frequency once they grow older. This is a shift in style. When shifts in style are common to many speakers in a community as they grow older, this is referred to by Labov (1994) as age grading.

Labov distinguishes different ways to investigate change. In the panel method, individuals are measured at two different points in their lives (about twenty years apart for example). If individuals have changed in these years, there are two possible types of explanation: either age grading has taken place (individuals change language with age just like older generations did before them) or there could be



a communal change (a change in all generations that has not occurred before). Another method is the apparent time method: one compares different generations to each other at the same moment. If different generations are the same there is either stability or communal change (all generations underwent the same change), if the generations differ this is either the effect of age grading or of generational change. Labov (1994) advocates the use of trend studies: apparent time studies (comparing different generations) at different points of time. Combining the two methods allows the researcher to interpret the different sources of change and stability. Although Labov (1994) indicates that his model is not intended for contact situations, it is important to realize that similarity and dissimilarity between generations may have multiple causes. If there is information on earlier versions of the HL and on language change in the HL, this is relevant for creating a balanced baseline. Table 6.1 gives an overview of the types of stability and change possible at the level of the individual and the community, as proposed by Labov (1994).

**Table 6.1** The types of stability and change possible on the level of the individual and of the community (based on Labov, 1994)

	Individual	Community	Linguistic domain
<i>Stability</i>	Stable	Stable	
<i>Age-grading</i>	Unstable	Stable	Conscious types of style shifting
<i>Generational change</i>	Stable	Unstable	Morphology Sound system
<i>Communal change</i>	Unstable	Unstable	Lexicon Syntax

## 6.7 Measuring proficiency and assessing linguistic profiles

Section 6.2 described different potential baseline groups, while Sections 6.3–6.6 discussed factors that affect HL use and acquisition and hence proficiency in that language. This section looks at ways to categorize heritage speakers and the groups they can be compared to in terms of their proficiency. Proficiency is often not explicitly taken into account in contact linguistics, because of the baseline problem discussed in Section 6.2. For existing tests, usually the monolingual standard variety is the baseline, and testing HL speakers in that variety does not always seem useful, and is often felt to be unfair. Perhaps because of that, proficiency was/is often only measured in a crude way. Sociolinguists may compare generations, and that these differ in proficiency at a very general level is usually obvious, but sometimes these differences are not explicitly tested. Nevertheless, many ways of testing proficiency exist; test batteries for measuring the proficiency of HL speakers are reviewed in great detail in Silva-Corvalán and Treffers-Daller (2016); see

also Birdsong, Gertken, & Amengual (2012) on bilingual profiles and Bosker et al. (2014) on using the perception of fluency by native speakers.

We can distinguish three functions that tests may have, namely (1) tests to differentiate between groups of native speakers in terms of proficiency; (2) tests to differentiate between different types of native speakers in terms of language dominance; and (3) tests to compare heritage speakers to other learner groups such as second language learners.

### 6.7.1 Cloze test

An often used measure of proficiency in second language assessment, and sometimes used in HL assessment, is the cloze test. The participant reads a segment of prose in which some words are systematically deleted from the text, and has to fill in the blanks. The test is quick, cheap and often considered reliable since research has shown that scores on the cloze test correlate well with other measures of second language proficiency (Stansfield & Hansen, 1983). Of course, cloze tests can only be used with literate HL speakers.

Montrul (2004b, 2010, 2011) found that heritage speakers' scores on the cloze test (actually part of the 'DELE' test battery, as discussed below) correlated very highly with accuracy on other linguistic tasks testing knowledge of tense, aspect, mood, and gender agreement. Note, however, that since its first application to bilingual learners (Carroll, Carton, & Wilds, 1959), there has been considerable uncertainty about the language skills and the cognitive processes that are involved in the cloze test (Stansfield & Hansen, 1983).

A frequently used cloze test in research on Spanish heritage speakers is the cloze part included in the DELE test (Diplomas of Spanish as a Foreign Language). These diplomas are 'official titles certifying degree of proficiency and mastery of the Spanish language, granted by the Ministry of Education, Culture and Sport of Spain'. These tests are frequently used for second language learners and are therefore suitable for researchers who would like to compare heritage speakers and second language learners with similar levels of proficiency. Many widely taught languages have standardized cloze tests available as part of a larger standardized test battery, for example, *Test Deutsch als Fremdsprache* (TestDaf), *Cambridge English Language Assessment for English*, *Hanyu Shuiping Kaoshi* (HSK) for Mandarin Chinese, *Certificado de Proficiência em Língua Portuguesa para Estrangeiros* (CELPE-Bras) for Portuguese, and *Test of Proficiency in Korean* (TOPIK) for Korean. You can also upload a text of your choice to specialized websites such as <<http://l.georges.online.fr/tools/cloze.html>> which will then develop a cloze test for you.

Valdés (1995) has questioned the validity of tests developed for second language learners when used with HL speakers because they are not meant for

the situations in which HLs emerge. HL speakers are not L2 speakers, often have learned the language only in oral settings, and they may have specific dialectal backgrounds. The specific test Valdés (1995) criticizes is the ACTFO (the American Council on the Teaching of Foreign Languages Oral Interview), but her arguments also apply to other tests developed for second language learners. Polinsky (2008b) suggests that the effectiveness of L2-tests for HL speakers differs per language. She claims that heritage Russian in the United States is close to the Standard language because most Russian immigrants are highly educated and were taught standard Russian in school. She claims that the ACTFO for Russian is therefore suitable for heritage speakers. Montrul agrees with Valdés that in oral tasks it is difficult to assess heritage speakers with the same tests as L2-speakers, but she found that the written part of the DELE predicted outcomes in other linguistic domains well for the heritage speakers.

### 6.7.2 Fluency measures

An increasingly popular way of categorizing heritage speakers is to investigate their overall fluency, for example by measuring the number of words they utter per minute (the words per minute or WPM rate). This measure takes more time than the cloze test, but it is possible to use it with illiterate speakers and its results are precise. Speech rate measured in words per minute has been shown to be one of the factors that best predict fluency in second language acquisition (Riggenbach, 1991). Of course, material needs to involve a specific task, also to make data comparable across speakers. Probably most popular have been picture descriptions and video retellings.

Examples of studies that have used the WPM-rate for heritage speakers are Polinsky (2008b), Nagy (2015), Moro (2016) and Irizarri van Suchtelen (2016). All these researchers look at variation within the group of heritage speakers; none of these compared heritage speakers to second language learners. If heritage users are able to and dare to speak, this method can yield a good predictor of fluency and of linguistic proficiency more generally. In order to measure speech rate heritage speakers are asked to describe pictures of objects or of a visual narrative, as in the Frog Story (Mayer, 1969) or Little Red Riding Hood (cf. Montrul, 2004a), clips or short videos (cf. Moro, 2016; Irizarri van Suchtelen, 2016). Their speech is transcribed in software packages like ELAN which aligns speech sounds with transcriptions and enables the researcher to filter out irrelevant pauses and speech of the interviewer and thus facilitates in calculating number of words per minute precisely. Polinsky (2008a) investigates retention versus loss of the three-gender system in Russian and shows that the speakers with the lowest speech rates are also the ones who have reduced the Russian three gender system to a two gender system.

Irizarri van Suchtelen (2016) calculated not only the WPM-rate in his heritage speakers but also their ‘uh-rate’, defined as ‘the total number of tokens which indicated *uh*-like sounds (such as *uhm*, *ah*, and *eh*) divided by the total number of words’, in the entire recording of a participant. Another study applying such a measure is Riggenbach (1991, p. 438), who found hesitation phenomena to be ‘salient in determining fluency level’. Irizarri van Suchtelen (2016) shows that lower WPM rates correlate with higher uh-rates in heritage speakers. He did not find such a correlation for first generation immigrant speakers nor for homeland speakers, and interprets the findings as supporting the assumption that both measures are related to the same underlying factors, i.e. ‘the deterioration in automaticity as a consequence of the history of exposure to the HL’. There are almost no differences between *uh*-rates of first-generation immigrants and homeland speakers, which suggests a ‘ceiling’ level with little variance. In other words, it may indicate that these speakers display an ‘ordinary’ or ‘default’ amount of hesitation in their speech (as opposed to an ‘above default’ rate in heritage speakers), which does not vary much between individuals. WPM rates do show individual variation in the non-heritage group, related to individual factors such as personality. To correct for personality effects some researchers measure WPM not only in the HL, but also in the dominant language (see Moro, 2016). Not all researchers find correlations between grammatical structures and WPM, for example, Nagy, Chocie, & Hoffman (2014). It is possible that this is an effect of the high proficiency of the heritage speakers she tested or that variation in WPM may differ per task (cf. Kagan & Friedman, 2003, p. 544 fn. 5).

### 6.7.3 Lexical proficiency tasks

If HL users are reluctant or unable to speak, another possible diagnostic test is a lexical task, such as lexical decision tasks or picture naming tasks. A test often used to compare heritage speakers to second language speakers is the lexical decision task in DELE or some variant of the Peabody Vocabulary Task. Both tests are suitable for non-literate heritage speakers and the tests are quick. In a lexical decision task, participants hear words and non-words and they have to press one button if they think a word is a real word and another button if it not. The more words that are correctly analysed as real words and the quicker these words are distinguished from non-words the more proficient the speaker is. Like cloze tests, lexical decision tasks are often included in standard test batteries. The Peabody Receptive Vocabulary test lets participants hear one word and at the same time see four pictures; the participant has to point to the picture that matches the word. An advantage of these tests is that they are widely used so norm interpretations are readily available. Lexical proficiency tasks from standard test batteries are good for

comparison between HL learners and second language learners. Some researchers have developed lexical proficiency tasks especially for HL users. We will look closer into two of those.

Polinsky (2006) tested Russian heritage speakers in the US on a set of eleven structural features and found that their scores on structural items correlated significantly with their lexical proficiency. She measured lexical proficiency with an oral Swadesh-list translation task. The Swadesh-list was originally developed in the field of historical linguistics to establish whether languages were genetically related. The list refers to basic words, which are usually not borrowed. If languages have cognates for many words on the Swadesh list this can be taken as evidence that the languages are genetically related. Polinsky (2006) transformed the Swadesh list in an oral translation task for heritage speakers. If the concepts the words stand for could be easily drawn, pictures were shown to the participants. These had to translate all words from the 100-word version of the Swadesh list. Every correct answer yielded 1 point, if the correct root was chosen but the wrong affix 0.5 point was given. No answer or a wrong answer yielded zero points (for the complete list with translations see Polinsky, 2006, pp. 201–203). The lower the score, the lower the proficiency of the person tested. Speakers with a score of 88 or higher turned out to perform much like homeland speakers on the eleven structural domains that were tested. Following the tradition in creole studies, Polinsky refers to these speakers as acrolectal speakers. Speakers with a score of 82–88 showed some structural problems (and are referred to as mesolectal speakers) and speakers with a score of 70–82 were most likely to have problems in the structural domain.

**Table 6.2** American Russian: Continuum of speakers (based on Polinsky, 2006, p. 253)

Speaker type	Basilectal	Mesolectal	Acrolectal
Score	70–82	82–88	88–90+

The goal of the lexical task devised by Polinsky was to allow comparison between different HL users. Acrolectal speakers are expected to perform closer to homeland speakers than a speaker who is in the basilectal range.

Another lexical task related to the knowledge of core vocabulary, the timed body part naming task, was developed in the HALA (Hawai'i Assessment of Language Access) project. The semantic domain of body parts is very basic and therefore available in all languages. Borrowing from other languages is less likely in this domain: the nose is less likely to give rise to a loanword than the I-pad. Because basic body parts are learned early, slow or blocked access to those words implies attrition.

The test is set up as follows: The researcher shows the participants pictures of body parts which vary with respect to the level of detail: from face to chin to

cheekbone. The more general words (face) tend to be learned earlier and tend to be easier accessible than the more specific words. Participants are instructed to say what they see in the picture as fast as possible and are specifically told to try to avoid ‘eh’ and to not use a determiner, just the bare noun. Reaction times are measured. The test does not require literacy and because the domain is so basic, it can be used for all languages.

The goal of the test is not to compare speakers to each other, but to compare the relative strength of two languages within an individual. The idea is that if a speaker is quicker in accessing words in language A than in language B, that language A is their dominant language. The underlying idea is that the frequency of use of the language and of the specific words in that language determines the activation strength of those words. The activation strength determines speed of access. Slower access to words suggests less use. Note that dominance is different from proficiency: speakers can be very proficient in two languages or not so proficient in either language (cf. Gertken et al., 2014). It is important to know about language dominance because it is possible that effects of cross-linguistic interference are more related to language dominance, i.e. to asymmetry between the mastery of the languages, than to language proficiency per se. Moreover, some heritage researchers are interested in language vitality and if words are accessed much slower in the HL than in the dominant language this suggests a decrease in vitality of the language.

O’Grady et al. (2009) used this test with Korean heritage learners, and expected the differences in reaction times to become larger when referring to more specific body parts, as these are used with lower frequency. The language with the fastest reaction time was the dominant language. Interestingly, differences between the languages became stronger in the low-frequency domain. The researchers also measured accuracy and some speakers were extremely accurate in both Korean and English. Had one only looked at the accuracy of their responses no asymmetry in dominance would have been detected between their Korean and English, whereas the relative distance between reaction times in the two languages did point to lesser use of Korean. Other methods to test language dominance are self-reports of language use per communicative setting (asking about language use with, for example, friends), sentence repetition tasks (Flege, MacKay, & Piske, 2002), word recall tasks (Golato, 2002), and lexical diversity scores in elicited speech samples (Treffers-Daller, 2011; for a broader overview see Gertken et al., 2014).

#### 6.7.4 Sociolinguistic background questionnaires

Questionnaires are a final useful means by which to categorize heritage speakers (usually in addition to linguistic tasks such as the ones reviewed above; see also

Chapter 5). These may include questions on self-rated proficiency, language use including information on language use per situation (home, church, school, sport club, neighbourhood etc.) and per interlocutor (parents, partner, friends, colleagues), language attitudes, language history (including age of onset of learning the HL and the dominant language), the social network, ethnic identity, experiences of discrimination or acceptance, and personal information such as gender, age, educational and professional background and SES (socio-economic status of participant and of parents).

The goals for which questionnaires are used differ across studies. Polinsky and Kagan (2007) describe a biographical questionnaire that facilitates correct placement for language courses; they find that the WPM-rate together with biographical information provides the most effective way to place learners in a language class. The BLP (Bilingual Language Profile) is intended to measure language dominance. Lindsey (2013) compared the HALA body parts test results to those of the BLP and found that the BLP may detect dominance shift later than the HALA. Knowing about language dominance is relevant for policy makers and language maintenance programs, but it may also be useful for answering questions about the linguistic profile of HL speakers. Do we observe effects of linguistic interference more in speakers who are dominant in the other language or not?

Almost all case studies of minority languages use some form of a questionnaire. The results can be used to identify the participants that have the right background characteristics to form the baseline, but also to map the variation in usage patterns within the HL community.

Knowing about the sociolinguistic background of the speakers facilitates understanding variation within the group and meaningful comparison across groups. Van Osch et al. (2015) for example, compare proficient heritage speakers of Spanish in the Netherlands to proficient second language learners of Spanish in the Netherlands. One goal was to find out to what extent Age of Onset is a relevant factor for proficient bilingual speakers. Two speaker groups were matched not only on the basis of linguistic tests (a lexical decision task and the cloze part of the DELE test), but also on their sociolinguistic profiles. The heritage speakers and the second language speakers had similar SES and all had spent some time in Spanish speaking countries in the recent past, for example. The more similar the speakers in the two groups are the more likely it is that any differences between the groups really can be attributed to Age of Onset effects.

Instead of making the groups as similar as possible, it is also possible to have a sample in which participants differ and overlap in different ways, and to calculate which factors have the most predictive power in accounting for variation. Montrul and Sánchez-Walker (2013) gathered information on many background factors of their participants to study the possible effects of these factors. They then

divided the group into speakers who performed like monolingual speakers on the test (which focused on whether or not Spanish object markers were dropped; proficient speakers would not drop them) and contrasted them to speakers who tended to drop the object markers. For every item on the questionnaire they checked whether the non-omitters (the monolingual-like group) and the omitters (the non-monolingual-like group) were evenly distributed. For example, they asked participants whether they speak Spanish only or both Spanish and English to their parents and found that the more monolingual group significantly more often spoke only Spanish to their parents.

## 6.8 Conclusion

This chapter discussed ways of investigating the presence or absence of changes in how heritage speakers use their HL. It gave an overview of possible groups to compare heritage speakers to, such as other heritage speakers, first generation immigrants, homeland speakers and second language speakers, and it gave an overview of factors that need to be taken into account to ensure a maximally precise and relevant comparison. Such factors include age of onset, stylistic repertoire, usage patterns, socio-economic status, dialect background and social connections.





# Heritage language phenomena and what triggers them

## 7.1 Introduction

Probably most HL researchers' focus of study is some kind of variation in the HL, resulting from change and producing further change. A secondary focus has been the retention of archaic structures, but this emphasis has now faded into the background. Recently some attention has also been paid to stability: what parts of the heritage language are stable across homeland speakers and various heritage speakers and why (Aalberse & Moro, 2014; Polinsky 2018a; Polinsky 2018b)? Below we focus on domains of change.

This chapter is divided into three main sections and one comparative final section. First, we survey some representative phenomena that have been studied (7.2) in the various linguistic domains. The third section discusses language internal factors that cause change and variation, including reduced input in and reduced use of the HL (7.3). Then we discuss changes that are more directly related to influence from the dominant language (7.4), external factors. Section 7.5 concludes the chapter by comparing the effects of these two different factors. To what extent can they really be kept separate, and to what extent do they interact?

### Main goals of this chapter:

To present an overview of the types of changes which occur in HLs.

To understand what language internal factors have an impact on change.

To understand what kinds of changes are induced by contact with another language.

To discuss whether we can keep these two main sources of change separate, and if so, how?

## 7.2 Phenomena studied

Change has been reported across the various linguistic domains. We give some examples per domain to give the reader a flavor of the changes that are apparently possible in HLs. We limit ourselves to illustrative examples, and do not attempt anything resembling full coverage of research results here.

### 7.2.1 Phonology

Some studies report that in the domain of phonology speakers with early exposure to an HL have an advantage over adults who learn or relearn the language. This advantage holds for perception as well as for production (Au, Knightly, Jun, & Sun, 2002; Oh, Jun, Knightly, Jun, Oh, & Au, 2003). We will zoom in on production here, since it has been documented better.

Recent studies that describe the phonology of heritage speakers include Queen (2012), Allen & Salmons (2015), Kupisch et al. (2015), Kupisch & Rothman (2016), Kupisch & van de Weijer (2016), Nagy (2015). Kupisch et al. (2015) consider the global accent of bilinguals and find that bilinguals are more often than not judged to have a foreign accent in their HL, but not in their dominant language. Native speakers find it harder to detect the origin of their accent, and judgment on their accents lie in between the values for L1 and L2 speakers. This supports the idea that heritage speakers have an advantage over second language learners in the domain of phonology. Pierce, Boas, and Roesch (2015) study the absence of fronted vowels such as the [y] in Texas heritage German. They relate this to the general markedness of these vowels, their absence in the dominant language English and their weaker status in the dialects of German on which the heritage variant was based. However, dialect studies suggest that this feature was absent in the regions in Germany where the later HL speakers came from. This points to the issue of the baseline as discussed in the previous chapter: fronted vowels may not have been present in the German varieties that arrived in Texas. Changes in HLs can often be attributed to a number of factors, which all may play a role simultaneously.

Kupisch et al. (2015), Kupisch and Rothman (2016) and Allen and Salmons (2015) all describe aspects of voicing in HLs, as well as features related to voice distinctions such as aspiration. Nagy (2015) reports that the voice onset time in the Russian and Ukrainian of HL speakers in Toronto has become more English like across generations whereas no such change occurred among Italians in Toronto. Kupisch and Rothman (2016) report on a set of studies on German-French and German-Italian simultaneous bilinguals. For both groups they investigated onset of voicing at the beginning of the word. The VOT in German is longer than in French and Italian. They found that heritage speakers of Italian and of French in Germany produced longer VOT in their Italian or French, respectively, compared to bilinguals who were dominant in French or Italian. The difference in length was not significant in the French-German group, whereas it was in the Italian group in the sense that the Italian heritage speakers in Germany had a significantly longer VOT than the German-Italian bilinguals in Italy. Allen and Salmons found a very small difference in voicing aspects between heritage Norwegian and Norwegian from Norway. To summarize, VOT in heritage speakers

can be affected but it does not have to be that way. It would be interesting to see what factors influence this.

Ahn et al. (2017) study the perception of several phonemic contrasts in Korean by heritage speakers in the US. These are a nasal versus lateral sonorant contrast (/n/ – /l/), also present in the dominant language English, a lax versus tense stop contrast (/t/ – /t\*/), which resembles the English /t/ – /d/ contrast, and a nontense versus tense fricative contrast (/s/ – /s\*/), which has no corresponding contrast in English. The heritage speakers' perception of L1-specific phonemic contrasts such as /s/-/s\*/ did indeed vary in the HL group (in contrast with a monolingual control group) as a function of the age of arrival, when contact with Korean became less intense.

### 7.2.2 Lexicon

In earlier literature on language contact it has been documented that HLs undergo various kinds of lexical influence from the other language. The main distinction is generally drawn between overt lexical influence, in the form of loanwords, and covert borrowing. The latter mostly takes the shape of novel meanings for existing words (semantic extension) or of novel combinations (loan translation).

We have seen in Chapter 4 that HLs generally borrow words from the dominant language. As such words are used by increasing numbers of speakers; they become established loanwords, enriching the HL vocabulary. At the same time, though, they may also have the effect of pushing out HL words, especially if such words have a meaning that is identical or similar to that of the loanword. We also saw that codeswitching may involve longer chunks of foreign lexical material, often because whole chunks are inserted from the other language. Prime candidates for such chunks are conventional combinations, i.e. collocations or multiword units.

Sometimes, such chunks are translated into the HL, so that a foreign-inspired collocation results, realized with HL words and morphemes. Such renditions are called loan translations.

#### *Loan translation*

Constructions and expressions from the dominant language can be directly translated into the HL. For example, Heritage Turkish speakers in the Netherlands often use the word *ev doktoru*, a combination of the words for 'house' and 'doctor', marked with the third person possessive morpheme *-u* as is required for compound nouns in Turkish. This word does not exist in baseline Turkish, and is a translation of Dutch *huisarts* 'house doctor'. The best English approximation for this concept is 'general practitioner', but in the context of the Dutch health care system the word has its own connotations, of the first person to go to for

treatment. In the Netherlands, most people are registered with one particular *huisarts*. Heritage Turkish has many such literal translations from Dutch, and their existence is one of the reasons why heritage speakers stand out as speaking ‘differently’ when in Turkey.

Marjolein Poortvliet (p.c.) describes a direct translation of English particle verb combinations in the heritage Dutch of a lady in Michigan who had left the Netherlands more than fifty years earlier. The American Dutch lady asked her *Wil je morgen inslapen?* intending to say ‘Would you like to sleep in tomorrow’. The grammar is Dutch with the particle *in* used pre-verbally, but the morpheme combination is based on English. Dutch actually has the verb *inslapen*, but it means either ‘to fall asleep’ or ‘to die in one’s sleep’. Another word, *uitslapen* (‘sleep out’), covers the concept of ‘sleeping in’.

### *Loan extension*

Under influence from the dominant language, HL words sometimes get used in ways in which the baseline variety does not use them, but which are similar to how their equivalents in the dominant language are used. American Russian speakers extend the meaning of the Russian word *neudobno* (‘uncomfortable’) to indicate both physical as well as psychological discomfort, whereas in baseline Russian it expresses physical discomfort but not embarrassment (Jarvis & Pavlenko, 2008). This extension is motivated by the interlingual identification between the words *neudobno* and *uncomfortable*.

In work by Schoenmakers-Klein Gunnewiek (1997, 1998) on Heritage Dutch in Brazil the focus is on such contact-induced meaning extension. Speakers use the verb *pakken* (‘to take’) in a broader range of contexts than Dutch speakers in the Netherlands do. In baseline Dutch, *pakken* (‘take’) implies that the subject acts intentionally and has control over the situation. Saying that *someone pakt de trein* (‘takes the train’) is fine, but *een ziekte pakken* (‘taking an illness’) is not, because one has no control over becoming ill. Schoenmakers-Klein Gunnewiek (1997) reports that HL speakers in Brazil do use the expression *een ziekte pakken* under the influence, presumably, of Brazilian-Portuguese *pegar* (‘take’), which is not limited to contexts of intention and control.

Whereas the two examples above are semantically motivated, loan extensions can also be motivated by phonological similarity. Pap (1949) describes that Portuguese *grosseria* meaning ‘rude remark’ in Portuguese has come to also mean ‘supermarket’ in Portuguese in the US due to phonological similarity with the word *grocery*. The word *humeroso* (‘capricious’) has come to also mean ‘humorous’.

For obvious reasons, semantic extension through phonological resemblance is more common when HL and dominant language are closely related, as they will share more cognates. For example, Clyne (1970) shows that partial semantic

correspondence and phonological similarity has motivated the use of *denken* ‘think’ in Heritage German in Australia at the expense of the verb *glauben* ‘believe’. Whereas Germans in Europe are likely to say *Ich glaube dass er kommt* ‘I believe he is coming’, Germans in Australia are more likely to say *Ich denke dass er kommt* ‘I think he is coming’, using the cognate of the word more commonly used in the English expression.

Hülya Şahin (2015) studied Papiamentu and Turkish as heritage languages in the Netherlands. One study (Indefrey et al., 2017) concerned the expression of static topological relations in heritage Turkish for two groups of speakers, Turkish-dominant and Dutch-dominant. In both bilingual groups, differences compared to monolingual speakers were observed. Dutch-dominant bilinguals showed more congruence between translation-equivalent Turkish and Dutch topological relation markers. Turkish-dominant bilinguals extended the use of a topologically neutral locative marker.

### 7.2.3 Morphology

Typical examples of morphological changes in HL contexts involve case marking on the noun phrase, for example in Hungarian as spoken in the United States (Fenyvesi, 1995/1996). While European Hungarian has a complex case system, with between 17 and 22 cases (depending on how one counts them), in heritage Hungarian many case markers are absent. Most affected are the essive and other locative cases. Also, one finds substitutions, again affecting the locative case system, where highly specific cases such as elative (‘out of something’) are replaced by more general cases such as ablative (‘away from something’). The basic movement/non-movement distinction is maintained, but the fine-grained meanings of the Hungarian locative cases are lost. Similar examples can be cited for ergative case in Dyirbal as an HL in Australia, as cited by Schmidt (1985), and the genitive case in heritage Scottish Gaelic, as analyzed by Dorian (1981; see also the discussion in Muysken, 2008, pp. 143–151).

In addition to nominal morphology, there is abundant evidence for morphological change involving verbal endings, as documented throughout this volume.

### 7.2.4 Syntax

There is a considerable set of studies dealing with HL syntax, mentioned throughout this volume. We will only select one source here, not discussed so far. Hartling (2016) describes cases of the contact-induced redistribution of syntactic structures. For example, whereas Danish speakers in Denmark would use the generic pronoun

in *man spise* ‘one eats’, as in A to express generic statements, heritage speakers in Argentina use the passive voice *spises*; is eaten’, as in B, like Spanish would:

A:	<i>det er et problem for man kan ikke spise penge</i> it is the problem since one can not eat money	TRADITIONAL DANISH
B:	<i>det er et problem, for penge spises ikke</i> it is the problem since money eat.itself not ‘it is a problem, because you can’t eat money’ <i>es un problema porque el dinero no se come</i>	DANISH IN ARGENTINA  SPANISH EQUIVALENT

Hartling also shows that the use of reflexive *se* in Spanish is often directly reflected in Danish *sig*.

This is similar to what is reported by Adjemian (1983; cited in Pavlenko & Jarvis, 2002) for English–French bilinguals. Their bilingual participant says ‘they want to fight themselves against this [tuition increase].’ The use of the reciprocal *themselves* is motivated by the subcategorization pattern of the French verb *battre* (‘to fight’) which selects a reflexive pronoun. It seems to be exactly the same pattern as we saw in heritage Danish in Argentina.

Finally, Villerius (2019) studied Javanese as a heritage language in the South-American former Dutch colony Surinam, where there has been a Javanese community since the 1870s. The Javanese migrants became bilingual first in Javanese and Sranantongo, and later on Dutch also became an important language in their repertoire. Topics focused on are locative constructions, multi-verb motion constructions, transfer events, and the expression grammatical voice (as it relates to focus). Villerius shows that Dutch and especially Sranantongo exerted a significant structural influence on heritage Javanese (see also Villerius, Moro, & Klamer, to appear).

In the following two sections we will describe internal factors in HL changes, having to do with the HL itself (7.3) and external factors (having to do with the contact with the dominant language(s) (7.4).

### 7.3 Language internal factors: Changes in the input for new generations of speakers

In this section we survey a number of language internal factors that have been suggested to account for the changes that occur in HLs. One idea in describing HLs is that heritage speakers cannot acquire all the structures a monolingual learner can acquire because of limited input and limited use. This is sometimes referred to as ‘incomplete acquisition’ (Levine, 2000; Montrul, 2008); another term is ‘partial acquisition’ (O’Grady, 2011). The term incomplete acquisition is problematic.

The intention of the term is that a structure that was offered in the input is not acquired at all, or that it is not acquired robustly and therefore lost again subsequently (Montrul, 2008; Silva-Corvalán 2018; Domínguez et al., in press). Polinsky (2008a) refers to the latter situation as ‘early attrition’. Early attrition would affect more linguistic domains than attrition later in life. Montrul (2008) even speaks of a critical period for language forgetting: the younger a learner is when he or she stops speaking a language, the more profound the loss will be. We can visualize the perspective of incomplete acquisition and early attrition as follows: The baseline contains a wide range of constructions and the HL contains a subset of these constructions.

Baseline	C1	C2	C3	C4	C5	C6
Heritage language	–	C2	–	C4	C5	C6

**Figure 7.1** The relation between the base line and the HL in the incomplete acquisition model

Authors like Putnam and Sánchez (2013), Kupisch and Rothman (2018), Bayram et al. (in press) and Putnam (2019) problematized the term ‘incomplete acquisition’ because of the harmful effects of the negative connotations of the term on heritage speakers and policy makers and because of the scientific backdrop of the term (see Silva-Corvalán, 2018 and Domínguez et al., in press, for an alternative view). Heritage grammars are referred to as ‘coherent grammars on their own’ (cf. Polinsky 2008) and the idea is that what is coherent cannot be incomplete (Bayram et al., in press). Additionally, what is the yardstick for calling a language complete (Putnam, 2019)? Moreover, ‘incomplete’ suggests that aspects of grammar are lacking, but sometimes structures have just been reanalyzed and have not been lost (Bayram et al., in press). For example, Moro (2015) describes the rise of grammatical categories like finiteness and definiteness in heritage Ambon Malay in the Netherlands which is more of an addition to the grammar than a loss. The label ‘divergence’ catches this change better.

Various authors describe a wider range of scenarios explaining differences between heritage speakers such as Polinsky (2018b) and Domínguez et al. (in press) apart from not acquiring every structure that is in the monolingual variety. Many authors stress the influence of the dominant language on the HL. Nagy (2015) emphasizes that language variation is an inherent part of all language acquisition (monolingual and bilingual) and that variation should not be placed in a deficiency perspective. Alternative labels proposed to describe the differences in the heritage grammar are ‘divergent’ grammar and for those aspects of the language that seem affected by lack of exposure, *unconsolidated* or *unstable* grammar (Putnam, 2019).



Polinsky (2018a, p. 9) notes that early research on heritage language focused strongly on divergence between heritage and homeland grammars, whereas stability in many areas is actually also very interesting (cf. Aalberse & Moro, 2014).

The term ‘incomplete acquisition’ suggests that a structure was actually present in the input. Pires and Rothman (2009) and Kupisch and Rothman (2018) argue, however, that what looks like incomplete acquisition may rather be complete acquisition of reduced input (see also Putnam & Sánchez, 2013). In that case, the language had already changed before the current generation was born. They report on heritage speakers of Brazilian Portuguese who do not use inflected infinitives, unlike baseline speakers. Pires and Rothman (2009) show that the absence of the inflected infinitive is not related to the HL situation per se (as heritage speakers of European Portuguese do use inflected infinitives) but related to the type of input heritage speakers receive. Heritage speakers of Brazilian Portuguese do not have inflected infinitives in their input, because for them use of the language is restricted to the home. Inflected infinitives are only part of the more formal academic register in Brazilian Portuguese and heritage speakers do not have access to this register. In European Portuguese, on the other hand, it is part of the spoken language. Pires and Rothman (2009) refer to this kind of change in HLs, resulting from restricted input, as ‘missing-input competence divergence’.

When structures typical of the academic register are not acquired, this represents a form of stylistic reduction in the HL repertoire. Stylistic reduction seems typical for HLs (see also Garrett, 2005 and some of the contributions to Dorian, 1989).

The term ‘incomplete acquisition’ is thus often met with objection, also from sociolinguists. They argue that incompleteness implies deficiency, sends out the wrong message to speakers and policy makers and that the language of heritage speakers should be regarded as ‘different’ rather than ‘deficient’. Moreover, it is argued that it unclear what complete means in language acquisition. If parts of grammar are in the input of heritage speakers but not (always) in their output, this grammar can be described as unstable. In short, heritage speakers receive less input than the baseline (quantitatively different) and they receive different input (qualitatively different) in the sense that it is often limited to the informal domestic register. Moreover, the language use of their parents may already show signs of attrition and cross-linguistic interference. Both qualitative and quantitative aspects of the input affect how people speak. Below we will discuss five factors that help determine which elements of the HL are likely to change as an effect of limited input and limited use: order of acquisition (7.3.1), frequency (7.3.2), optionality (7.3.3), restrictive use (7.3.4), and transparency (7.3.5).

### 7.3.1 Order of acquisition

One hypothesis on the likelihood of change in heritage speakers concerns the order of acquisition. What is acquired late is prone to loss. First, if acquisition is indeed incomplete or unstable it follows that what is normally learned last will not be learned well or at all. The hypothesis was formulated by Jakobson (1941) as the Regression Hypothesis; it was paraphrased by Montrul (2008, p. 69) as: ‘acquisition is the mirror image of attrition: structures that are acquired late in L1 acquisition will be the first to be affected in L1 attrition.’ Others who have discussed the regression hypothesis include Gonzo & Saltarelli (1983) and Keizer (2010a, b).

It is assumed that what is acquired late is somehow marked, while unmarked structures are acquired early. Ideas on markedness are particularly well developed for the domains of phonology and morphological inflection. For example, coronal sounds are less marked than non-coronal sounds (cf. Paradis & Prunet, 1991; Montrul, 2008, p. 70). The coronal sounds are acquired first, and the regression hypothesis predicts that they are more resistant to change than non-coronal sounds.

### 7.3.2 Frequency

High frequency of use of a language is obviously an important factor facilitating language acquisition and language maintenance, protecting against attrition (see Schmid, 2009). Frequency also plays a role in explaining what is vulnerable within HLs. Studies on the neurological aspects of activation have demonstrated that words that are used more frequently are more easily available and retrieved to the speaker than words used less frequently (the Activation Threshold Hypothesis in Paradis, 2007). The simplest hypothesis on the relation between frequency and change in heritage speakers is captured by the ‘Input Strength Hypothesis’ which O’Grady et al. (2011) formulate as follows

#### INPUT STRENGTH HYPOTHESIS

The most frequently encountered forms are acquired first and are the most accessible for language use throughout life.

If this hypothesis is correct, those elements that are frequent in the early input would be stable in heritage speakers, while infrequent forms and constructions would be unstable.

Frequency can explain some developmental patterns. Concerning lexical development, it is shown that toddlers learn frequently presented novel words faster than infrequently presented ones (Schwartz & Terrell, 1983; Gershkoff-Stowe, 2002) and that adults retrieve and recognize high frequency words faster than less frequent counterparts (Jurafsky, 2003). Similar effects have been observed

for syntactic constructions as well: compared to less frequent sentence types, the ubiquitous subject – verb – object (SVO) pattern of English manifests a faster learning curve in the course of acquisition, is less likely to be lost in the case of aphasia, and is less likely to be derailed in the course of processing (Townsend & Bever, 2001; Dick, Bates, Wulfeck, Utman, Dronkers, & Gernsbacher, 2001; Ferreira, 2003). Ambridge et al. (2015) is a recent overview of the role of frequency in child language acquisition.

Some inflectional phenomena also show effects of input strength. Drawing on detailed longitudinal data from three children, Maratsos (2000) observed that learners come to systematically use frequently heard irregular past tense verbs like ‘went’ and ‘saw’ before less common ones such as ‘sank’ or ‘won’. This suggests a frequency effect – irregular past tense forms have to be encountered a certain number of times before they are mastered, and the more frequently occurring forms are acquired first. A case study on heritage speakers of Mandarin Chinese shows that the most frequently used classifiers in Mandarin Chinese are over-generalized while less frequent classifiers are lost. This loss can simultaneously be related to acquisition order: the less frequent classifiers are also acquired later (Aalberse & Moro, 2014).

In other cases, however, the effect of input on the developmental profile is much harder to discern. For instance, the definite article ‘the’ is the most frequent word in the English language; yet it comes to be used productively at a relatively late point. Brown (1973, p. 358) reports that use of the articles ‘the’ and ‘a’ attains 90% accuracy much later than several less frequent morphemes, including the progressive marker *-ing*, the plural suffix *-s*, the possessive morpheme *-’s*, and at least some irregular past tense forms.

Why is frequency not always the key to developmental patterns? O’Grady et al. (2011) hypothesize that learning takes place only to the extent that children encounter situations in which forms can be successfully linked to a corresponding semantic function, thereby creating mappings. They revise the Input Strength Hypothesis as follows:

INPUT STRENGTH HYPOTHESIS (REVISED):

The most frequently instantiated form – meaning mappings are acquired first and are the most accessible for language use throughout life.

In order for an instantiation of a form – meaning mapping in the input to be useful to a learner, two conditions must be met: the form must be audible, and the meaning must be discernible. Crucially, neither condition can be taken for granted, especially in the case of ‘grammatical morphemes’ (verb inflection, determiners, auxiliary verbs, etc.).

### 7.3.3 Optionality

Many changes described in HLs are related to optionality in some way. The central idea is that what is optional in language is more vulnerable to change than what is categorical. Hypotheses that take some type of optionality into account in explaining change include the hypothesis of Indeterminacy (Polinsky, 2011), Interface Hypothesis (Tsimpli & Sorace, 2006), the Smaller Domain Principle, the Vulnerability Hypothesis (de Prada Pérez & Pascual y Cabo, 2012), and the hypothesis of Transparency (O’Grady, 2011). However, there is a long tradition of precursors to this in the study of language contact and codeswitching: Aikhenvald (2002) makes a distinction between system-preserving and system-altering changes (concluding that most changes are of the former kind, i.e. you rarely get entirely new structures in a language). Heine and Kuteva’s framework (2003) privileges the elaboration of existing structures (‘minor use patterns’). Myers-Scotton (2006) allows for deviations from the canonical integration of foreign words (‘code-switches’) into the Matrix Language structure by the overuse of possible but rare structures. Poplack’s Equivalence Constraint (1980) in codeswitching also rules out ungrammatical structures, ungrammaticality referring to things that violate structures categorically.

Let us start with the hypothesis of Indeterminacy. Polinsky (2011) defines indeterminacy as follows:

INDETERMINACY HYPOTHESIS

Form X is suitable for multiple syntactic contexts, and the same syntactic context allows for more than one form

Differential Object Marking in Spanish is a typical case of indeterminacy (Montrul & Bowles, 2009). Thus the Spanish accusative object marker *a* could be considered indeterminate because it also marks datives (*lo doy a Juana* ‘I give it to Juana’) and infinitives (*voy a comer* ‘I am going to eat’), and because it only occurs with some objects, typically animate definite (*veo a Juana* ‘I see Juana’ versus *veo la casa* ‘I see the house’).

The Interface Hypothesis states that if different modules of language interact (modules include phonology, lexicon, morphology, syntax, etc.), this aspect of the language will pose processing problems and therefore be more vulnerable than aspects of the language that operate within one module. External interfaces, e.g. interaction with non-linguistic modules, are expected to be the most vulnerable.

INTERFACE HYPOTHESIS

Narrow syntax < Internal interfaces < External interfaces

Van Osch et al. (2014) apply the Interface Hypothesis to subjunctives in Spanish. Epistemic verbs in Spanish require the use of the indicative, as in *Yo sé que vas* /\**vayas conmigo* ‘I know that you go with me’ while volitional verbs require the subjunctive: *Quiero que \*vas/ vaya conmigo*. ‘I want that you go with me.’ Since the choice between the indicative and the subjunctive is determined by the matrix verb, i.e. within “narrow” syntax, which defines the structure of the clause, this should not be vulnerable to change in heritage speakers of Spanish, and this indeed is what is found. In contrast, in clauses with negated epistemic, perception and communication verbs, the choice of mood in the embedded clause depends on the speakers’ commitment to the truth of the proposition. An example of a commitment indicative is *Pedro no dice que es / #sea su culpa* ‘Pedro doesn’t say that it is his fault’. The implication of the indicative is that I actually know it is Pedro’s fault. An example of a non-commitment subjunctive is *Pedro no dice que #es / sea su culpa* ‘Pedro doesn’t say that it is his fault’. The implication here is that the speaker indeed does not know whether it was Pedro’s fault or not.

In a way similar to the Interface Hypothesis, the ‘Smaller Domain Principle’ predicts that smaller domains are easier for the language user, leading to the following implicational scale or hierarchy (Reuland, 2011; Koornneef et al., 2011).

SMALLER DOMAIN PRINCIPLE

Narrow Syntax < Logical syntax (C-I interface) < Discourse

The Smaller Domain Principle is somewhat more specific than the Interface Hypothesis in its predictions.

If we follow the ideas advanced by Reuland (2011) and Koornneef et al. (2011), linguistic encoding formed in components farther to the left on the hierarchy is ‘less costly’ in terms of processing and production than those towards the right because fewer considerations need to be taken into account. These predictions have been formulated to hold for competent speakers, who have fully acquired a given language. For the purposes of this book, this suggests that we should expect heritage speakers to show different degrees of difficulty with elements that belong in the different components of the hierarchy. In particular, we expect that phenomena that involve semantic and discourse ‘computation’, i.e. paying attention to semantic factors and discourse considerations, will be more difficult than phenomena governed primarily by structural syntactic constraints. Within the semantic and discourse components, we expect a further difference: semantic computation should be easier than the computation of discourse-related elements. Laleko & Polinsky (2016) cite a number of earlier researchers (Givón, 1979; Langacker, 1987, 2008; Reinhart, 1983, 2006; Grodzinsky & Reinhart, 1993; Frazier & Clifton, 1996) who formulated broadly similar claims, from different theoretical perspectives.

The pragmatic information of commitment that has to be integrated as part of the meaning implies an external interface. It turns out that such interfaces are vulnerable in HLs. This has been formulated as the Vulnerability Hypothesis (de Prada Pérez & Pascuál y Cabo, 2012):

VULNERABILITY HYPOTHESIS

Categorical distribution ↔ variable distribution

Less vulnerable ↔ more vulnerable

The idea is that the cause of vulnerability is not the interface as such but the variable distribution of the data that affect the choice between subjunctive and indicative. Vulnerability could also describe the accusative marker just mentioned.

At several places in this book we have mentioned the idea that semantic transparency plays a role in determining outcomes in HLs, as e.g. proposed by O'Grady et al. (2011) for the loss of case marking as a clue to grammatical meaning in heritage Korean. The Transparency Hypothesis can be formulated as follows:

TRANSPARENCY HYPOTHESIS

One-to-one form meaning mappings are easiest both to acquire and to retain.

The difficulty of optionality in grammatical constructions is also illustrated by problems in dative case marking in heritage Hindi. Montrul, Bhatt, and Bhatia (2012) show that dative marking is stable for marking recipients as indirect objects, as in 'I give him a book', where 'him' receives dative case), but that dative subjects are unstable in the speech of heritage speakers. This vulnerability is explained as the result of optionality: whereas recipients are always encoded with dative case (non-optional), subjects are more frequently encoded with nominative case. This is likely to cause uncertainty in heritage speakers.

Taking one step backwards, possibly the Smaller Domain Principle and the other principles mentioned can be subsumed under a general notion of ECONOMY (as in Optimality Theory or Minimalism, see Chapter 5), which would also yield a further range of predictions regarding HL production (cf. Irizarri van Suchtelen, 2016). All theories share the assumption that features involved in optionality, involving one to many mappings and many to one mappings, are difficult to acquire and difficult to retain.

From a usage-based perspective, the suggested explanation is that what is categorical is well entrenched and pre-empts the use of any alternative, while what is variable can easily shift in frequency of use, becoming more or less frequent. Slight shifts in how the word or structure is used (semantic extension, shift in combinability as in loan translations, and shifts in pragmatic markedness) come with frequency shifts. This is illustrated in work on Differential Object Marking

(Montrul & Bowles, 2009; Guijarro-Fuentes & Marinis, 2011; Irizarri van Suchtelen, 2016), where various factors are seen to constrain the output of different speaker groups (see 6.3).

#### 7.3.4 Restricted use

Heritage speakers use their language differently from baseline speakers, and this affects the development of the language. Aalberse and Moro (2014) explain an aborted case of grammaticalization, involving the Malay possessive marker *punya*, as the result of the different usage situations of the language. Grammaticalization means that a lexical item becomes more functional. It often implies a phonological reduction of the form as well. One factor driving grammaticalization is that frequent words tend to become shorter in form (see Principle of Least Effort in Zipf, 1949, and the OT constraint *ECONOMY* in Lestrade, 2010) and more general in meaning. Heritage speakers, however, make limited use of the language and words are therefore activated with a lower frequency than in baseline speakers.

A related phenomenon is predictability due to co-occurrence. Piantadosi et al. (2011) have shown that speakers tend to shorten the most predictable words, those carrying little information content. Heritage speakers, like late bilinguals, have more difficulties in predicting upcoming words in speech, since they use word combinations with lower frequency, and therefore they are less likely to reduce pronunciation. As pointed out by Haspelmath (2008, p. 47): ‘predictability allows shortness of coding, while non-predictability requires explicitness of coding’.

The third factor that explains the lack of grammaticalization in the heritage group is the great value heritage speakers attach to perceptual ease. Ernestus (2000, p. 24) shows that elision in casual speech is related to the relative importance that speakers attach to ease of articulation over ease of perception. Speakers who favor ease of articulation tend to reduce their speech, while speakers who favor ease of perception are less likely to reduce their speech. Baseline speakers tend to belong to the former category and heritage speakers tend to belong to the latter category. The idea that heritage speakers tend to value ease of perception is also captured by the Explicitness Hypothesis. Heritage speakers may be more likely to lack the confidence that their message will be understood properly, and therefore introduce more overt elements that are supposed to guide the hearer in processing than monolinguals would (cf. Polinsky, 2006). Laleko and Polinsky (2016) refer to the ‘silent problem’ and note that heritage speakers have most trouble with linguistic segments that are covert, not spelled out, and represented by some kind of silent exponent, be it at the level of morphological encoding (null morphemes), lexicon (null pronouns), or contextual deletion such as ellipsis.

The heritage situation does not always lead to less grammaticalization. For example, the progressive in Spanish is used more in informal speech than in formal speech (Torres Cacoullos, 2000). Because heritage speakers mainly use the language in informal settings, change towards higher use of the progressive may accelerate in these constructions in HL speakers. Indeed, we find many reports on the overgeneralization of progressives in HLs (Shi, 2011; Putnam & Sánchez, 2013; Brown & Putnam, 2015; Irizarri van Suchtelen, 2016; Moro, 2016).

## 7.4 Cross-linguistic influence: External factors

The previous section focused on language internal factors in explaining changes that occur in heritage varieties. A second possibility is that HL structures interact with structures from the dominant language, causing transfer. We will discuss four common forms of cross-linguistic transfer, or ‘interference’: the dominant language acting as a filter for grammatical categories in the HL (7.4.1), convergence through a shift in distribution (7.4.2), loan translation and semantic extension (7.4.3), and the adaptation of grammatical categories via the dominant language (7.4.4). It should be stressed that compared to internally caused changes, changes due to influence from a dominant language have been underexplored in HL studies, in contrast to most studies on language contact in the broader sense. Many examples have been documented but the findings generally have not been subjected to systematic comparative experimental study.

### 7.4.1 Filter of grammatical categories via the dominant language

Grammatical categories that are absent in the dominant language prove vulnerable in the HL. For example, Albirini, Benmamoun, and Chakrani (2013) argue that the reason why the dual is lost in heritage Palestine and Egyptian Arabic is that the dominant language English does not have a dual marker. Absence of a category does not automatically lead to loss, however. For example, in heritage Chinese in Canada and in the Netherlands, classifiers are to a large extent retained despite the absence of classifiers in the dominant languages (cf. Nagy et al., 2014; Aalberse & Moro, 2014).

### 7.4.2 Convergence through a shift in distribution

Some claim that the effect of the dominant language is more indirect. When there is optionality in the HL (also see 7.1.4), the option that is shared with the dominant language will win out. This hypothesis is also known as the alternation



hypothesis (Jansen, Lalleman, & Muysken, 1981), and has gained wide acceptance in the HL literature. The facilitating role of partial overlap is also described by Montrul (2004a) and Silva-Corvalán (1994, 2014). The increase in the use of shared constructions is also referred to as ‘structure preserving change’ by Backus (2013) and as ‘frequential copy’ by Johanson (2002, 2008).

The notion of system preserving change was applied to heritage Malay by Moro and Klamer (2015). They show that the DOC construction (a three-place predicate structure without overt marking of the recipient) was already allowed in Malay but that its use increases in heritage Malay in the Netherlands due to cross-linguistic interference.

Another example is the use of psych verb constructions in heritage Spanish in the Netherlands (see Irizarri van Suchtelen, 2012). Spanish allows both option A with a subject experiencer and option B with a dative experiencer. Whereas monolingual controls showed a preference for B, bilinguals showed a tendency towards preferential use of A, the only structure that is available in Dutch.

A: SUBJECT EXPERIENCER

*olvidó su llave.*

forget.PAST.3SG 3SG.POSS key

‘He forgot his key.’

(HiG2-J)

B: DATIVE EXPERIENCER

*Se le olvidaron las llaves.*

REFL DAT.3SG forget.PAST.3PL DET.F.PL key.PL

Lit.: ‘The keys were to him forgotten.’

(HiG2-H)

Many cases of convergence have been documented in the literature, of which we have only mentioned a few here.

### 7.4.3 Loan translations and semantic extensions

In Section 7.2 we discussed the phenomenon of loan translations. As with loan-words, loan translations by definition can only appear in a language by virtue of external influence. Heritage speakers may exhibit this phenomenon in quite extensive ways, and this may be linked to how meanings are conceptualized.

Schoenmakers-Klein Gunnewiek (1997, 1998) hypothesizes that HL speakers are affected by conceptualization strategies in the dominant language. Even when they retain the lexicon of the HL, the concepts attached to the words are those of the dominant language such as described above in the example of *een ziekte pakken* (‘taking an illness’) in HL Dutch in Brazil. Flecken (2010) shows that conceptualization also affects grammatical features. Grammatical aspect marking, for example, is linked to particular perceptions of reality. The Conceptualization

Hypothesis then predicts that those structures that are conceptualized differently in the HL and in the dominant language will be vulnerable to change. This idea was also explored in Backus and Dorleijn (2008), where lexical and grammatical convergence are placed on a continuum.

#### 7.4.4 Contact induced grammaticalization or additive borrowing

An important impulse for the study of cross-linguistic influence has come from the work of Heine & Kuteva (2003). Cross-linguistic influence can also take the form of grammatical categories from the dominant language being integrated into the HL. Perhaps surprisingly, this path is not often discussed in the HL literature, while it is extremely well studied in contact linguistics. Incorporation of grammatical categories from the dominant language into the other language is clearly a possible outcome of bilingualism.

By far the majority of documented cases concerns changes to existing categories, rather than the borrowing of categories or grammatical distinctions that did not exist yet in the HL before contact. Aikhenvald (2002) refers to the most common kind of change as ‘system-preserving change’; new categories would be ‘system-altering’ (see Backus, 2005 for a review). Trudgill (2011), following Nichols (1992) and Aikhenvald (2002), also discusses language contact as a possible source for grammatical changes. Whereas adult second language acquisition typically leads to simplification, Trudgill (2011) sees *early* bilingualism as a possible source of contact induced complexification. He relates the presence of linguistic areas to early bilingualism. Mixing languages is one way in which languages would become more complex. Whether mixing occurs is not related to proficiency but to specific sociolinguistic practices (Li Wei, 2007, p. 5). In our data we found some examples of the use of Dutch inflectional markers in heritage Chinese. For example, Chau (2011) describes a Dutch heritage speaker of Cantonese who uses Dutch inflection in her Cantonese as shown in (3).

*Keoi5 go3 snaar laan6 zo2-t*  
 佢 個 snaar 爛 咗-t  
 SG CL snaar string tear ASP-T  
 ‘His guitar string broke.’

Queen (2012) reports that Turkish-German bilinguals in Germany use German and Turkish intonational patterns in both their German and their Turkish and add a pragmatic load they do not have in the monolingual varieties of either language. Yip and Matthews (2007) show transfer of grammatical elements of Cantonese Chinese in the English of their children that resemble features of Singapore English. Both in the variety of their children as well as in Singapore English the adverb

*already* is used similarly as the Chinese aspect particle *le* and the adverb *ever* is used as an experiential particle. Jeanine Treffers-Daller (p.c.) has remarked that Turkish returnees use evidential marking in their German. She said that evidentiality became such an important part of conceptualization while using Turkish that the speakers retained the category even though German does not encode evidentiality grammatically. This leads to utterances in German that include the Turkish suffix *-mis*. In the example below *-mis* indicates that the speaker wasn't there when Michael sent the letter:

*Michael hat den Brief losgeschickt-mis*

Michael has the letter sent-EVI

'Michael sent the letter.'

Heine and Kuteva (2003) summarize much of the literature on convergence and describe many examples of contact induced 'grammaticalization'. Crucially, their examples do not involve bringing in a new grammatical category into a language: they are mostly cases where an existing category undergoes extension, usually because it gains wider distribution and greater frequency, all under the influence of its equivalent category in the other language. All these characteristics are typical of grammaticalization processes, hence the use of this term. Heine and Kuteva (2003) assume that many generations are needed to complete a process of contact-induced grammaticalization. This may explain the relatively low number of described cases of added complexity. There is some preliminary evidence of a distal demonstrative developing into a definite marker in heritage Mandarin (cf. Aalberse, Zhou, & Andringa, 2017) and the same goes for Ambon Malay *-nya* which also seems to be turning into a definiteness marker (cf. Moro, 2016).

## 7.5 Comparing internal and external factors

This chapter has compared perspectives on vulnerability and stability in HLs in terms of both language internal and external or contact factors. Independently of the characteristics of the dominant language, some structures seem vulnerable to change. Forms that are infrequent, have non-transparent form meaning mappings, require integration of different modules, are marked and acquired late are all prone to change. If we take the cross-linguistic interference perspective, possible vulnerabilities are categories that are absent in the dominant language. If there is optionality in the HL, the structure shared with the dominant language will win out. Highly automatized grammatical categories in the dominant language can become integrated into the HL. Finally, semantic characteristics of translation equivalents

in the dominant language and collocations from the dominant language can be transferred to native words in the HLs.

It is not always easy to distinguish sources of change. For example, if we find that the subjunctive is used less in heritage Spanish in the Netherlands, this could be an effect of markedness (the subjunctive is more marked than the indicative, see Andersen, 1989, 1991), but it could also be an effect of the absence of subjunctives in Dutch. The loss of the dual could also be interpreted in terms of markedness (the value is considered marked and acquired late) as well as the result of absence of the category in the dominant language. Cantonese Chinese has an inverted word order in ditransitive structures with the verb 'to give'; (e.g. 'I give book him' rather than 'I give him a book'). This structure is under pressure in heritage Cantonese. We can relate the vulnerability to age of acquisition (this order is acquired late), non-transparency and infrequency (all other ditransitive structures take a different word order) and to absence of the structure in the dominant language.

Grammaticalization of certain categories could be contact induced but could occur spontaneously as well. For example, there are dialects of Chinese that, like Dutch heritage Chinese, show the use of classifiers as definiteness markers. Some aspects of change are convincingly related to cross-linguistic interference, such as shift in conceptualization, but and some other changes are clearly not related to cross-linguistic interference. For example, Polinsky (2008c) shows that heritage Russian speakers lose functions of the relative markers (heritage speakers can only use relative clauses in which the relative marker functions as subject) although both Russian and English can use relative markers with object functions. The loss of allomorphy also seems unrelated to cross-linguistic interference.

One important factor we have not discussed is saliency. Various different types of saliency have been discussed in the literature, but the concept needs further elucidating. Acoustic saliency is one input-related factor that affects acquisition success (see O'Grady, 2011 and O'Grady et al., 2011). Acoustic salience is related to audibility. The better a form is heard the higher the chance that it is retained. O'Grady (2011) manipulated acoustic saliency experimentally. He demonstrates that when the saliency of the Korean accusative marker *-(l)ul* is enhanced in an experimental setting by manipulating the volume, the duration and the pitch), heritage speakers perform better in a comprehension task. Polinsky (2011) shows that phonologically heavy case markers are overgeneralized in heritage Russian at the expense of phonologically light case markers. There is also quite a bit of work on pragmatic saliency, related to information structure (Duranti, 1994).

## 7.6 Summary

The take home message of this chapter is that changes in HLs can be due to both language internal factors, including reduced input in and reduced use of the HL, and to influence from another language. Whenever a change in a heritage is described, ask yourself what the possible explanations for this change may be, and how to construct a tight argument for a particular explanation. Often these causal factors are difficult to separate; explanations often involve interaction between them.

# Grammatical models and research paradigms

## 8.1 Introduction

This chapter gives an overview of theoretical approaches to variation and change in HLs and of the linguistic models they use to interpret and to predict language change. We will briefly discuss the history of these models, their aims and their interpretation of heritage data.

We will begin the survey with generative linguistics, which has been applied to HL studies from the perspective of the impact of language acquisition scenarios (8.2). We then turn to variationist models, which take variability in production as their focus of interest (8.3), and to Optimality Theory, which tries to model variation and change in terms of the ranking of competing constraints which are in themselves universal (8.4). From a usage-based perspective, finally, it is also important to link HLs to the fact that language is an inherently dynamic phenomenon, constantly undergoing changes, in interaction with other languages that share the same social ecological space, and with a considerable amount of internal variation (8.5). Within the usage-based perspective, the type of descriptive models used often link up with the functional and cognitive considerations that inform linguistic typology.

It is also important to note that the research paradigm that a researcher chooses for a particular project often helps determine the nature of the data collection. What are considered good data in one framework may be less useful for another one. Nonetheless, by making the point of departure of particular studies explicit, the possibilities of comparing results across paradigms increase.

### **Main goals of this chapter**

To present theoretical models and linguistic research paradigms.

For each model to present its origins, aims, and main concepts.

To illustrate each model with an HL case study, in terms of data selection and method, analysis, and interpretations.

## 8.2 Generative grammar

Many studies in modern linguistics as it developed in the 1960s and 1970s were carried out in the generative framework, and some influential research on HLs has been carried out within this framework. It is impossible to even attempt to do justice to this model in a few paragraphs, given its long history, research output, and large number of practitioners.

The generative perspective has been linked particularly to theories about acquisition. HL speakers are first language learners (early acquisition in a natural setting) who share situational aspects with second language learners, crucially the existence of a strong contact language. In addition, their language output also exhibits characteristics of monolingual first language learners and of second language learners. By comparing heritage speakers' linguistic outcomes to both the output of second language learners and monolingual first language learners, as well as comparing the learning environments of these different learner groups, generative researchers have tried to gain more insight into the role of the age at which L2 acquisition started, the quantity and the quality of the input, the quantity of the output, the nature of bilingualism in general and of the role of language dominance in particular. HLs are therefore an ideal test case for developing grammatical models and the theories of language acquisition and language change that they are associated with.

### 8.2.1 Outline

The origin and aims of generative linguistics are well known, of course.

#### *Origin*

The roots of the generative approach lie in the 1950s, in the work of Noam Chomsky (born 1928), a theoretical linguist who did much to revolutionize modern linguistics and strengthen its cognitive underpinnings. Chomsky has published scores of books and articles, but his *Aspects of the theory of syntax* (1965) remains a key reference. Among many other such books, you may want to consult Isac and Reiss (2013) for a more general overview of the developments and implications of the generative research program.

#### *Aims*

The aim of generative linguistics is to account for the knowledge of language as a uniquely human and domain specific mental capacity. Chomsky's work, which is primarily focused on syntax although the other components of language are not excluded, has served to stress the fact that the structure of language is multi-layered

(traditionally referred to in terms of deep and surface structure) and hierarchical, involving nested constituents.

In the original framework, there were phrase structure rules, building the structures in the clause, and transformations, which had the power to alter these structures. An example of a phrase structure rule + transformation is the rule of Verb Second. This rule captures the observation that in Dutch all verbs are placed in sentence final position (b)–(d) apart from the finite verb in main clauses (a), as shown in the following sentences illustrating the position of the verb in Dutch:

- (a) *Wij eten een appel.*  
we eat an apple
- (b) *Wij hebben een appel gegeten.*  
we have an apple eaten
- (c) *Omdat wij een appel eten, ....*  
because we an apple eat
- (d) *Omdat wij een appel hebben gegeten, ....*  
because we an apple have eaten

As is clear from (a)–(d) all verbs follow the object *appel* ‘apple’ apart from sentence (a), where it precedes the object. The transformational perspective on this is that the basic underlying word order in Dutch is OV and that the finite verb in main clauses moves to the second position.

Children, when learning the language of their immediate environment, have to rely on the input they get, but this input can only be processed because the children have access to rich innate knowledge of what to expect. The child is particularly sensitive to this innate knowledge in a specific phase of its development, the so-called **critical period**. For this reason, the **age of acquisition** of a language and the learning context is crucial to the model, as became clear in Chapter 7 on possible sources of variation in HL speakers.

Also crucial is that there may be **indeterminacy** in the input: given their multi-layered nature, some surface patterns may be interpreted in terms of several different underlying structures, possibly leading to variation in new patterns being produced by the child. This is related to the assumption of the **poverty of the stimulus**, the idea that the language learning child is able to reconstruct the same grammar as its parents have in spite of limited input. Indeterminacy in HL would predict that strings which may have multiple structural interpretations are more likely to undergo change in an HL setting.

A number of other features of generative linguistics are relevant to the research on HLs. The central idea behind **parameter theory** is that there is a limited set of



principled points of variation on which languages differ, and which are crucial for the language learning child in determining the nature of highly complex language input (e.g. Hyams, 1986). If, as a result of bilingual convergence or attrition, parents produce patterns in their daily speech which differ from the previous baseline, e.g. in a quantitative sense, this may lead to parameter resetting by the child. Cases in point may be **pro-drop** (see e.g. Montrul, 2004a) or **basic word order**, which have been frequently studied in HL research. Parameter theory is not universally adopted any more as a model for linguistic variation, but it is relevant to many research publications in this domain.

Another set of theoretical assumptions from generative linguistics that have played a role in HL research is the assumption of the **modular organization** of our linguistic knowledge. Under this assumption, each module of our language capacity (think syntax, lexicon and morphology, phonology, semantics, pragmatics, although other more specific proposals have been made as well) has its own primitives and functions as a self-contained unit. However, there have been various interpretations in the theory of what the modules actually are: classical grammar 'components' such as syntax or morphology, or more abstract notions involving cognition, vision, etc. as in the work of Fodor (1983) and Jackendoff (1997). In this view grammar provides the optimal output to ensure interpretation/legibility of representations to the sensori-motor and conceptual-intentional interfaces.

The idea of modules thus naturally leads to the **interface hypothesis** (see Chapter 7): if modules are self-contained and have their own primitives and organization, they need to interact in production and perception, and the cognitive costs involved may lead to indeterminacy and change in an HL context. As we have seen in Chapter 7, and will engage with more below, some HL studies have appealed to the interface hypothesis to explain phenomena of change (Sorace, 2011).

Many accounts of variation between different languages, and hence of language change, in this tradition, have turned to **functional categories** and the strength of feature specifications in this domain. Thus, for instance, changes in definiteness and gender have been appealed to in order to explain changes in agreement patterns in heritage Spanish (van Osch et al., 2014).

Finally, many generative linguists support the assumption that there is a difference between our linguistic **competence**, the somewhat abstract knowledge of the rules and patterns of our grammar, and **performance**, the practice of using this knowledge in actual production and comprehension. It may be, however, that this distinction cannot really be made, and that syntax cannot be viewed as a cognitive system separate from language processing (cf. e.g. Philips, 2013), though many linguists do adopt this assumption.

Since competence has a somewhat privileged status in generative linguistics, researchers often directly access **grammaticality judgments** of the speaker

population under study, which are assumed to reflect linguistic competence directly. This leads to research designs in which subjects are asked to carry out grammaticality or acceptability judgment tasks (see Chapter 5). Sometimes these judgments concern sentences that are hard to find in corpora or are virtually absent in ordinary speech.

### 8.2.2 Case study

A typical case study in the generative paradigm is Laleko and Polinsky (2016), who investigate case and topic markers in heritage Japanese and Korean. We selected this study because it is a good example of an investigation that follows from deductive reasoning. Various possible explanations for vulnerability are translated into predictions of areas of syntax where heritage speakers may be relatively vulnerable. This specific study is a follow-up to Laleko & Kawamura (2011), which reported on a story retelling task and showed that heritage speakers of Japanese overused nominative particles and underused topic particles. This finding was in line with the hypothesis that narrow syntax (linked to case marking, such as use of the nominative) is more robust than syntactic domains sensitive to discourse considerations (see Chapter 7). The latter involve, for example, the marking of sentence topics, the constituent that an utterance is ‘about’, such as the phrase in italics in *‘The swimming pool maintenance man, have you seen him lately?’* The study we now focus on tries to delve deeper into the question what it is that makes narrow syntax more robust than discourse marking. It tests the general hypothesis that syntactic marking is more robust than discourse marking, as well as specific sub-hypotheses that could explain the differences in robustness between narrow syntax and discourse-sensitive syntax.

#### *Data selection and method*

The umbrella hypothesis examined in this study is that formal features associated with the syntactic component present fewer challenges for bilingual populations than discursive features which also involve pragmatic knowledge. The differences between the older study by Laleko and Kawamura and the study we focus on here concern the expanded methodology and the more precise hypotheses. Rather than only looking at the differences between topic particles and nominative particles, this study divides the overall hypothesis into three sub-hypotheses, namely (1) the integration difficulty hypothesis, (2) the structural complexity hypothesis and (3) the contextual embedding hypothesis; we will be able to explain these below. Rather than just looking at the distinction between topic markers and nominative case markers, the authors look at different uses of these markers and compare predictions, following from the three sub-hypotheses. With regard to nominative

particles, they compare descriptive markers in which the nominative indicates a structural role versus nominatives that encode an exhaustive interpretation, in which all members of the group intended are included. With regard to topic particles, they compare contrastive particles which function to signal a special contrastive relation between two or more elements to thematic topics whose main functions include reference and discourse tracking. The latter group is divided into anaphoric topics that are associated with theme maintenance and generic topics, which create a theme.

*Integration difficulty hypothesis.* This is a popular hypothesis in generative work (Laleko & Polinsky, 2016); it is also referred to as the ‘interface hypothesis.’ This hypothesis states that discourse related domains are more difficult than domains related to narrow syntax because of the integrative nature of discourse related domains: the syntactic knowledge needs to integrate with information from other cognitive domains such as discourse. This causes a higher processing load and that makes interface domains more vulnerable for HL speakers and other bilingual speakers. The integration hypothesis predicts overall vulnerability of topic markers because they are all connected to discourse marking. It does not predict differences between different types of topic markers, because they are all sensitive to the syntax discourse interface. The integration hypothesis predicts more difficulty with the exhaustive marker uses of the nominative than for subject use of nominative, because only the former is connected to discourse information.

*Structural complexity.* Often discourse related structures are also structurally more complex; possibly it is not the integrative nature of discourse phenomena that makes them more difficult to process but their structural complexity. In this light, Laleko and Polinsky (2016) assume that the degree of embedding of a structure constitutes its complexity. The deeper a structure is embedded, given a particular theory of clause structure, the more complex it is and the more difficult it is to learn. Constructions in higher clausal projections are more difficult to learn ( $[_{CP} \dots [_{TP} \dots [_{VP} \dots]]]$ ) than projections lower in the tree. Constructions at the interface make use of higher projections, so both predict difficulty with interface structures, but for different reasons. Like the integration hypothesis, this hypothesis predicts more difficulty with the exhaustive markers. Unlike the integration hypothesis, however, it predicts different outcomes with regard to thematic topics and contrastive topics. Because thematic topics are assumed to be associated with the highest position in a tree structure, they are assumed to be more complex and thus cause more difficulties than contrastive topics.

*Memory load/contextual embedding.* Having to store more information in memory during a syntactic computation can cause learners to have processing

difficulties. Structures at the interface always involve high demands on memory because information from different contexts needs to be combined (contextual embedding), so in a sense it is not clear if it is the integration work or the pressure on the memory load that causes interface constructions to cause more problems for bilingual learners. The general prediction of the two hypotheses is thus identical. However, the contextual embedding hypothesis predicts that generic topics are less vulnerable than anaphoric topics, because anaphoric topics are more dependent on previous discourse than generic topics. This specific prediction does not follow from the integration hypothesis.

This set of predictions is tested by presenting heritage speakers and monolinguals with sentences involving topics and nominatives in all conditions. The idea is that the structures that receive ratings from the heritage speakers that are most unlike the ratings monolinguals provide are the most vulnerable structures. Data were elicited by using Amazon Mechanical Turk (Gibson, Piantados, & Fedorenko, 2011; Sprouse, 2011). Here are three types of topics discussed in the paper:

---

*'The swimming pool maintenance man, have you seen him lately?'*

Anaphoric topics    Yeah, *the maintenance man* just came by with the bill.

Generic topics        Oh, *these maintenance men*, they are never around when you need them.

Contrastive topics    No, but *the gardener* came by yesterday.

---

### Analysis

Heritage speakers and second language learners deviate more from the ratings of native speakers with topics than with subjects. This is in line with all hypotheses that fall under the general umbrella of the vulnerability of interface phenomena. What is not predicted by the integration hypothesis, but what does follow from the contextual embedding hypothesis (and to a certain extent from the structural complexity hypothesis) is the finding that anaphoric topics cause more problems than generic topics. Anaphors require more contextual information than generic topics (while both require integration of information) and are thus expected to be more vulnerable from the perspective of memory load.

Laleko and Polinsky (2016) give a number of suggestions for how to further differentiate between the possible sources of vulnerability for interface-related phenomena. For example, the contribution of syntactic complexity can be examined in different structural contexts within the same interface-level, e.g. by comparing subject and object topics in bilingual speakers of Japanese and Korean. If only the interface matters, than subjects and objects would be predicted to be equally difficult, but if syntactic embedding matters, object topicalization should prove more difficult than subject topicalization, because topic objects require more movement operations and thus are structurally more complex than subject topics.

To disentangle the effects from the integration and contextual embedding the authors looked at topic particles in the main clause and in embedded clauses. Topic particles in embedded clauses have only one interpretation, namely a contrastive one, whereas in main clauses they can be contrastive or thematic. If the contextual embedding hypothesis is correct, the topic particle in main clauses should cause more difficulty than the particle in embedded clauses, because its interpretation needs to be resolved on the basis of more contextual information. The integration difficulty hypothesis does not predict differences between the two contexts. Table 8.1 lists the predictions per hypothesis and the outcomes. The symbol '>' stands for 'more difficult to process' and '=' for 'equal ease of processing'. For example, all three hypotheses predict that topic particles are generally more difficult to process than nominative particles. Only the contextual embedding hypothesis predicts that anaphoric topic particles are more difficult to process than generic topic particles, because anaphoric particles require listeners to use information from prior discourse.

**Table 8.1** The predictions of the umbrella hypothesis (adapted from Laleko & Polinsky (2016))

	Integration Difficulty Hypothesis	Structural Complexity-Hypothesis	Contextual Embedding Hypothesis
<i>Topic particle</i>	anaphoric = generic = contrastive	thematic > contrastive	anaphoric > generic
	☒ not supported	☹ partially supported	☑ supported
<i>Nominative particle</i>	exhaustive > descriptive	exhaustive > descriptive	exhaustive > descriptive
	☑ supported	☑ supported	☑ supported

### *Implications*

This study excels in translating theoretical claims into testable predictions. The tested domains are motivated by theory. The method used is a grammaticality judgement task via the computer. It concerns data on structures that may not surface much in recordings of naturalistic conversation; studies like this help understand why certain structures are vulnerable and further develop testable predictions that could feed back into linguistic theory.

### 8.3 Variationist sociolinguistics

In the 1960s a second important research paradigm developed that is relevant for HL studies: variationist sociolinguistics originally developed by William Labov. It focused on the actual appearance of language in natural speech rather than on underlying structure. This model has led to much careful descriptive research.

#### 8.3.1 Outline

##### *Origin*

Variationist sociolinguistics owes its development to the pioneering work of William Labov (born 1927), who researched sound change first on the island of Martha's Vineyard and then in New York City. One of his best-known books is *Sociolinguistic Patterns* (1972b), which outlines many of the main principles of his approach. Labov is also well-known for his work on African American Vernacular English (AAVE). In more recent times he has studied sound changes across the United States, but many of his former students and close colleagues have also studied other languages such as Brazilian Portuguese and Canadian French and have focused on morphosyntax as well as phonology.

##### *Aims*

The crucial emphasis within variationist sociolinguistics lies on studying the regularity that underlies patterns of variation and change and how this ties in with the social stratification of society. This should make us understand how languages continually change while at the same time remains the stable web that keeps a speech community together.

Originally, variationist sociolinguistics tried to remain close to generative grammar by adopting **variable rules**. These were not unlike the syntactic transformations or the phonological rules of the generative framework. However, the conditions under which they apply could be weighted by quantitative constraints; these also play a role in Optimality Theory as presented below in 8.4. To give an example from phonology, the rule omitting word final /t/ in several languages, including English and Dutch, would apply more frequently if that /t/ is preceded by another consonant than if it is preceded by a vowel, and more frequently if the element is part of another morpheme than if it has a separate meaning (*daft* versus *walk-ed*). These conditions can be formulated as quantitative constraints.

Even though each individual is unique, in variationist sociolinguistics the concept of the **speech community** plays a central role, as does the degree to which speakers can claim group membership. Speech communities impose variable norms on speakers, through the constant interaction within the community. These

communities can be thought of as complex interactive **networks** of speakers. It is obvious that networks are relevant to heritage speakers as well; speakers participate to a greater or lesser degree in various networks. Some speakers maintain tight ethnic networks, while others may form their close ties in the dominant language community.

The variable norms that exist within a speech community also lead to **style shifting**: speakers speaking differently depending on how much attention they are paying to their speech, often sensitive to whether outsiders are present or whether a formal register is called for. This implies that (a) heritage speakers may vary as individuals in the type of language they produce; and (b) when gathering speech data, the style dimension should be taken into account and held constant across speakers or speaker groups.

Methodologically, the choice to focus on speech as it is found leads to a preference for audio recordings of **spontaneous conversations** in naturalistic settings.

### 8.3.2 Case study

We illustrate the variationist method for HL studies with Nagy et al. (2014).

#### *Data selection and method*

Nagy et al. (2014) investigate two variables, Voice Onset Time (VOT) of word initial voiceless stops in stressed syllables with an /a/ or an /o/ in the nucleus, and null subject pronouns in conversational speech. They do so in several generations of heritage speakers of various languages, all living in Toronto. English data are available from the same speakers and an ethnic orientation questionnaire was filled out by all participants to give information on the extent to which they use the HL and how they affiliate (positively or negatively) with their ethnic identity and their HL. The languages investigated are Heritage Cantonese, Italian and for null pronouns and Heritage Russian, Ukrainian and Italian for VOT. The authors use multivariate analysis to interpret the data, considering the linguistic factors that are known to affect VOT or null subjects. For example, for null subjects the person and the number of the subject, the tense and aspect of the verb, the presence or absence of negation, and the presence or absence of preverbal clitics are taken into account. For VOT, the linguistic factors tested are the place of articulation (p/t/k) and the effect of the following vowel (/a/ or /o/). Analyzing the effects of these factors is important because differences per group may be related to different usage contexts, rather than to a difference in VOT or the use of null subjects per se. Note that this is always important when naturalistic data are used, because contexts can differ more widely than in experimental settings.

### *Analysis*

Taking the analysis of the Cantonese data as an example, the proportion of null subjects in the raw data shows that first generation Cantonese speakers use more null subjects than second or third generation heritage speakers. Nagy et al. (2014) state that although the raw data percentages indicate a significantly higher use of null subjects in the first-generation group, this difference disappears when the linguistic context is taken into account: first generation Cantonese speakers happen to use more linguistic contexts that favor null subjects, i.e. contexts in which the subject has already been introduced as a topic of discussion. When context is considered, they find no difference between generations of speakers. Thus, the difference between the generations appears to be pragmatic rather than grammatical.

Apart from looking at the mere presence or absence of null subjects, they investigated the effect size, if any, of the various conditioning factors and their weight and ranking in the different generations. Again, the results were that generations did not differ from each other: the same linguistic factors played a role in the choice for a null subject and their effects were the same size and were ranked similarly in both generations.

They did find two small differences, however. First, in the Russian data the person hierarchy in favouring null subjects differed significantly between heritage speakers and first-generation speakers. Whereas the latter favoured null subjects most in third person contexts, followed by second and third, the heritage speakers favour null subjects in the order  $2 > 1 > 3$ . Heritage speakers also used more null subjects in negated sentences. These differences did not relate to the frequency of use of Russian and they could not be explained by the ordering of the conditioning factors in English. Nagy et al. (2014) interpret the absence of a relation with language choice as an indication that it cannot be incomplete acquisition that explains the difference, since proficiency in English would probably be related to language choice patterns. The absence of similarity with the conditioning factors in English may be a sign that cross-linguistic transfer cannot explain the data either. In short, they state that there are hardly any differences between generations and that where differences are attested these are language internal differences rather than ones caused by attrition or incomplete acquisition.

For the VOT they find that while the three generations of speakers of Italian in Toronto do not differ from homeland speakers, the Ukrainian and Russian speakers tend to get longer VOT's per generation, thus becoming more similar to English VOT's. Since not all heritage groups lengthen their VOT to become more English-like and since there is no relation with scores in the ethnolinguistic identity questionnaire, Nagy et al. again suggest that there is no evidence for incomplete acquisition. There is change, but this change could very well be regular change rather than contact-induced change.



### *Implications*

Different from other heritage studies, Nagy et al. (2014) report stability in the domains they studied and show that the few differences that they did find did not relate to language use patterns, making incomplete acquisition less likely as an explanation. They argue for the possibility of regular language change rather than contact-induced change, a conclusion a number of other studies in the variationist paradigm have drawn as well (Poplack, Zentz, & Dion, 2012; Torres Cacoullous & Travis, 2015).

The data-gathering methods in variationist studies may help heritage speakers to perform to their abilities. Efforts were taken to make the speakers feel at ease, the first half hour was not recorded to give speakers time to warm up and the investigators were heritage speakers themselves and they knew the people they interviewed. Taking conditioning factors into account when interpreting the data, helps to distinguish “real” differences, from the effect of the conditioning factor. The study shows the relevance of naturalistic data. Experimental data may cause heritage speakers to perform below their abilities. A drawback of the use of naturalistic data is that speakers may avoid constructions that they find difficult. If the construction does not occur, it is difficult to know whether this was just coincidence or the result of avoidance.

## 8.4 Optimality theory

Optimality Theory or OT is a linguistic model that stresses that the actual forms of language result from the interaction of potentially conflicting constraints on the speaker’s linguistic output. Differences between languages result from different rankings of these constraints.

### 8.4.1 Outline

There is not much work on HL within this paradigm, but it has explanatory potential, if one subscribes to its basic premises.

### *Origin*

Optimality Theory arose as a model around 1993, when the computational linguist Paul Smolensky and the phonologist Alan Prince joined forces. Their joint work appeared in book form in 2004 (Prince & Smolenski, 2004) but in the preceding decade the work done on OT was reported through the web-based *Rutgers Optimality Archives*. While the primary domain to which OT has been applied is phonology, there have been extensions to morphosyntax (Bresnan, 2000; Grimshaw, 1997; Kuhn, 2001; Bousquette, Putnam, Salmons, Frey, & Nützel, 2016) and

semantics (Blutner, de Hoop, & Hendriks, 2006). Legendre et al. (2016) present work on ‘bidirectional OT’, in which constraints on production and comprehension can ‘bidirectionally’ influence each other. Production is seen as a step from meaning to form (morphology and syntax) and comprehension as a step from form to meaning (semantics and pragmatics).

### *Aims*

The primary aim of OT is to explore the ways in which very general constraints on sound patterns (e.g. that syllables ending in a vowel are preferred over those ending in a consonant) interact, especially how one constraint can be overridden by a conflicting constraint.

While the ranking is language specific and accounts for variation between languages, OT assumes that there are **Universal constraints** that are operant in all languages, as illustrated in the case study presented below.

Recent work on **stochastic variation** in OT (Anttila, 1997; Boersma & Hayes, 2001; Nagy & Reynolds, 1997) makes this framework a promising one for the development of a formal analysis of variable data, where quantitative differences can be taken into account. The stochastic approach allows for quantitative weighing of different constraints; this too will be illustrated below.

For HL studies this approach is characterized by the idea that the changes observed in HLs tend not to be unique to these, but the result of global tendencies shared by all languages, but possibly more pronounced in heritage speakers (see also Bousquette et al., 2016).

### 8.4.2 Case study

We illustrate the use of a stochastic OT model in HL studies with the work of Koontz-Garboden (2004) on the variable use of Spanish progressive aspect.

#### *Data selection*

Koontz-Garboden (2004) reanalyzes data gathered by Klein (1980) using a stochastic OT model. Klein (1980) described the use of synthetic and analytic verb forms in monolingual Spanish and in the Spanish of Spanish-English bilinguals. The analytic form as shown in A is consistent only with progressive meaning. The Spanish synthetic form as shown in B is consistent with two readings: a habitual one and a progressive one.

#### A. ANALYTIC FORM

*Mira, está saliendo ahora el sol.*

look is coming-out now the sun.

‘Look the sun is coming out now’

## B. SYNTHETIC FORM

*Mira, sale            ahora el sol.*  
 look comes-out now the sun.  
 ‘Look the sun comes out now’.

The analytic and the synthetic form thus share a progressive meaning. This overlap in meaning between the synthetic and the analytic forms is absent in English where the synthetic form is only used for the habitual and where all progressives are encoded by the analytic form. Klein (1980) shows that bilinguals move towards the English system in the sense that they encode more progressives with an analytic form, so they are more likely to use A in progressive contexts than monolinguals. Moreover, she shows an effect of context: if the context does not hint at a progressive reading, speakers are more likely to choose the analytic form than if the context already indicates a progressive reading. The effect of context is significant only in the monolingual group, but both groups show the effects as shown in Tables 8.2 and 8.3.

Table 8.2 Utterances used progressively with less context

Less context	Monolingual	Bilingual
Synthetic	13.6% ( $n = 24$ )	5.6% ( $n = 13$ )
Analytic	86.4% ( $n = 152$ )	94.4% ( $n = 221$ )

Table 8.3 Utterances used progressively with more context

More context	Monolingual	Bilingual
Synthetic	60% ( $n = 39$ )	9.4% ( $n = 5$ )
Analytic	40% ( $n = 26$ )	91.6% ( $n = 48$ )

The effects of context and of bilingualism are both quantitative effects: bilinguals are significantly more likely than monolinguals to use the progressive form and monolingual speakers are more likely to use the analytic form when there are no contextual clues to indicate the progressive reading, but these effects are never categorical.

*Analysis*

Koontz-Garboden (2004) argues that the way variation is treated in OT makes this framework a promising one for studying quantitatively conditioned variation. To Koontz-Garboden the OT-approach allows room for variation while also making use of universal formal constraints.

Koontz-Garboden suggests four sets of partly competing constraints. The first set relates to ‘faithfulness’ and requires the overt expression of all aspectual distinctions. The second relates to markedness: typologically more marked categories are more likely to be overtly encoded than unmarked ones. The third set relates to economy: as speakers we use what takes the least cognitive effort. The fourth set relates to pragmatics: we encode information with maximally distinctive forms if the context fails to provide enough cues to the correct interpretation. Koontz-Garboden shows that the two speaker groups rank these constraints differently. In other words, while both groups have the same universal constraints at work, the weight of each differs per group and the different ranking accounts for quantitative differences between the two groups.

Before we turn to the actual analysis, let us briefly look into their content and how they work. The faithfulness set contains only one faithfulness constraint, namely:

FAITHFULNESS CONSTRAINTS

MAX: input attribute/value pairs are lexically realized in the input

The faithfulness constraint supports the overt encoding of aspectual distinctions. The markedness set contains two constraints that go in the opposite direction: do not encode certain distinctions.

MARKEDNESS CONSTRAINTS

\*PROG: \*PROG + & \*HAB –

\*HAB: \*PROG – & \*HAB +

\*PROG goes against the overt expression of the progressive, \*HAB goes against the overt expression of the habitual. If the MAX constraint is ranked highest all distinctions will be overtly encoded. If, on the other hand, \*PROG is marked higher than MAX, the progressive will not be overtly encoded, because it goes against \*PROG.

Koontz-Garboden also posits two economy constraints based on the work of Sells (1997, 1998), namely:

ECONOMY CONSTRAINTS

\*X<sup>0</sup>

AFFIX<sup>3</sup> >> AFFIX<sup>2</sup> >> AFFIX<sup>1</sup>

The first constraint (\*X<sup>0</sup>) implies that languages strive for the smallest syntactic structures. The idea is that a synthetic structure has less syntax than an analytic structure in A and thus the analytic structure in A violates \*X<sup>0</sup>, whereas the synthetic structure does not. The second constraint focuses on morphological richness and states that languages want as few affixes as possible. Languages differ, however, in whether it is morphological or syntactic economy that is more important. They can also differ in their cut off point: some languages may want to avoid all affixes,

other languages resort to analytic structures when the alternative has more than two affixes, others when there are more than three affixes etc.

PRAGMATIC CONSTRAINTS

\*INDET

Finally, the constraint \*INDET (for ‘indeterminacy’) requires that if the context does not make clear what the interpretation of a form is, the form has to be as explicit as possible. So when there is little or no context suggesting a progressive reading, \*INDET pushes the use of the overt progressive form.

Koontz-Garboden explains the difference between stochastic OT (Boersma & Hayes, 2001) from standard OT as follows. First, in contrast to standard OT where constraints are ranked in a list, constraint ranking in stochastic OT is along a continuous ranking scale, so that constraints can be closer to, or farther apart from one another, as illustrated in Figure 8.1.

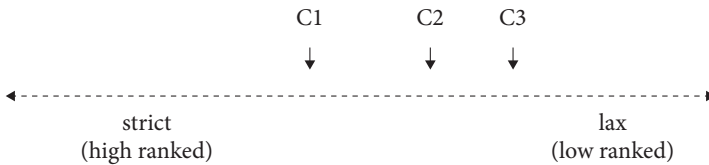


Figure 8.1 Continuous ranking scale (adapted from Koontz-Garboden, 2004)

Figure 8.1 shows that constraint 1 is further apart from the other constraints, whereas constraints 2 and 3 are relatively close to each other.

The second innovation is ‘stochastic candidate evaluation’, as illustrated in Figure 2. When we adopt this model, ‘the position of each constraint is perturbed by a random variable so that the relative rankings of constraints can be disturbed, with the possible degree of disturbance following a normal distribution.’ The more the constraints overlap in their distribution, the more likely it is that their relative rankings can be disturbed on a particular evaluation. Changes in relative rankings on particular evaluations may lead to variation in the output.

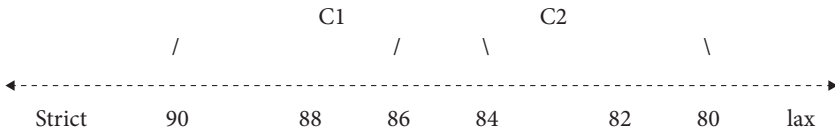


Figure 8.2 Relative constraint rankings vary (adapted from Koontz-Garboden, 2004)

Let us now move to Koontz-Garboden’s analysis of the Spanish data. Koontz-Garboden fed Klein’s data into UGLA (a general learning algorithm developed by Boersma & Weening, 2002). The learning algorithm assigned the values to all constraints listed in Table 8.4:

**Table 8.4** Bilingual grammar compared to monolingual grammar

Bilinguals constraint	Value	Monolinguals constraint	Value
*HAB	124.726	*HAB	126.262
*AFFIX2	118.604	*AFFIX2	120.655
MAX	110.614	*INDET	111.068
*X <sup>0</sup>	106.171	*X <sup>0</sup>	108.276
*INDET	83.125	MAX	107.459

The main differences between the bilingual and the monolingual grammar are the ranking of \*MAX and \*INDET and the ranking between \*MAX and \*X<sup>0</sup>. Because \*MAX is ranked so highly for the bilinguals they are more likely than monolinguals to encode progressives explicitly and less likely to use the underspecified imperfect form. The low ranking of \*INDET indicates low pragmatic sensitivity in the bilinguals.

OT-analyses contain so-called tableaux: a table with possible input forms and a list of constraints where violations are marked with \* and where fatal violations (that is violations that are so severe that a form is not allowed to surface) with \*!. The form that is chosen is marked with a pointing finger (☞). Let us consider these forms in tableau form for one context, namely the monolinguals who want to express progressive in a context with few cues towards this progressive meaning. A thing to keep in mind is that the tableau contains universal constraints. Some of these constraints may not connect to actual surface forms in Spanish, but they still appear because it is assumed in OT that all languages have all these constraints. For example, the tableau contains the options ‘synthetic progressive’ and ‘analytic imperfective’ although Spanish does not actually show the existence of such forms.

**Table 8.5** Tableau for monolinguals in a less context situation

PROG	*HAB	*AFFIX2	*INDET	*X <sup>0</sup>	MAX
☞ Analytic, PROG				**	
Synthetic, PROG		*!		*	
Synthetic, IMP			*!	*	*
Analytic, IMP			*!	**	*

In a ‘less context situation’ avoiding indeterminacy is relatively important, so \*INDET favors overt progressives (PROG) over the underspecified imperfectives (IMP). Because \*AFFIX2 is always ranked higher than \*X<sup>0</sup> in Spanish an analytic form is chosen over a synthetic form. In stochastic OT, the distribution between constraints is important. The distance between \*AFFIX2 and \*X<sup>0</sup> is so large that the analytic form will always be chosen to express progressivity overtly.

This choice is categorical. The distance between \*INDET and \*X<sup>o</sup> is smaller and they partly overlap. This means that in most situations with less context the overt progressive form will be chosen but in some cases \*INDET will be overruled by X<sup>o</sup> and the synthetic imperfect will be chosen for the progressive meaning. The choice between the synthetic form and the analytic form is thus variable. If there is more context for interpretation, the \*INDET constraint is less important and \*X<sup>o</sup> is likely to win out and yield more synthetic imperfectives. A difference between monolingual and bilingual speakers is the ranking between MAX and \*X<sup>o</sup>. MAX requires the explicit encoding of progressive meaning. This could explain the high incidence of the use of the progressives in bilinguals. The pragmatic constraint \*INDET is less important in bilinguals than in monolinguals.

### *Implications*

Koontz-Garboden propagates the use of OT as a means to deal with variation while at the same time keeping track of formal universals. It is a way to bridge the strengths of the variationist and the generative model. His model has implications for cross-linguistic variation and for language contact studies. With regards to cross-linguistic variation: Koontz-Garboden shows that X<sup>o</sup> and AFFIX<sup>3</sup>>>AFFIX<sup>2</sup>>>AFFIX<sup>1</sup> are in competition with each other. X<sup>o</sup> should be positioned somewhere in relation to AFFIX<sup>3</sup>>>AFFIX<sup>2</sup>>>AFFIX<sup>1</sup>. It follows from his assumptions that it is impossible to have an analytic form to express imperfectivity (the basic form with the least affixes) and a synthetic progressive form (a more derived form which will always have one affix more than the imperfect). With regards to language contact, it may be possible to find generalizations in the competition between constraints in language contact situations as shown for example by Kusters (2003). It is possible that language contact situations tend to demote pragmatic constraints (cf. Silva-Corvalán, 1994, 2015) or that they promote morphological economy above syntactic economy (Kusters, 2003).

## 8.5 Usage-based models

Finally, HLs are also studied using from a usage-based perspective, the dominant approach in Cognitive Linguistics and Construction Grammar.

### 8.5.1 Outline

#### *Origin*

It is probably fair to say that the first articulation of a usage-based approach was by Ronald Langacker (born 1942) in his 1987 book *Foundations of Cognitive Grammar*,

*Volume 1*; see Langacker (2008) for a recent update. Also influential is the work by Joan Bybee (born 1945), who further expanded this model; cf. Bybee (2010).

### *Aims*

Central to all these models is the idea that linguistic knowledge is not viewed as innate knowledge of a fixed set of rules, but as emergent mastery of many items and constructions and of how they can be combined, all on the basis of experience ('usage') and innate cognitive skills. Contrary to the generative tradition, the properties of language structure, change and variation are explained out of out of principles of **general cognition** rather than specifically linguistic cognitive properties, and with reference to the different patterns of usage that hold for individuals and social groups.

The basic elements for linguistic description are linguistic **units** or **minimal pairings** of **sign** and **meaning**, such as words, collocations, and constructions.

Very important is the role of **frequency of use**, since this helps determine the degree to which units are mastered by individual speakers or **entrenched** in their minds. At the level of communities and languages, units emerge through **conventionalization**, as the same unit becomes entrenched in the minds of individuals that often interact as they are part of the same speech community.

Frequent use and strong entrenchment of an item may lead to **grammaticalization**, the process by which units gradually develop from content words to function words in the language usage system. Doing so they may undergo phonological reduction and become shorter and less marked. They may also lose part of their more concrete specific meaning and become more abstract in their meaning.

While in the generative tradition there is a specific role for the child learning a language as the creative transformer of the system based on adult input, in usage-based models there is no special role for the child. In fact, most changes will emerge through ordinary **adult usage**. This has important implications as well for HL studies, since it is usage rather than acquisition and age of onset that are claimed to play the key role, though these factors may, of course, affect patterns of usage.

### 8.5.2 Case study

A good example of this approach is the work reported in Doğruöz & Backus (2009) and Backus, Doğruöz, & Heine (2011).

#### *Data selection and methods*

Doğruöz and Backus (2009) investigated informal speech in a corpus of spoken Turkish in the Netherlands and in Turkey. Native speakers listened to samples of Dutch Turkish (or 'NL-Turkish') speech to confirm or disconfirm that constructions



that sounded “unconventional” to the researchers indeed sounded unconventional to them too (though not necessarily ungrammatical). They investigated to what extent certain constructions are specific for Turkish in the Netherlands or differ markedly in frequency from the Turkey Turkish (or ‘TR-Turkish’) corpus compiled as control data. In this way they found several constructions that appeared more frequently or only in NL-Turkish. We discuss two examples:

*Semantic intensification of negative adjectives.* The dominant language Dutch can use the adjective *erg* (‘bad’) to intensify the meaning of another adjective as shown below in example (4). This use is calqued into Turkish.

NL-Turkish	<i>Bir sene</i>	<i>acaip kötü</i>	<i>sıcak-tı</i>	<i>burası</i>	
	one year	very bad	warm-PST	here	
Dutch	<i>Er was een jaar toen</i>	<i>was</i>	<i>het heel erg warm.</i>		
	there was one year then	was	it Very bad warm		
	‘One year it was incredibly warm here.’				

*Specific interpretation of the use of an indefinite determiner.* Turkish has a weakly grammaticalized indefinite marker (derived from the numeral *one*). This marker can only be used to indicate indefinite non-specific nouns. For example, it is possible to say *akustik bir gitar* (‘acoustic one/a guitar’) in saying something like ‘an acoustic guitar is better than an electric one’, but it is impossible to use the marker *bir* to indicate an indefinite but specific guitar (e.g. ‘My neighbour had a guitar that I used to play’). In the last utterance the speaker knows which specific guitar he’s talking about, and as a result Turkish would have a bare noun. In the dominant language Dutch, however, an indefinite marker would be used. Turkish speakers in the Netherlands extend the use of *bir* to specific-indefinites.

### *Analysis*

The notion construction plays a central role in the analysis of Doğruöz & Backus (2009). With construction they mean a structural unit that is partially filled lexically. For example, [*erg* A] ([‘very A’], literally [‘bad A’]) is a construction in Dutch to express an intensified meaning. The claim is that heritage Turkish in the Netherlands is developing Dutch-influenced constructions with the help of perceived semantic equivalence between Dutch and Turkish lexical items. For example, *kötü* is mapped onto Dutch *erg* ‘bad’ and *bir* is mapped onto Dutch *een* ‘a’. Thus, Dutch grammar is not copied as such, but complex lexical units are translated into Turkish.

### Implications

Doğruöz and Backus (2009) show that we can look at HL usage and competence at a constructional level, combining lexicon and grammar. They show that differences between native speakers and heritage speakers are not absolute but gradual. Some constructions are less likely to occur in TR-Turkish than in NL-Turkish, but they are not impossible. Speakers can differ in the frequency in which they use certain constructions and this shift in frequency can be contact-induced. Backus et al. (2011) relate the synchronic variation they attest to diachronic cases of contact-induced change. They show that if language contact lasts long, i.e. centuries, it can profoundly change the typological profile of the language. They reference languages that were once without articles but after many centuries of language contact developed them, e.g. Sorbian (under German influence). They suggest that such typological shifts begin as changes at the level of individual constructions.

## 8.6 Summary discussion: Integrating the models

We have treated these different grammatical models or research paradigms as separate strands of research, and this seems necessary given their very different assumptions. However, there are also points of contact between these approaches, and individual authors can combine and highlight different sets of assumptions. For example, the work by Pires & Rothman is generative in nature but also emphasizes the role of input factors and the work by O’Grady et al. (2011) departs from a usage-based perspective, but stresses that input is not all. The approaches can be arranged on a broad continuum with two poles, roughly as in Table 8.6:

**Table 8.6** Continuum between different models

Formal grammar	◀ Stochastic OT grammars variationist models	▶ Usage based grammar
Innate domain specific principles	◀ Interaction between different kinds of principles	▶ Domain- general cognitive principles
Competence and judgments	◀ Structured ecologically valid experiments	▶ Performance and spontaneous usage
Frequency irrelevant	◀ Frequency interacts with other cognitive factors	▶ Frequency crucial

The challenge is to see where these models connect, and for which type of phenomena which model provides the most insightful explanation. As formal grammar models sometimes start making use of quantitative data and usage-based grammar models are being formalized, what seemed to be irreconcilable theoretical differences may start vanishing.



# Language processing in multilingual speakers

## 9.1 Introduction

In this chapter we discuss HL speakers from the perspective of multilingual processing and try to bring to bear some of the findings of experimental psycholinguistics on these speakers. As discussed in for example Chapters 4 and 6, many HL speakers display (1) processing difficulties, such as speaking slower and more hesitantly than non-HL speakers, and (2) forms of ‘interference’ or ‘co-activation’ of the dominant non-HL language, such as language transfer, contact-induced linguistic innovations and/or codeswitching (see Chapter 4). Both these aspects of being an HL speaker can be studied from a psycholinguistic perspective.

But what is a psycholinguistic perspective? The main aim of psycholinguistics is to obtain insight into the mental processes and knowledge structures underlying language use: language production, language comprehension, language acquisition, and, of course, the interplay between these aspects of language use (de Groot, 2011; Gaskell, 2007; Dell & Chang, 2014). Thus, the object of study in psycholinguistics is not language per se, but the language user *in relation to his/her cognitive system*.

There is a vast research literature that needs to be considered when we address the psycholinguistics of language processing in heritage speakers. An important reason for this is that, within the field of psycholinguistics, no principled distinction is made between HL speakers, bilinguals, multilinguals and second language learners. That is, HL speakers, like second language learners, can simply be seen as a special kind of bilingual or multilingual speaker, based on which we can study the cognitive mechanisms of learning and speaking more than one language (see also Chapter 1 of this book for more information on definitions of heritage speakers and the way in which these definitions are shaped by the focus of interest of a research field). The term ‘HL speaker’ is therefore not used much in the psycholinguistic literature. Instead, psycholinguistic research tends to focus on a wide variety of speaker-specific background variables like language dominance, age of acquisition, language history, and proficiency to analyse the extent to which these variables influence language processing. HL speakers can be regarded as a group of bilingual speakers in which these background variables have specific values (e.g., becoming bilingual at a young age, relatively low proficiency in the HL, etc.).

Consistent with the terminology from the psycholinguistic literature, in this chapter, we will base most of our discussion on theories and research on bilingual and multilingual language processing, in which the term ‘bilingual’, like ‘multilingual’, is loosely used as referring to a speaker of two or more languages, which includes HL-speakers (see Appel & Muysken, 1987; Bhatia & Ritchie, 2008; Grosjean & Li, 2013; Kroll & de Groot, 2005; Li Wei, 2007, for overviews). This does not mean, of course, that differences between different kinds of bilingual speakers are ignored in psycholinguistics. Indeed, variation between different bilingual speakers is an important issue in this field, but this variation is mostly and increasingly operationalized in terms of a combination of speaker-specific background variables rather than group comparisons.

How can we characterize different kinds of bilingual / multilingual speakers (including HL speakers) from a psycholinguistic perspective, and what are important speaker-specific factors that may influence linguistic choices and processes in the various kinds of multilinguals?

In Section 9.2 we consider core theoretical notions of language processing in bilingual speakers, focusing on the work inspired by researchers such as Levelt (1989) and Pickering & Garrod (2004). Section 9.3 discusses core findings of language processing in multilinguals, focusing on three domains: cross-language interactions (9.3.1), processing differences between bilinguals and monolinguals (9.3.2), and language switching and inhibition (9.3.3). In Section 9.4, we discuss various factors that influence language processing in bilinguals. Section 9.5 presents a psycholinguistic perspective on the issue of age of acquisition (see also Chapter 6). Section 9.6, finally, presents some conclusions and issues for further research.

#### Main goals of this chapter

- To present the core notions in language processing of multilingual speakers
- To survey the main findings in processing research relevant to HL speakers

## 9.2 Core notions in research on language processing in bilingual speakers

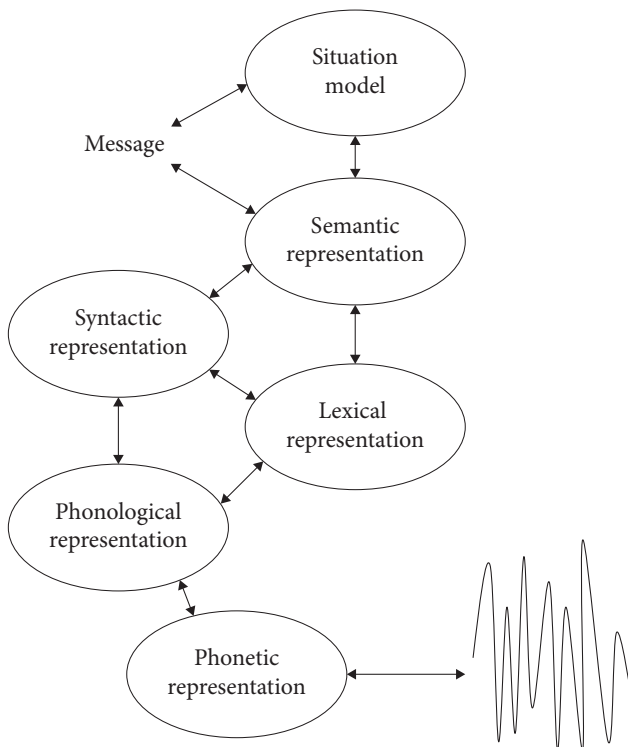
One of the major research themes in studies on multilingual language processing is how we can account for the fact that bilinguals are in principle able to keep their languages apart in language production when needed, but are also influenced by their knowledge of multiple languages when producing language (cf., de Bot, 2004; Kootstra, van Hell & Dijkstra, 2009; Myers-Scotton, 2006; Poulisse & Bongaerts, 1994). These questions are often investigated in terms of *models* of the bilingual language user. Most of these models are adaptations of monolingual ones (e.g.,

de Bot, 1992, is an adaptation of Levelt, 1989; Hartsuiker & Pickering, 2008, is an adaptation of Pickering & Branigan, 1998). To understand the bilingual models, it is first important to introduce the core notions of the monolingual models.

According to most models (see e.g., Brown & Hagoort, 1999, for an overview), language production and comprehension takes place by means of the *activation* of linguistic ‘representations’ (pieces of information) in our mind, at multiple *processing levels*. Language production, for example, is typically seen as a multi-staged process from conceptualization to articulation (e.g., Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; though see also Pickering & Garrod, 2004): Based on the discourse situation at hand, a speaker conceptualizes a preverbal message that s/he wishes to communicate. The speaker then activates and selects the appropriate words and sentence structure associated with the meaning of this message, and subsequently retrieves the appropriate phonological forms to turn words and sentence structure into a phonetic pattern that can be articulated. Altogether, this process involves socio-contextual, semantic, syntactic, phonological, and phonetic processes that are all tuned to one another (see Figure 9.1). Language comprehension can be seen as a similar process of activating and selecting linguistic representations, but now the order of the process is reversed: A listener comprehends a stream of sound by activating and selecting the right phonemes, words, grammatical structure, and meaning associated with this stream of sound, which is then interpreted on the basis of the discourse situation at hand.

Possibly the most important mechanism to understand in these models is the mechanism of *activation* of linguistic representations. It is assumed in many models that linguistic representations (e.g., words, constructions, morphemes, etc.) have a certain *level of activation*, which is among other things based on frequency of usage: the more frequently a word or construction is used, the higher its level of activation in the mind will be (see also the discussion of the usage-based approach in Chapter 8). This level of activation determines the ease with which this specific linguistic representation can be selected during language processing. Thus, when words and/or constructions are used during linguistic tasks such as speaking, listening, or reading, these words and/or constructions may become active and available faster or slower, depending on their frequency of usage (see also Kootstra, Dijkstra, & Starren, 2015).

An important aspect of the activation metaphor is that linguistic representations are not activated and selected in a ‘vacuum’; rather, linguistic representations in the mind are supposed to be interconnected with each other, like neurons in a neural network (this is where the term ‘connectionism’ comes from). What is interesting about neural networks is that they often form clusters of highly interconnected neurons with specific functions. When this function needs to be executed, neurons in that cluster will be activated, and most importantly, the connections



**Figure 9.1** A schematic presentation of the processing levels involved in language production and comprehension (loosely based on Levelt, 1989; Levelt et al., 1999; Pickering & Garrod, 2004)

between these neurons will spread activation across interconnected neurons, leading to a co-activation of neurons, which facilitates the efficiency with which the function is executed. With respect to language use, one can imagine a neural network of interconnected linguistic representations based on shared functions or characteristics, which will therefore co-activate each other. Thus, for example, the words ‘table’, ‘chair’, and ‘desk’ are all nouns representing furniture, and may therefore probably co-activate each other. Similarly, the words ‘racket’ and ‘rocket’ may also co-activate each other, because of their shared form characteristics. Finally, spreading of activation across linguistic representations will also take place with representations that often co-occur. Thus, for example, activation of the ditransitive verb ‘to promise’ in sentences like ‘I promised her a kiss’, which selects for both an object (the kiss) and a recipient (her) and is claimed to also lead to the activation of the ditransitive syntactic structure (which is a linguistic representation at the syntactic level), which in turns activates other ditransitive constructions ‘I gave her a book’. Similarly, activation of the intransitive construction ‘I slept’ may lead

to the activation of the intransitive syntactic structure (i.e., agent-verb), which will then activate other intransitive constructions like 'I sneezed'. This interconnectivity between linguistic representations facilitates the staged process of language production and/or comprehension. Researchers differ from each other on the question whether activation proceeds in stages or takes place instantaneously at all levels.

Now that some important notions from the monolingual literature have been introduced, it is time to talk about bilingual language processing. How do the general notions from the monolingual literature apply to bilingual language processing? Again, the metaphor of *activation* of linguistic representations in an *interconnected network* is of central importance.

Let us first discuss activation level. As introduced above, the level of activation of linguistic representations (and therefore the ease with which these representations can be used) differs depending on, among other things, frequency of usage. It is not unlikely that the frequency of usage of linguistic units in a specific language is lower in bilinguals than in monolinguals, for the simple reason that bilinguals must divide their language use over two (or more) languages in their daily lives, leading to a relatively low frequency of usage per language, compared to monolinguals (e.g., Gollan, Montoya, Cera, & Sandoval, 2008). This will also be the case for HL speakers. Therefore, the level of activation of units in both a bilingual's languages may be lower than that of monolinguals, leading to, for example, processing difficulties as reflected in lower speech rate and/or hesitant speech in a language used less frequently. This is indeed what has been found in heritage speakers: HL-speakers tend to speak more slowly than non-HL speakers, often with pauses and word finding difficulties (see Chapter 6). The goal of psycholinguistic research on bilingualism is to explore which factors influence this level of activation of linguistic representations of both languages (see Sections 9.3. and 9.4. for some core findings in this respect).

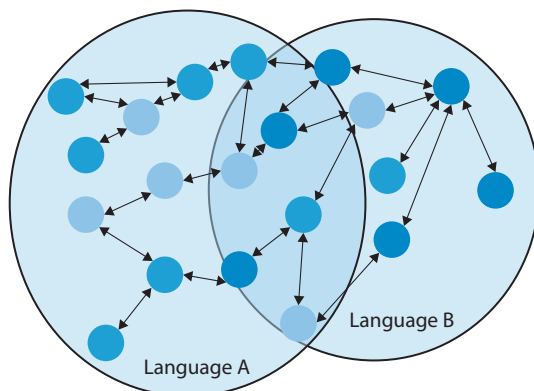
We now turn to the notion of an interconnected network of linguistic representations. A major question in the psycholinguistics of bilingualism (see e.g., Kroll, Dussias, Bice, & Perrotti, 2015) concerns the extent to which a bilingual's languages are shared (i.e., one big network of linguistic representations from both languages) or separate (i.e., multiple networks of linguistic representations for each language separately). If both languages are shared in one big interconnected network in the mind, then it is likely that the activation of linguistic representations from one language may also lead to the activation of linguistic representations with similar characteristics or functions from the other language (i.e., co-activation across languages, or *cross-language activation*). This would then be reflected in cross-linguistic influences in the language use of bilinguals. If, however, both languages are represented as separate networks, then cross-language influences caused by



cross-language activation are relatively difficult to explain. Therefore, research on cross-language influences in bilingual language use is illuminative for theories of bilingualism, but also, at a more general level, for insight into the fundamental interactivity of the human mind (Kroll, Bobb, & Hoshino, 2014). Many studies in this area therefore focus on the question how and to what extent bilinguals may differ from one another in the degree to which their languages are stored and accessed separately, and which variables influence these processes.

To sum up, the level of activation of linguistic representations in a bilingual's languages may differ as a function of the frequency of usage, and these linguistic representations may or may not be interconnected in the bilingual mind. Phenomena like codeswitching and cross-language influences in the speech of both bilingual children and adults, including HL speakers, (e.g., Jarvis & Pavlenko, 2008; Muysken, 2000; Unsworth, 2013; other chapters in this book) are strong indicators, however, that the linguistic representations of the two languages are interconnected. These cross-linguistic interactions reveal that, as explained in the previous paragraph, linguistic representations can be activated in parallel, leading to cross-language activation.

The abovementioned metaphor of the neural network with clusters of sub-networks is especially useful when we are dealing with bilingualism. How is it possible that on the one hand people can keep their languages apart but on the other hand experience considerable cross-linguistic interactions? The idea is that each language constitutes a sub-network or cluster in the complete network of linguistic representations, as presented in Figure 9.2 (de Bot, Lowie, & Verspoor, 2005; Kootstra, van Hell, & Dijkstra, 2009; this idea is equivalent to the idea that linguistic representations are 'tagged for language', e.g., Kootstra & Rossi, 2017; see Putnam, Carlson, & Reitter, 2018, for a related account of an integrated multilingual architecture). If a person then wants to speak a specific language, that person can do so by activating linguistic representations from the sub-network of that specific language. But these sub-networks also always form links with sub-networks from other languages (see Figure 3), because of which bilingual speakers will be influenced by their other language from time to time. In this way it is possible to account for the fact that most bilingual speakers are very capable of speaking one specific language only (an exception are some cases of elderly bilinguals suffering from Alzheimer's), but at the same time may be influenced by their other language while codeswitching or even in a purely monolingual mode.



**Figure 9.2** An example of sub-networks of languages within a general network of linguistic representations. The dots represent linguistic representations, the colors of the dots represent different levels of activation of different linguistic representations, and the arrows represent the (bidirectional) interconnections between the linguistic representations. This figure shows that the sub-networks of languages may form separate networks but are also linked to each other. Based on de Bot, Lowie, & Verspoor (2005)

### 9.3 Core findings

Most psycholinguistic research is done by means of experiments. As noted in Chapter 5, experiments are studies in which data are collected under standardized circumstances (often a computer task done by a large number of bilingual participants) in which specific variables of interest are either controlled or manipulated. The advantage of experimental methods is that they enable researchers to test a systematically manipulated sample of linguistic materials on a large sample of participants in a controlled situation, thus allowing for quantitative inferential analyses. A disadvantage of experimental methods is that they typically generate rather artificial data (recall the discussion of ecological validity in Chapter 5) and thus run the risk of compromising the phenomenon under study (Gullberg et al., 2009). Still, experimental data can provide valuable information, especially when they are considered in combination with more natural forms of data (like natural speech recordings). The findings presented below are mostly based on experimental data.

We focus on three types of findings in this chapter:

- Cross-language interactions during language use
- Processing difficulties in bilinguals' language use
- Language switching and inhibition

### 9.3.1 Cross-language interactions

A key finding in psycholinguistic research on bilingualism is that it is nearly impossible for a bilingual to completely switch off one of his/her languages (see e.g., de Groot, 2011; Grosjean, 2008; Kroll, Bogulski, & McClain, 2012, for reviews). Psycholinguistic research has provided evidence that most sorts of language information, for instance, with respect to pronunciation, words, word formation and grammar, meaning and information structure in bilinguals are sensitive to cross-language interactions. Below, we focus on findings at the word and sentence levels.

Consider first the word level. In many studies the question is addressed how words are processed that have a form overlap between several languages. Words that resemble each other in different languages in their orthography or pronunciation, but have different meanings are called ‘false friends’, or also interlingual homographs (writing) or interlingual homophones (pronunciation). Examples are the homographs COIN (English) – COIN (French ‘corner’). Words that are similar in form and in meaning are called ‘cognates’ in psycholinguistics; such words often have a common etymology or have been borrowed into one of the languages. So-called ‘identical cognates’ are translation equivalents that are orthographically identical and often phonologically very similar across languages, like FILM (English) – FILM (Dutch, French, Polish, and many other languages) (Peeters et al., 2013). ‘Non-identical cognates’ are less similar in their orthography and phonology, such as TOMATO (English) – TOMAAT (Dutch) – TOMATE (French, Spanish).

These ‘bilingual’ words are very handy for studying cross-language interactions in the bilingual mind, often with very simple computerized experiments, such as lexical decision tasks. In this task, bilingual participants are presented with a list of words and pseudo-words that are presented one by one on a computer screen. For each word the participants must indicate as quickly and accurately as possible whether they think the word is an existing word or not. Speed and accuracy of performance on this task are recorded and analyzed as measures of language processing. In lexical decision experiments on cognates and false friends, participants are presented with these types of words, in addition to control words that do not overlap between the languages. The logic of these experiments is quite simple: if cognates and false friends lead to different reaction times and accuracy scores than matched control words, then this means that the bilingual status of these words influences language processing; this can only be explained by cross-language activation processes. If, however, cognates and false friends do not lead to different reaction times and/or accuracy scores than matched control words, then this means that the bilingual status of these words does not influence language processing. Then there is no evidence of cross-language activation.

The results on studies with cognates and false friends lead to compelling evidence for cross-language activation processes. Studies with false friends have shown that their processing is typically slower when compared to matched one-language control words (Dijkstra, 2005). The degree of interference depends on the relative frequency of usage of the false friends in the associated languages. The strongest interference effects have been observed for low-frequency L2 words with high-frequency L1 false friends (e.g., Dijkstra et al., 1998; Smits et al., 2006). Studies with cognates have reported that these are usually processed faster than comparable control words ('cognate facilitation') and that the degree of facilitation depends on the cognate's degree of cross-linguistic similarity. This holds for words in isolation (Dijkstra et al., 2010) and in sentences (Van Assche et al., 2011). As with false friends, how frequently a cognate is used in a language also plays a role (Peeters et al., 2013).

There is growing evidence for cross-language activation across a range of different language structures and units. Cross-language interactions have not only been found in experiments at the word level, but also at the sentence level. Interpretation at the sentence level has for example been studied in interpretation tasks where bilinguals and second language learners are confronted with a simple sentence (e.g., 'the woman touched the girl' or 'the man touched the stone'), of which they must indicate which argument is performing the action. Many studies have found that second language learners are strongly influenced by L1 processing preferences, such as those due to word order, while performing the sentence interpretation task in their L2 (e.g., Gass, 1987; Harrington, 1987; Liu et al., 1992; McDonald, 1987).

However, when L2 learners become more proficient in their L2, they gradually develop more target-language-like sentence interpretation strategies. Cross-language interactions at the sentence level are not completely absent, however, in more proficient L2 learners. Even learners with a high proficiency in both languages show mutual influence between L1 and the L2 (e.g., Liu et al., 1992).

It is also important to take cross-language syntactic priming into account. Cross-language interactions at the sentence level have been shown to result from priming (e.g., Bernolet et al., 2012; Loebell & Bock, 2003; Hartsuiker et al., 2004; Kootstra & Doedens, 2016; Weber & Indefrey, 2009; see Kootstra & Muysken, 2017, for a special issue). With cross-language syntactic priming we mean the process in which a bilingual's syntactic processing of an utterance is influenced by the syntactic structure of a recently processed utterance *in a different language*. You hear a pattern in one language (for instance, a passive structure), and then you are inclined to use that same pattern not just in the same language (monolingual priming), but also in the other language. For example, Hartsuiker et al. (2004) asked pairs of Spanish-English bilingual subjects to describe cards depicting

transitive events (X HITS Y) to each other in a dialogue game. One of the participants was a confederate (a research assistant hired by the experimenter), and s/he produced utterances according to a prescribed format. The 'real' participants tended to use similar sentence constructions (X HITS Y, Y IS HIT BY X) as the confederate, even though the confederate used Spanish and the real participants used English. An example of a priming study that focused specifically on HL speakers is described by Fernández, de Souza, and Carando (2017). They used a priming task to study the occurrence of contact-induced linguistic innovations in the Spanish of Spanish-English bilinguals from Córdoba, Argentina (non-HL speakers, non-contact setting) and New York (HL speakers, contact setting). They found that the New York bilinguals produced more innovations than the Córdoba bilinguals. Another point to make is that cross-linguistic structural priming is also observed in children. Vasilyeva et al., (2010), for example, found that simultaneous Spanish-English bilingual children (which can of course also be regarded as HL speakers; cf., Kupisch & Rothman, 2016) produced more passive sentences in English after the experimenter had just produced a passive in Spanish (see Hsin, Legendre, & Omake, 2013 for similar results). Altogether, these studies show that cross-linguistic structural priming is a ubiquitous aspect of bilingual language use. Indeed, multiple studies have now shown various forms of cross-linguistic priming across different languages, from L1 to L2 and from L2 to L1, with different syntactic structures and constituents, and using different tasks (see e.g., Hartsuiker & Pickering, 2008; van Gompel & Arai, 2017, for reviews).

We can conclude that there is compelling psycholinguistic evidence for systematic and frequent cross-language interactions in bilinguals, including HL speakers. This evidence can easily be related to other evidence of cross-language interactions in bilinguals and HL speakers, such as codeswitching and cross-linguistic transfer (see other chapters in this book).

### 9.3.2 Processing differences

Another line of evidence in the psycholinguistics of bilingualism derives from processing differences between monolinguals and bilinguals in language production and comprehension. The basic finding from this research is that language use in bilinguals is typically slower and more effortful than it is for monolinguals (see also the findings in Chapter 6 on speech rate in HL speakers). Most of this evidence is based on experimental studies at the word level; however, in many studies the target is the majority language rather than the total knowledge over both languages. For example, based on standardized vocabulary tests, many studies have found that bilingual children often have a smaller vocabulary size for each separate language than monolingual children (e.g., Mahon & Crutchley, 2006; see

Bialystok, 2009, for a review). This can have consequences for their language processing skills. For example, in a series of studies, de Zeeuw and colleagues tested speed and accuracy of Dutch word reading by Turkish-Dutch bilinguals in Dutch primary schools (de Zeeuw et al., 2012, 2013, 2015). They did this by means of the lexical decision task, in which Dutch monolingual children and Turkish-Dutch bilingual children were presented with a list of Dutch words and pseudo-words that were presented one by one on a computer screen. For each word the participants had to indicate as fast and accurately as possible whether they thought the word was an existing Dutch word or not. Speed and accuracy of task performance are recorded and analyzed as measures of word knowledge and lexical retrieval, which are essential predictors of one of the most important academic skills, namely reading (see also Perfetti & Hart, 2002). De Zeeuw and colleagues found that, in general, the Turkish-Dutch bilingual children were both slower and less accurate than their monolingual peers on this task. This is consistent with other research on the same population of bilinguals, which found that bilingual children in primary school typically lag in reading skills and vocabulary knowledge (e.g., Cremer & Schoonen, 2013; Durgunoglu & Verhoeven, 1998; Verhallen & Schoonen, 1998; Verhoeven, 2000; Vermeer, 2001). Although this would need to be confirmed in meta-analysis of data from studies across very diverse bilingual populations, research from other bilingual populations, using other measures, appears to confirm this pattern of results. Hoff et al. (2012), for example, studied lexical and grammatical development in young bilingual and monolingual children, and found that the monolingual children had higher vocabulary and grammar scores than the bilingual children. Note, however, that this was only the case in single-language comparisons; the monolingual and bilingual children had similar vocabulary scores when bilinguals' total vocabulary (i.e., for both languages) was considered.

Not only do bilinguals often lag in receptive skills like word reading, but also in productive skills like word naming. This has most notably been found in picture naming studies. Picture naming is one of the ways used most to study language production (see e.g., Bock, 1995, for more information on methods and methodologies in language production research). It is the 'production alternative' to lexical decision: a participant is presented with a list of pictures that appear one-by-one on a computer screen. The participant must name the picture as accurately and quickly as possible, into a microphone. Naming speed and accuracy is recorded and analyzed as a measure of lexical retrieval in language production. It has been found in many picture naming studies that bilinguals and L2 learners are generally less quick and less accurate than monolinguals in picture naming (see Hanulová, Davidson, & Indefrey, 2011, for a review; see also Bialystok, 2009). Since picture naming is a miniature version of real language production (i.e., the picture represents a concept that has to be put into words and then articulated), these picture

naming studies provide important experimental support of phenomena that also take place in real-life language production in bilinguals and HL speakers, such as reduced speech rate (Polinsky, 2008b).

An important question to answer is now: why would the language production and comprehension of bilinguals be slower and more error-prone than in monolinguals, particularly in their weaker language? This question may have several answers. Based on the findings presented above, a likely answer would be that bilinguals simply have had less exposure and experience to each of their languages than monolinguals, leading to the situation that the frequency with which target-language words and linguistic structures have been used and encountered is generally lower in each language of the bilinguals than for monolinguals. This is indeed what Hoff et al. (2012) found in their research: all measures of vocabulary and grammar development they analyzed in their study with Spanish-English bilingual children could be traced back to the relative amount of input in the language. Importantly, however, it could also be the case that cross-language activation and the need to control for it slows down processing as well (see also Bobb & Wodniecka, 2013).

There are thus two lines of evidence for different processing rates in bilinguals. How to account for them? Frequency of input and language use but also competition and cross-language activation and controlling for this competition are involved (see also Putnam & Sánchez, 2013, for more thoughts on this). More research is needed to find out more about how these processes work together in shaping bilingual language production, comprehension, and acquisition.

### 9.3.3 Language switching and inhibition

Other evidence for the way bilingual processing works comes from language switching studies (e.g., Costa & Santesteban, 2004; Meuter & Allport, 1999; Verhoeve, Roelofs, & Chwilla, 2009; see Meuter, 2005, 2009, for reviews). These studies all use variations of the same experimental task, in which bilingual speakers are asked to name pictured objects or numbers one by one as they are presented on a computer screen. External cues (e.g., the background color of the presented item) induce participants to use either one or the other language in naming the item. The stimulus lists are created such that participants must switch languages or not between the naming of one item and the next one. The cost of switching is measured by subtracting the naming latencies of switch trials from non-switch trials. Generally, the magnitude of the switch cost has been found to depend on a combination of the direction of switching and the participants' relative language proficiency: In non-balanced bilinguals, the switch cost is higher when participants must switch from their non-dominant language to their dominant

language than vice versa, whereas in balanced bilinguals, switch costs are more symmetrical. This result is regarded by many researchers as evidence of inhibition processes in language production: To be able to speak one language, it is necessary to inhibit activation of elements from the non-target language. When the non-target language is the dominant language, inhibition takes a relatively high effort. Recovering from this inhibition of the dominant language will then also take a relatively high effort, which explains the high switch cost when switching back to the dominant language.

Importantly, it is tempting to directly relate these laboratory findings of language switching to the natural phenomenon. Based on the findings above, one would then assume that codeswitching is costly behavior. This direct link between language switching and codeswitching is too simple, however. One could even argue that language switching studies are more informative about how to keep languages apart than about how to go back and forth between languages (i.e., codeswitching). Indeed, research on the cognitive mechanisms of codeswitching has indicated that codeswitching involves more than executive control or opportunistic language production planning; rather, it involves knowledge of both grammars, socio-pragmatic mechanisms, adaptive control, and other mechanisms of language use that are also involved in monolingual language use (see e.g., Fricke & Kootstra, 2016; Gollan & Goldrick, 2016, 2018; Green & Abutalebi, 2013; Li Wei & Green, 2014, Kootstra, van Hell, & Dijkstra, 2010; for more discussion). Thus, whereas it is evident that inhibition of the non-target language plays a crucial role in language processing in multilingual speakers, it is crucial to also take other factors into account.

### 9.3.4 Summary of preceding discussion

Let us return to models based on the activation metaphor: elegant and close to neural processes in the brain. They form a nice point of departure for research. The fact that bilinguals can keep their languages apart suggests that they can selectively activate or inhibit items from different languages. On the other hand, codeswitching and other types of cross-linguistic interaction, such as language transfer, imply that they also co-activate items from different languages at the same time. Cognitive models of bilingual processing must assume that codeswitching and other forms of cross-linguistic interaction, as well as the ability to keep languages apart, are based on the same underlying cognitive architecture and involve the same mechanisms.

The same models have been suggested to account for cross-language activation phenomena in real life, such as codeswitching, hesitant speech and impacted speech rate, cross-linguistic influence in second language learners, foreign accent,



contact-induced language change and convergence, phenomena that play an important role in HL speakers.

## 9.4 Factors influencing language processing in bilinguals

The findings described in this chapter indicate that the language processing system in bilinguals is shared for both languages and open to cross-language interactions (i.e., cross-language activation). We now describe several factors at different levels of language processing that have been shown to influence the likelihood of cross-language interactions in bilingual language use (see also Kootstra, van Hell, & Dijkstra, 2009; Kroll, Bobb, & Wodniecka, 2006; Bultena, 2013, for more information). Importantly, these factors also include variables that are crucial in differentiating different types of HL speakers, such as language dominance and age of acquisition (see also Section 9.5.)

First, socio-contextual factors play an important role. Their impact has most specifically been addressed in the so-called language mode hypothesis by Grosjean (e.g., Grosjean, 2001, 2008). Grosjean has argued that, amongst other things, interlocutors, the physical location, and the functionality of the discourse have considerable impact on the overall state of activation of the bilingual's languages. This will affect the way bilinguals process their different languages and it will differ across conversational settings. Evidence of the language mode hypothesis has been found in both naturalistic and experimental data. Naturalistic evidence may for example be seen in the seminal study by Blom and Gumperz (1972). Based on qualitative analyses of conversations between residents of a Norwegian town, Blom and Gumperz pointed out that the choice between the standard and a dialect variety of Norwegian, including codeswitching, was influenced by the topic of the conversation and by the socio-cultural identity of the speakers. They thus emphasized that language choice and codeswitching carry social meaning and are determined by the social setting (as has subsequently been shown many times; see also Chapters 3 and 4). This evidence could clearly be interpreted in terms of the language mode hypothesis, although probably these researchers assume speaker agency to play an important role.

Experimental evidence supporting the language mode hypothesis can for example be found in comparisons between task settings in experiments. That is, even though it is only vaguely like a natural discourse setting, the setting in which an experiment takes place can still be seen as an instantiation of a real-life discourse situation (Green, 2011). Discourse settings in real life can be more or less multilingual. Similarly, experiments can have a single-language context or a dual-language context. Single-language contexts refer to experiments in which only one language

needs to be used to perform the task, like a task in which participants name objects in one language, and all instructions are also in that language. In a dual-language context the experiment requires participants to use two languages to perform the task, such as a task in which participants name objects in multiple languages. Also, more than one language may be used in the instructions. Cross-language effects occur in both single and dual-language contexts, but they have been found to be more prevalent and/or stronger in dual-language contexts than in single-language contexts (see e.g., Christoffels et al., 2007; Dijkstra & van Heuven, 2002; Hatzidaki et al., 2011). Perhaps dual-language contexts can be interpreted as the experimental parallel to a bilingual language mode setting from the perspective of activation.

In addition to these comparisons between single-language and dual-language contexts in experiments, there is also more direct experimental evidence for the influence of socio-contextual factors on cross-language interactions, namely in experimental studies on codeswitching. Kootstra, van Hell, and Dijkstra (2012) used *confederate-scripting techniques* to study how dialogue partners adapt their language to each other. In confederate-scripting, two participants (one of whom is a confederate, as already described, an actor instructed beforehand by the experimenter and whose linguistic behavior has been scripted, unbeknownst to the real participant) sit facing each other. They each have a laptop in front of them and perform a dialogue game in which they describe to each other pictures they are presented with on their laptop. In the Kootstra et al. study, the confederate code-switched in a picture description in 50% of the cases. The real participant was completely free to use whatever language or combination of languages (s)he wanted to use in the picture descriptions. Now, did the confederate's codeswitching influence the tendency of the real participant to code-switch? Indeed, the real participants' picture descriptions were significantly more often code-switched when the confederate had just code-switched in the previous turn, compared to when the confederate had not code-switched in the previous turn. Thus, the confederate's switching influenced the participants' tendency to code-switch, which illustrates the impact of socio-interactive processes on codeswitching behavior.

In addition to discourse-situational and socio-interactive effects, cross-language activation processes are also influenced by the degree of lexical and syntactic overlap across languages (see e.g., de Bot, Boersma, & Isurin, 2009; Kootstra et al., 2009, for reviews). Indeed, the influence of the overlap across languages is one of the central notions in research on cross-language interaction. Overlapping items or structures are associated with more than one language. The occurrence of cross-linguistic interaction is more likely with such items than with language-specific structures.

The distinction between overlapping and language-specific structures is reflected in various approaches to cross-linguistic influence. In experimental

psychology, for example, selective versus non-selective lexical access can be studied by comparing the processing of overlapping lexical items to the processing of language-specific items (cf., e.g., Dijkstra, 2005), as we saw in the discussion above on false friends. In the same vein, second language acquisition research has assumed that cross-linguistic influence (or: transfer) results from “the *similarities* and *differences* between the target language and any other language that has been previously [...] acquired” (Odlin, 1989, p. 27, our italics), and syntactic overlap is also crucial in theories on cross-linguistic influence in bilingual first language acquisition (e.g., Hulk & Müller, 2000). In addition, findings of cross-linguistic structural priming have been found to be the strongest and most consistent in cases where the word order of both languages involved is the same (see Van Gompel & Arai, 2017, for a review). Finally, numerous accounts of codeswitching are based on similarities and differences in the ways in which different languages map meaning onto form (e.g., Deuchar, 2005; Muysken, 2000; Poplack, 1980). A typical example is Poplack’s equivalence constraint, which has been experimentally confirmed in Kootstra et al. (2010).

Cross-language interactions as described above also turn out to be influenced by *relative language proficiency*. Van Hell and Dijkstra (2002) have linked the size of the ‘cognate facilitation’ effect in trilinguals to their proficiency in their third language. When Dutch (L1) university students with English as a second language (L2) and French as a third language (L3) performed a Dutch lexical decision task, a cognate facilitation effect was found for Dutch-English cognates, but not for Dutch-French cognates. Since the trilingual participants were all native speakers of Dutch (L1) with a better knowledge of English than of French, this was to be expected. In a second experiment, trilinguals with much better knowledge of French were tested (i.e., Dutch university students of French). For this group, facilitation effects in Dutch were obtained for both Dutch-English (L1–L2) and Dutch-French (L1–L3) cognates. This study shows that relative proficiency in another language (L2 or L3) affects the size of the cognate facilitation effect in their L1, at least for this type of population.

As in bilingual word processing studies, cross-language interaction in sentence production is influenced by relative language proficiency as well. Schoonbaert et al. (2007) tested Flemish learners of English and found that cross-language syntactic priming occurred both from L1 to L2 and from L2 to L1. However, enhanced priming, due to lexical repetition between prime and target (the ‘lexical boost’ effect) was only found from L1 to L2; not from L2 to L1. Likewise, Kootstra and Doedens (2016) tested Dutch learners of English and found that long-lasting effects of cross-language syntactic priming were only found from L1 to L2 and not from L2 to L1. L1 structural preferences associated with specific verbs (the verb alternation bias, for instance GAVE MARY A BOOK versus GAVE A BOOK TO

MARY) influenced L2 syntactic choices in picture descriptions, while L2 structural preferences influenced L1 syntactic choices to a lesser extent. In a related way, Kootstra, van Hell, and Dijkstra (2012) also found that lexical priming in the production of bilingual sentences was stronger in high-proficient bilinguals than in low-proficient bilinguals. Finally, it should be noted that findings of relative language proficiency have also been found in the child bilingualism literature: both cross-linguistic influences and general measures of vocabulary and grammar development in these children are related to measures such as language dominance and quantity and quality of input (cf., e.g., Unsworth, 2013, for an overview).

## 9.5 The issue of age of acquisition

Psycholinguists have also addressed the influence of the age of acquisition, which of course plays an important role when it comes to the notion of HLs. The consensus view is that the sensitivity to learn a second language decreases as you get older. However, it appears that the earlier assumption in terms of a strict ‘critical period’ for L2-learning is too strong. Sensitivity does not decline catastrophically after puberty, but gradually, and learning another language at a later age is still possible.

When babies begin learning a language they are sensitive to the *phoneme distinctions* that exist in the different languages. Already after half a year their perception of non-native phonemic contrasts declines while that of specific native contrasts (for instance, Voice Onset Time) increases (e.g., Werker et al., 1981; Kuhl, 2004). The older they get, the more difficult they find it to learn the sounds of other languages. However, this development is not unitary but depends on the specific acoustic and articulatory sound features involved (Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005). Nevertheless, when learners are motivated and try hard, a very high level of pronunciation accuracy can still be reached also in later stages of life (Bongaerts et al., 1997).

The age effect is stronger for the acquisition of the *grammar* of a foreign language than for vocabulary is less sensitive. Ullman (2001) explains this in terms of two types of memory. The first type of memory stores facts and events (declarative memory); this is important for our ‘dictionary’, i.e. words. The second type of memory is involved in learning and actions (procedural memory) and is important for the acquisition of sequences. Grammar involves word order and sound sequences, in fact sequences more generally, and learning grammar therefore involves the second type of memory. This second type of memory would be sensitive to hormonal changes, and it would operate a bit better before puberty. In this view one learns a language somewhat differently at a later age relative to an earlier age: somewhat less ‘intuitively’ and with a more conscious desire to derive

rules. Seen from this developmental perspective, it can be understood why (late) L2 learners reveal themselves as non-native especially in terms of their pronunciation and grammar (see e.g., Diaz et al., 2008).

In a classic study, Johnson and Newport (1989) tested the English proficiency with a grammaticality judgment task of 46 native Korean and Chinese speakers who had arrived in the United States between age 3 and 39. A variety of English syntactic constructions was involved, including past tense, plural, third person singular, present progressive, and determiners. Those who arrived earlier clearly performed better relative to later arrivals. The advantage was linearly related to age of arrival up to puberty (set at age 16), but variable and unrelated to age of arrival after puberty. This effect held for all grammatical structures tested, although there were considerable performance differences between constructions and individuals as well (for follow-up studies see Birdsong & Molis, 2001; Liu, Bates, & Li, 1992).

Schmid (e.g. 2011, 2012) has studied a ‘real life’ event that underlines this point. At the brink of the Second World War, in the years 1938–1939, a very large number of Jewish children were transferred from Germany to England by charity organizations to save them from persecution, the so-called ‘*Kindertransport*’. The age of the children varied from very young (2 y) to post-puberty (17 y). Since the children ended up in foster parent families, they often lost all opportunities to speak German, and in many cases, they were reluctant to do so anyway. However, sometimes the children escaped with other family members, and as a consequence some German communication took place after migration. Starting in the 1990’s, testimonies of these people have been recorded in extensive archives, and this makes it possible to discover the precise role played by the age at which the children acquired the foreign language and the moment at which they started to forget their mother tongue. Schmid showed that the age at which the children left Germany was more important for the characteristics of their heritage German than their subsequent contacts in England with other family members. Schmid found that their language problems were especially concerned with the activation and access to German words, rather than with the loss of general linguistic knowledge or the replacement of their first language by a second one.

## 9.6 Concluding remarks and perspectives for codeswitching research

From a psycholinguistic perspective, language use (i.e., comprehension and production) involves that the activation and selection of linguistic representations from an interconnected network of representations that is shared for both languages. Language processing in such a mental network is inherently interactive and dynamic: cross-language connections and parallel activation of languages

explain why and how cross-linguistic interactions such as codeswitching and language transfer can occur, and there are multiple speaker-specific, discourse-situational, and linguistic variables that influence the ease of language processing as well as the likelihood of cross-linguistic influence. With respect to HL speakers, speaker-specific variables such as age of acquisition and language dominance can be of specific interest and are crucial for theory formation. However, these speaker-specific variables should ideally be studied in combination with socio-contextual and linguistic variables to facilitate a more complete, processing-based account of the cognitive mechanisms that govern language use in HL speakers. It is a challenge for psycholinguistic research to embed its experimental designs in real-life settings, accounting for other intervening variables, and thus make such research even more relevant for the understanding of HL speakers.

The future in this area of research lies in further modeling the dynamics of bilingual language processing. What are the relations with language change, cognitive skills, developments in the individual across the lifespan, and educational programs? What is clear in any case is that the bilingual speaker, and a fortiori the heritage speaker, is not the simple result of putting two languages into one person. At the very least, theories will need to account for the continuous interaction that is often observed. This becomes clearest from codeswitching studies with HL speakers e.g. Clyne (2003), who has documented cases like:

*En we reckoned Holland was too smal voor ons.*  
 ‘And we reckoned Holland was too small for us.’

Note that in the example above the etymologically related words ‘small’ in English and *small* ‘narrow’ in Dutch are conflated. In data from Papiamentu-Dutch HL speakers (Muysken, Kook, & Vedder, 1996) we have cases such as:

*Bo ta sinta na bo gemak?*  
 2SG PRES feel LOC 2SG ease  
 ‘Do you feel at ease?’

This last case involves just one Dutch word, but reflects a Papiamentu calque of the Dutch expression *voel je je op je gemak?* which contains that same word. We will return to this type of interaction in the next chapter.



# Heritage languages in a post-colonial setting

## Focus on Papiamentu

### 10.1 Introduction

Many HLs spoken in immigrant communities do not have a history of contact with the dominant language. There was, for example, very little contact between Danish and Turkish before people of Turkish descent started to settle in Denmark (Jørgensen, 2003) as part of the labor migration movement of the 1960's. However, in several cases there already was a longer contact history: the case of the post-colonial HLs. Hindi has been in contact with English since the British invasion of India and the subsequent founding of Imperial India as a British colony. It now interacts with English in migration settings both in the UK and in the United States. Malay has been in contact with Dutch through much of the colonial history of Indonesia, and then after the migration of the speakers of Ambon Malay to the Netherlands contact between the two languages simply intensified, as the community partially shifted to Dutch.

Similarly, the complex relation between British English, West-Indian English and Caribbean Creole in the UK is linked to their colonial and post-colonial relationships in the Caribbean (Sebba, 1993). If the dominant language is also the original lexifier language of the creole, however, as in the case of English and London Jamaican, it is sometimes hard to distinguish between loans from the dominant language into the HL and the original lexicon of the HL.

The effect of long-time contacts on an HL will be illustrated in this chapter with the case of Papiamentu, the creole language of the Caribbean islands Aruba, Bonaire and Curaçao. Papiamentu and Dutch have been in a complex colonial and post-colonial relationship ever since the mid-17th century. In this chapter we describe the nature as well as the linguistic consequences of this relationship. The focus in this chapter is not on the status of Papiamentu as a creole, something not relevant to the discussion here, but on the fact that it had a long contact history with Dutch already in the colonial setting. This chapter is partly based on Jacobs (2014), and draws on Vedder, Muysken, & Kook (1996) and Muysken, Kook, & Vedder (1996).



A typical quote showing some of the contact phenomena involved as well as the speakers' attitudes to the changes would be the following statement by someone from Curaçao:

*E hebruik di hollandse woorden ta duw nos Papiamentu achteruit*  
 DET use of Dutch words PRES push 1PL Papiamentu backwards  
 'The use of Dutch words pushes our Papiamentu backwards.'

(Brenneker, 1961, p. 148, 149, cited in Wood, 1970, p. 88)

The Dutch words in this Papiamentu utterance are in bold, and ironically the very utterance illustrates what the speaker is criticizing.

The structure of this chapter is as follows: after providing some general background information on the origins of Papiamentu (Section 10.2) and on the language itself (10.3), we give a brief sketch of the history of Papiamentu-Dutch contacts (10.4), focusing mainly on Curaçao, the island where the Dutch colonial activity was centered and where Papiamentu first emerged. Subsequently, we describe the linguistic contribution of Dutch to Papiamentu grammar (10.5) and then address the morphosyntactic integration of Dutch loanwords in Papiamentu (10.6). In the final Section (10.7), we shed light on (the nature of) the use of Papiamentu in the Netherlands.

#### **Aims of this chapter**

To present the main features of HLs which have a post-colonial origin and thus a long history of contact with a dominant language.

To show that lexical influence is often very substantial: some loanwords are quite old and, in the case of creole languages, may date from the period of genesis of the HL.

To illustrate that there may be several varieties of the HL dating from the very beginning of migration, which differ in the degree of influence from the dominant language.

To show that the influence of the dominant language may include structural influence.

## **10.2 Early history of Papiamentu**

The Dutch West India Company (WIC) took Curaçao in 1634 forcing the few Spaniards and Amerindians present at the time to leave for the mainland. Curaçao's linguistic history thus started from scratch in that year. The importation of high numbers of slaves from West Africa commenced in the 1650s and came to a halt in the (early) 18th century (see Postma, 1974 for details). There is broad agreement among scholars that Papiamentu emerged on Curaçao in that period, more precisely in the second half of the 17th century. In the 18th century, Papiamentu spread to Aruba and Bonaire through migration.

The first reference to Papiamentu is from 1704, when a German priest travelling to Curaçao mentions that the slaves there speak ‘broken Spanish’. Another priest reports in 1732 that, in addition to Spanish, Portuguese, and Dutch, the people of Curaçao speak ‘the language of the country’ and a legal deposition of 1737 mentions the use of a ‘creolse taal’ [‘creole language’] (Rupert, 2012, p. 214). The first attestation of the glossonym Papiamentu (<Poppementu>) is in a 1747 legal document (Hartog, 1968; Fouse, 2002; Rupert, 2012).

The basic vocabulary of Papiamentu is mixed Spanish-Portuguese (e.g. Papiamentu *blanku* ‘white’ < Spanish *blanco* vs. Pap. *pretu* ‘black’ < Portuguese *preto*). Hence, the classification of Papiamentu as an either Spanish- or Portuguese-lexifier creole is controversial. To avoid this controversy, scholars often label Papiamentu an Iberian-based (or Afro-Iberian) creole. Martinus (1996), Quint (2000) and Jacobs (2009 and elsewhere), however, have revealed linguistic ties with the Portuguese-based creoles of Upper Guinea, claiming this variety was transferred to Curaçao and subsequently relexified towards Spanish. The issue of the origins of Papiamentu will not play any role of importance in this chapter, though.

### 10.3 Background on Papiamentu and its status nowadays

Papiamentu is one of the official languages (next to Dutch) of Aruba, Bonaire, and Curaçao (the Leeward Islands of the former Netherlands Antilles), or ABC Islands. The language has an estimated number of 270.000 speakers, of whom 120.000 live on Curaçao, ca. 60.000 on Aruba and 10.000 on Bonaire. Approximately 100.000 people from the ABC islands live in the Netherlands, most of them Papiamentu-speaking (cf. Kook & Narain, 1993, p. 69). Papiamentu received official status (alongside Dutch and English) on Aruba in, 2003, with Bonaire and Curaçao following in, 2007.

Corresponding to the three islands we can distinguish three main dialects, which are mutually perfectly intelligible but do display some subtle differences in the lexicon, the phonology (e.g. unstressed word-final [u] in Curaçao and Bonaire typically corresponds to an [o] on Aruba: /papiamentu/ vs. /papiamento/), and a few aspects of the morphosyntax. Most existing grammars, grammatical sketches and dictionaries of Papiamentu are based on the variety of Willemstad, the capital city of Curaçao. The more recent dictionaries (Joubert, 1999; Ratzlaff, 2008; van Putte & van Putte-de Windt, 2005) do contain dialectal information for several lexical entries, but a comprehensive study of the linguistic variation between the islands is lacking to date.

Aruban Papiamentu is traditionally thought to be more influenced by Spanish due to strong contacts with the mainland, whilst Curaçaoan Papiamentu is usually

claimed to have been more influenced by Dutch. This is sometimes readily visible in the lexicon (e.g. Aruban Pap. *novia* ‘fiancée’ < Spanish *novia* vs. Curaçaoan Pap. *bruit* < Dutch *bruid*). However, the claim has never been properly investigated and is not always confirmed by the data: for instance, the Dutch-derived passive auxiliary *wòrdu* (and variants) is surprisingly more frequent on Aruba than on Curaçao (Eckkrammer, 2004). Perhaps, the Papiamentu of Curaçao has undergone less influence from other languages altogether in its post-formative stage.

On each island, and probably most clearly on Curaçao, there is quite considerable sociolectal variation. Some varieties are less influenced by language contact – in creole studies known as ‘basilectal’ varieties, typically corresponding to rural areas – whilst in the urbanized areas more ‘acrolectal’ varieties are spoken, which are influenced more heavily by Dutch, Spanish and English (cf. Wood, 1970, p. 87; Wood, 1971; Kramer, 2004, p. 154; Sanchez, 2005, p. 110 footnote).

The Antillean community in the Netherlands has a very diverse sociological profile, including well-established long-term residents of Antillean origin, students, and young people with little chance of employment and living in poor conditions. Kester and Hortencia (2010) and Kester and Fun (2012) present the results of sociolinguistic surveys on the relative importance of Papiamentu vis-à-vis Dutch among Curaçaoan and Aruban students in the Netherlands. Altogether, there is a strong appreciation for Papiamentu among Antilleans, even though different groups of Antilleans differ widely in the degree to which they use Papiamentu in their daily lives. There are also numerous cultural associations, such as Antillean student societies.

#### 10.4 A brief history of Papiamentu-Dutch contact

Within the Caribbean, the ABC Islands occupy a somewhat unique position: unlike elsewhere in the Caribbean, the inhabitants did not (or hardly) shift to the colonial language, Dutch. By contrast, for instance, Spanish became the native language in all Spanish Caribbean colonies. Furthermore, Dutch did not become the lexifier language for a creole on the ABC Islands, unlike e.g. English on Jamaica or French on Haiti. Dutch did provide the lexicon for creoles in other places in the Caribbean (Negerhollands on the Virgin Islands, Berbice Dutch Creole and Skepi Dutch in Guyana).

However, the social and economic ties with the Netherlands have always remained essential to the ABC Islands at large, enabling Dutch to retain its historical role as the language of administration, politics, and education and to exert influence on Papiamentu over a period of three and a half centuries and counting.

### 10.4.1 The 18th century

Papiamentu emerged among the slaves in the second half of the 17th century. During the 18th century it spread to the white European upper class consisting mainly of Dutch Protestants and a substantial Sephardic Jewish settler population. In the 18th century, Dutch Protestants constituted roughly 13% of the Curaçaoan population while Sephardic Jews made up 8%. The remainder consisted of slaves and freed slaves (Maurer, 1988, p. 192). The diffusion of Papiamentu among the white upper class is evidenced by two of the earliest written attestations of Papiamentu, dated 1775 and 1783. It concerns private letters by a Sephardic Jew and a Dutch female citizen of Willemstad (Jacobs & van der Wal, 2015).

Dutch-Papiamentu bilingualism in this early period involved the contact between Afro-Curaçaoan *yayas* (nannies) and children of Dutch slave owners. Dutch furthermore was the official language of government and trade; merchants of whatever background needed to have at least a basic command of Dutch (Rupert, 2012, p. 231). The role of Dutch in missionary activities in this period was minimal: the Christianization of the Afro-Curaçaoan population was largely (and often deliberately) left in the hands of Spanish/Venezuelan missionaries.

### 10.4.2 Increase of Dutch influence on Curaçao in the 19th and 20th centuries

However, in the early 19th century, following the slave revolts from 1795–1800 (Klooster & Oostindie, 2011), an increasing number of Dutch clergymen was sent to Curaçao to gain more influence in the Afro-Curaçaoan part of the island society. The first Dutch priest, Johannes Niewindt, arrived in 1824 (Fouse, 2002, pp. 127–131). Significantly, he chose to preach in Papiamentu, setting the example for his Dutch successors on the island. As printing press was introduced in the 1820s (Rupert, 2012, p. 246), the following decades saw Dutch clergymen (Catholic as well as Protestant) produce a significant corpus of evangelical texts in Papiamentu. In 1916, Eybers (1916) edited the first translation of the Bible into Papiamentu.

Parallel to the increased evangelical efforts, attempts were made to implement a sound educational system. Dutch was declared the official language of school instruction in 1819, but it would still take some time for a full-fledged educational system to develop (Fouse, 2002, p. 137). In 1849, Father Jacobus Putman (Putman, 1849) wrote a didactic grammar of Dutch in Papiamentu.

The contact between Dutch and Papiamentu in the religious and educational domains was an important incentive for new Dutch words and features to be borrowed into Papiamentu. For instance, the passive auxiliary Pap. *wòrdu* (< Dutch *worden*) can be seen to emerge in evangelical texts from after 1850 (cf. Sanchez,

2005; Jacobs, 2011). Pap. *wòrdu* is nowadays solidly integrated in the Papiamentu grammar (see Eckkrammer, 2004).

Meanwhile, Papiamentu continued to spread as the language of the wider community. Cohen Henriquez (1934) affirms that during the 19th century Papiamentu had already become widespread in Willemstad's two European residential neighborhoods Punda (Sephardic Jewish) and Otrobanda (predominantly Dutch Protestant). For 19th-century Otrobanda, he describes a distinct Dutch-lexified dialect of Papiamentu, whilst Punda was home to a more Iberian-lexified variety which resembles modern-day Willemstad Papiamentu more closely. This is illustrated here with two alternative versions of the same sentence, in which the Dutch loans in the upper version are underlined. The selection of the Pap. preposition *riba* 'on' (< Iberian *arriba*), whose use in more traditional Papiamentu is strictly locative) is a calque on the Dutch preposition *op* 'on' in the collocation [*verliefd op*] 'in love on' → 'in love with'.

		<i>let op</i>			<i>merke</i>			<i>jonglui-nan</i>
<i>si</i>	<i>bo</i>		<i>bon, lo</i>	<i>bo</i>	<i>ku</i>	<i>e</i>	<i>dos</i>	
		<i>buta atenshon</i>			<i>ripara</i>			<i>joven-nan</i>
if	you	pay attention	Well	FUT	you	notice	that	DET two young-PL
		<i>verlief</i>	<i>Riba</i>					
<i>ei</i>	<i>ta</i>			<i>otro</i>				
		<i>namorá</i>	<i>Di</i>					
there	be	in love	With	other				

'If you observe carefully, you will notice that those two young people are in love with each other'

Variation in 19th century European Papiamentu in Willemstad (Otrobanda top, Punda bottom) (Cohen Henriquez, 1934, p. 32, cited in Wood, 1970, p. 83)

A new phase of Dutch influence on Curaçao came with the establishment of an oil refinery of the Royal Dutch Shell on Curaçao in 1915, triggering waves of immigration from the Netherlands and a remodeling of the Curaçaoan school system with Dutch as the sole language of instruction. The Dutch linguistic educational policy was one of complete submersion; speaking Papiamentu was prohibited at the pain of physical punishment. Against this background, anti-Dutch sentiments and a desire for autonomy grew steadily during the 20th century, culminating in the famous 1969 labour rebellion (Fouse, 2002, pp. 147–150).

#### 10.4.3 The current situation

At present, Dutch still predominates on the ABC Islands in education and in formal domains such as politics, law and administration and it is still often preferred

over Papiamentu by local novelists (Dijkhoff, Kouwenberg, & Fat, 2006, p. 2107). A comparison of censuses from the 1980s and 2000s (Table 10.1) confirms Kook & Narain's (1993, p. 72) observation that "the proportion of people in Curaçao speaking only Papiamentu is gradually decreasing, while the proportion of people speaking only Dutch or both Papiamentu and Dutch are gradually increasing".

**Table 10.1** Census of 1981 and 2000/01 (adapted from Maurer, 1988, p. 143 and Kester & Fun, 2012, p. 238)

	Aruba, 1981	Aruba, 2000	Curaçao, 1981	Curaçao, 2001
Papiamentu	80.1	69.4	86.9	80.3
Dutch	5.0	6.1	6.8	9.3
Other languages (mainly English and Spanish)	14.9	24.5	6.3	10.4

On the other hand, as Dijkhoff, Kouwenberg, & Fat (2006, p. 2107) note, the Dutch language is "barred from social manifestations that may be termed 'local' or 'national', and in the arts, the non-print media (...) its role is negligible." Indeed, most radio stations and local television channels transmit in Papiamentu and most newspapers appear in Papiamentu as well (Kook & Narain, 1993, p. 71).

Altogether, in the second half of the 20th century and the early 21st century, Papiamentu has grown in its role as a local identity marker (Oostindie, 2008); the Papiamentu literary output has increased, and citizens have begun to voice concerns over the dominant role of Dutch in the school system at the expense of Papiamentu. However, pleas for introducing Papiamentu at all levels of education have not been met with great enthusiasm in the Curaçaoan society at large, for the simple reason that "proficiency in Dutch is still seen as providing access to advanced study and better job opportunities" (Dijkhoff, Kouwenberg, & Fat, 2006, p. 2107).

The double status of Dutch on Curaçao (that is, its comparatively low socio-cultural prestige vs. its high economic importance) and the corresponding language-political and educational issues is what Oostindie (1995) refers to as 'Curaçao's dilemma': whilst Papiamentu is increasingly propagated as the symbol of the culture of the ABC islands, the corresponding desire to introduce it in the political and educational realm at the cost of Dutch is not (yet) a real option.

## 10.5 Dutch influence on Papiamentu

The Papiamentu-Dutch language contact situation described in the foregoing is comparable to many (post-)colonial settings in which "[a] socially subordinate language borrows from a socially dominant language, whereas the reverse is

much less frequent” (Muysken, 2008, p. 301). This situation is also referred to as ‘asymmetric borrowing’ (cf. Matras, 2009, p. 193). The result of this borrowing is an unmistakable Dutch imprint on Papiamentu. Some of the most frequent colloquial expressions contain words of clear Dutch etymology, such as Pap. *hopi bon!* ‘great, very well!’ where *hopi* is derived from the Dutch noun *hoop* ‘heap’ → ‘a lot, much’, *masha danki!* ‘many thanks!’ where *danki* is derived from Dutch *dank* ‘thank’. Other Dutch loans are less easily discernible but nonetheless highly frequent, such as the colloquial address form Pap. *sua* (as in *ki tin, sua?* ‘what’s up, brother?’), derived from Dutch *zwager* ‘brother in law’.

Below we review those Dutch contributions generally considered to be an integral part of standard, mainstream Papiamentu as spoken on the ABC Islands. The most important references on this topic include Fokker (1914), Lenz (1928), Wood (1970, 1972), Maduro (1953), Andersen (1974), Kowallik & Kramer (1994) and Kramer (2004, p. 139–155).

### 10.5.1 Quantity and quality of Dutch loans

Estimates of the quantity of Dutch loanwords in Papiamentu vocabulary hover around 30% (Lenz, 1928, p. 210; Maduro, 1953, p. 143). This 30% is not equally distributed over all word classes. Around 69% of Dutch loans in Papiamentu consists of nouns, 18% verbs, 8% adjectives, 3% adverbs, and 2% other items such as prepositions, discourse markers, and interjections. (Calculations based on Wood, 1970.) This distribution corresponds roughly with universal tendencies of contact-induced change, which predict that content words are more easily borrowed than function words and that, within content vocabulary, nouns will be more prone to borrowing than verbs, verbs more than adjectives, etc. (Matras, 2009, p. 157).

Along with Dutch lexical items, several Dutch phonemes have been borrowed into Papiamentu, such as the vowels /ø/, /y/ as well as the voiced fricatives /v, z/, while the phoneme /x/ probably results from converging Spanish and Dutch influence (Holm, 1988, p. 316).

In principle, the more basic the vocabulary, the fewer Dutch loanwords we find. For instance, while the *total* number of Dutch loans in Papiamentu fluctuates around 30%, on the Swadesh 100 list (Swadesh, 1971, p. 283) – still commonly used as a yardstick for identifying basic or core vocabulary – only 3% to 9% percent is of Dutch origin. The reason why this cannot be determined more accurately is that most Dutch loans on the Papiamentu Swadesh list have an Iberian-based synonym. For instance, ‘leaf’ is translated by either Pap. *blachi* (< Dutch *blaadje*) or Pap. *foyo/a* (< Portuguese *folha*). The Dutch items with no Iberian variant are *nèk* ‘neck’, *santu* ‘sand’, and *hel/gel* ‘yellow’, from Dutch *nek*, *zand* and *geel*. Hence, we took 3% as the minimum. Again, this situation is comparable to many other

cases in which the speech of a prestigious outsider influences the the language with less prestigious, as was shown for Dutch Turkish in Chapter 4.

Unsurprisingly, borrowing from Dutch is rife in semantic domains where Dutch has historically played an important role, such as domestic vocabulary, construction work, and education; and they are predictably scarce in some other domains, such as religion, which is historically related to Spanish Catholicism (cf. Kramer, 2004, p. 141; Hesseling, 1933, p. 56; Martinus, 1990, p. 130). Maurer (1988, p. 183) provides table- and kitchen-related vocabulary as an example of a semantic field with overwhelming Dutch influence: Pap. *stul* (< *stoel* ‘chair’), *taflak* (< *tafellaken* ‘table cloth’), *sèrbète* (< *servet* ‘napkin’), *fòrki* (< *vork* ‘fork’), *glas* (< *glas* ‘glass’), *kopi* (< *kop* ‘cup’), *skòter* (< *schotel* ‘dish’), *panchi* (< *pannetje* ‘little pan’) all have Dutch etymologies.

### 10.5.2 Phonological adaptation of Dutch loans

The type and degree of phonological adaptation of Dutch loans can give interesting clues as to the time-depth of certain loans. As a rule, the more recent a loan, the less it will have been adapted phonologically. For instance, a loan which has maintained the original etymological velar fricative /x/ (e.g. Pap. *brùg* /brax/ ‘bridge’) can safely be assumed to be more recent than one that doesn’t (e.g. Pap. *hanchi* ‘alley, street’ < Dutch *gangetje*).

The borrowing of Dutch nouns is characterized by the non-meaningful transmission of the Dutch diminutive suffixes, with non-standard Dutch *-ie* (standard Dutch *-je*) corresponding to Pap. *-i* (e.g. Pap. *buki*, *hopi*, *tiki* < Dutch *boekje*, *hoopje*, *tikkie*), *-tje* to *-chi* (e.g. Pap. *borchi*, *sunchi*, *purunchi* < Dutch *bordje* ‘plate’, *zoentje* ‘kiss’, *sproetje* ‘freckle’), *-sje* to *-shi* (e.g. Dutch *poesje* > Pap. *pushi*) and *-pje* to *-pi* (Dutch *oompje* > Pap. *ompi*) (cf. Lenz, 1928, p. 151; Andersen, 1974, p. 209; Kramer, 2004, p. 142, 143). The Dutch diminutive suffixes were integrated as a phonological part of the root but convey no diminutive meaning in Papiamentu. Papiamentu diminutives are formed with the postnominal adjective *chikí*. Thus, Pap. *sunchi* ‘kiss’ → *sunchi chikí* ‘little kiss’.

Some Dutch loans have either been adapted to Spanish phonologically (e.g. earlier Pap. *dirèkt* ‘direct(ly)’, now *direkto*) or been replaced by Spanish equivalents (e.g. earlier Pap. *sport*, now *deporte*) (Martinus, 1990, p. 140; van Putte, 1999, p. 64).

### 10.5.3 Calques

One of the most fruitful ways in which Papiamentu has borrowed from Dutch is by means of calquing, as already illustrated with the example *verlief riba* ‘in



love with' above in 10.4.2. The following two examples also illustrate idiomatic expressions calqued from Dutch (Fokker, 1914, p. 57, spelling as in the original).

---

#### Calques on idiomatic expressions

---

Papiamentu	<i>é no por yuda</i>	<i>é ta kai flau</i>
Gloss	3SG.PRO NEG can help	3SG.PRO PRES fall faint
Dutch	<i>hij/zij kan't niet helpen</i>	<i>hij/zij valt flauw</i>
English	s/he can't help it	s/he faints

---

Not infrequently, entire extended idioms are replicated, often consisting of Dutch words or phrases, underlined below, embedded in a Papiamentu morphosyntactic framework (Brenneker, 1961, p. 148, 149; cited in Wood, 1970, p. 88):

---

#### Adaptations of extended idioms

---

Papiamentu	<i>e aap ta bini uit de mouw</i>	<i>nan ta stuur mi met een kluitje in het riet</i>
Gloss	DET monkey PRES come out of sleeve	3PL PRES send 1SG with a little clump into the reeds
Dutch source	<i>de aap komt uit de mouw</i>	<i>ze sturen me met een kluitje in het riet</i>
English	'The truth comes out.'	'They send me on a wild goose chase.'

---

Some Papiamentu lexemes have undergone meaning extension to cover the semantic properties of a corresponding Dutch lexeme. For instance, Pap. *gai* 'rooster' gained the secondary meaning of 'trigger', reflecting the double meaning of Dutch *haan* 'rooster; trigger', and Pap. *por* 'to be able to' has the additional sense of knowing a language, e.g. *mi no por hulandes bon* 'I don't speak Dutch well' (Maurer, 1988, p. 276), reflecting Dutch *kunnen*.

#### 10.5.4 Discourse markers and modal particles

Due to their pragmatic salience discourse markers are particularly prone to borrowing in situations of active bilingualism (Muysken, 2000, p. 112, 113; Matras, 2009, p. 157), such as on the ABC Islands. Among the Dutch-derived discourse markers most frequently heard in mainstream Papiamentu discourse we find Pap. *tòg* 'still, nonetheless', *ègt* 'really', *nèt* 'just', *eigenlùk* 'actually', *dus* 'thus, so', *kijk* 'look' and *gewon~hewon* 'simply'.

### 10.5.5 Prepositions and verb particle combinations

Prepositions play an important role in Papiamentu (unlike in many other Caribbean creole languages), as they do in Dutch. As a consequence, Papiamentu has borrowed and calqued quite copiously from Dutch in the prepositional domain, even though the most basic prepositions are of clear Iberian origin (Pap. *na* 'locative', *pa* 'for', *di* 'of', *riba* 'on', *den* 'in', *ku* 'with', *te* 'until').

Several Papiamentu prepositions take on functions modeled or calqued on their Dutch equivalents. The non-transparent use of the locative preposition *riba* 'on' for instance, is modeled on Dutch *op*:

Pap. *riba dos di april* ≈ Dutch *op twee april* 'on April second'  
(Kramer, 2004, p. 150)

A Dutch-derived preposition sometimes co-exists with an original Papiamentu equivalent, allowing speakers to make subtle semantic distinctions by choosing either one or the other. For instance, whilst 'to talk about' is usually translated in Papiamentu as *papia ofer* 'talk about' (with Pap. *ofer* < Dutch *over*), the variant *papia riba* is used in the sense of 'to gossip' (Guiselle Starink-Martina, p.c.).

Furthermore, contact with Dutch seems to have caused an increasing preference for prepositional phrases in place of a (grammatically more 'correct') direct object construction. Consider, for instance, the following cases:

---

#### Dutch-influenced verb + preposition constructions

---

Adapted Papiamentu	Dutch	Traditional Papiamentu
<i>pone presion riba un hende</i>	<i>druk op iemand zetten</i>	<i>presioná un hende</i>
put pressure on a person	pressure on someone put	pressure a person
'to put pressure on somebody'	'pressure someone'	
<i>el a bula ófer trankéra</i>	<i>Hij sprong over het hek.</i>	<i>el a bula trankéra</i>
3SG.PRO PERF jump over fence	He jumped across the fence	3SG.PRO PERF jump fence
'he jumped over the fence'		'he jumped the fence'

---

Papiamentu also directly borrowed several prepositions from Dutch. However, the majority of these appear to occur only in verb + preposition combinations based on Dutch particle verbs in which the preposition typically has a non-transparent meaning.

*Lo mi bèl bo op.*  
FUT 1SG ring 2SG up  
'I will call you.' (cf. Dutch *opbellen*) (Kouwenberg & Murray, 1994)

*tin algun hende dor*  
 have somebody through  
 ‘to be on to somebody’ (cf. Dutch *iemand doorhebben*)

In more traditional Papiamentu such particle verbs are largely absent, with the exception of constructions involving a movement verb and *bèk* ‘back’ (interestingly of English origin):

*Wanchu a bini kas bèk.*  
 John PERF come home back  
 ‘John has come back come.’

### 10.5.6 Passive and the agent phrase

Many Dutch prepositions borrowed into Papiamentu are not used independently, but only in combination with native verbs. Consider the composed preposition Pap. *dor di* [Dutch *door* ‘through, by’ + Pap. *di* ‘of’], which is used exclusively in the Papiamentu passive of the type [TMA marker + passive auxiliary + past participle].

---

#### Passive structures in Papiamentu

---

	<i>wordu</i>		<i>dor di</i>	
<i>e skòl</i>	<i>ta</i>	<i>finansiá</i>		<i>(e) gobièrno</i>
	<i>ser</i>		<i>pa</i>	
DET school	PRES be	financed by	(DET) government	
‘the school is (being) financed by the government’				

---

As noted previously (10.4.2), the Dutch-derived passive auxiliary *wòrdu* was integrated in Papiamentu during the 19th century. Papiamentu also acquired an Iberian-derived auxiliary, *ser*. Agentive BY-phrases are introduced either by Dutch-derived *dor di* or by the native preposition *pa* ‘by’. In principle, *dor di* is used then when the auxiliary is *wòrdu*, whereas *ser* is accompanied by agentive *pa*. However, there is no rule prohibiting the use of *wòrdu* in combination with *pa*, or of *ser* with *dor di*.

*E ministerio ta wordu koordiná pa ...*  
 DET ministry PRES be coordinated by  
 ‘the ministry is being coordinated by (...)’  
 (Telenotisia / TeleCuraçao 27–09–2012)

*Nos ta ser uzá dor di señor*  
 1PL.PRO PRES be used by gentleman  
 ‘We are being used by [you] sir.’

### 10.5.7 Other function words borrowed from Dutch

In addition to the passive auxiliary *wordu* and the BY-phrase preposition *dor di*, the relatively few functional items that Papiamentu borrowed from Dutch include the etymologically hybrid conjunction *dor ku* ‘because’ (calqued on Dutch *doordat*), the coordinate conjunction *of ‘or*’, and the modal verb Pap. *mag* ‘to be allowed to’ < Dutch *mogen*. The Papiamentu progressive expression [*ta bezeg ta* + verb], calqued on Dutch [ *bezig zijn te* + verb], may also be mentioned here.

Perhaps with the partial exception of the diminutive suffix discussed in Section 10.5.2, Papiamentu did not borrow any derivational affixes from Dutch. In the domain of inflection, the Dutch prefix *ge-*, creating a past participle from a verb, was borrowed into Papiamentu (see 10.6.2).

## 10.6 Morphological integration of Dutch nouns and verbs

Having discussed the quantity and quality of Dutch borrowings in Papiamentu, this section looks at how these borrowings interact with native Papiamentu morphology.

### 10.6.1 Nouns and nominalizations

Dutch loanwords follow the standard Papiamentu patterns for plural marking. Thus, Dutch-derived nouns in Papiamentu will only receive the Pap. plural marker (*-nan*) if the noun is [+plural, +definite]: Pap. *dos buki* two-book ‘two books’, but *e dos bukinan* [DEF-two-book.PL] ‘the two books’ (with Pap *buki* < Dutch *boek*).

There is no fully productive denominal morphology in Papiamentu. As noted previously, diminutives are built analytically using the adjective *chiki* ‘small’; the Dutch diminutive suffix *-je* (and variants) on words such as Pap. *sunchi* ‘kiss’ derived from the Dutch morphologically complex *zoen-tje* (kiss-DIM) is morphologically reanalyzed as a chunk and therefore non-meaningful. The semi-productive Iberian-derived prefixes *re-*, *des-*, and *in-* do not attach to Dutch loanwords (Kouwenberg & Murray, 1994, p. 30).

Several Dutch compound nouns, typically of the type [modifier + head noun] (i.e. left-branching), have been integrated into Papiamentu by translating the individual components and rendering them as right-branching [head noun + *di* + modifier] compounds (Fokker, 1914, p. 56, 57; Hesseling, 1933, p. 46; Van Putte, 1999, p. 95):

<i>bentana di dak</i> ‘loft-window’	< Dutch <i>dakraam</i>
<i>koker di pèn</i> ‘penholder’	< Dutch <i>pennenkoker</i>
<i>rampi di skeif</i> ‘sliding window’	< Dutch <i>schuifraam</i>
<i>outo di hür</i> ‘rental car’	< Dutch <i>huurauto</i>
<i>stul di zoya</i> ‘rocking chair’	< Dutch <i>schommelstoel</i>

We do also find wholesale borrowings of Dutch compounds (Kouwenberg & Murray, 1994, p. 34). These are not always transparent, however.

Dutch	Gloss	Papiamentu
<i>soldeer-bout</i>	welding iron	<i>sòldeerbout</i>
<i>thee-blad(je)</i>	tea tray	<i>teblachi</i>
<i>regen-bak</i>	(rain) water cistern	<i>rembak</i>
<i>kies-schijf</i>	rotary dial	<i>kiskeif</i>
<i>knoopsgat</i>	button hole	<i>konoskat</i>

There are no indications that the integration of these left-branching compounds has affected or will affect the right-branching orientation of Papiamentu compounds.

The two most productive Papiamentu derivational suffixes, *-mentu* (for action nouns) and *-do* (for agentive nouns), etymologically derived from Iberian *-mento* and *-dor*, combine freely with Dutch loan verbs, resulting in a large and potentially open class of action nouns (see examples below; cf. Birmingham, 1970, p. 127; Dijkhoff, 1993, pp. 148–151; Kowallik & Kramer, 1994, p. 157). Pap. *-do* and *-mentu* are insensitive to phonotactic patterns, attaching to Dutch loan verbs irrespective of phonological shape, tone melody, or syllable length.

V + *-mentu* ‘the act of V-ing’

*wak* ‘to watch’ > *wakmentu* ‘the act of watching’

*ferlof* ‘to engage’ > *ferlofmentu* ‘engagement’

V + *-dó* ‘somebody who V-s’

*yag* ‘to hunt’ > *yagdó* ‘hunter’

*hür* ‘to rent’ > *hürdó* ‘tenant, renter’

Some Dutch loans have been combined with – clearly non-productive – Iberian-derived nominalizing suffixes resulting in hybrid nouns, such as Pap. *bukeria* ‘book store’ from *buki* ‘book’ + *-eria* (Dijkhoff, 1993, p. 84; cf. Van Putte, 1999, p. 95; Kouwenberg & Murray, 1994, p. 29).

### 10.6.2 Verbs and inflection

The Iberian-derived gerundive suffix *-ndo* is becoming increasingly frequent in Papiamentu as a means of marking progressive aspect more emphatically: Pap. *papia* ‘to talk’ > *papiando* ‘talking’. Like the derivational morphemes *-mentu* and *-do*, the inflectional suffix *-ndo* too can be attached to any Dutch verb, regardless of syllable length: *zuai* ‘swing’ > *zuayendo* ‘swinging’, *fèrf* ‘paint’ > *fèrfiando* ‘painting’, etc. (Kouwenberg & Murray, 1994, p. 20; Sanchez, 2005, p. 75).

Unlike the suffixation of *-nan*, *-mentu*, *-dó* and *-ndo*, the formation of past participles is in fact sensitive to the morphophonological structure of the verb. This formation consists of a stress shift from the penultimate to the final syllable of the verb, e.g. Pap. *morde* ‘to bite’ > *mordé* ‘bitten’, *kanta* ‘to sing’ > *kantá* ‘sung’, etc.), but only those loan verbs that have been adapted to the native stress-tone pattern of bisyllabic verbs – stress and low tone on the penultimate syllable and a short high tone on the final syllable – appear susceptible to this mechanism. A good number of Dutch-derived verbs have in fact been adapted to the native stress-tone pattern; in analogy with Iberian-derived verbs (most of which end in *-a*), these typically end in a non-etymological *-a*, such as Pap. *lesa*, *fula*, *harka* from Dutch *lezen*, *voelen*, *harken* (cf. Hesseling, 1933, p. 52; Andersen, 1974, p. 206). These adapted loan verbs, then, in line with native verbs, form their past participles by means of a stress shift: Pap. *harka-harká* ‘rake-raked’, etc.

To accommodate past participle formation from loan verbs not adapted to the native stress-tone pattern, on the other hand, the Dutch-derived past participle prefix *he-* (and variants) was borrowed into Papiamentu: Pap. *wak-hewak* ‘watch-watched’, Pap. *fangu-hefangu* ‘catch-caught’, etc. (Martinus, 1990, p. 132, 133; Maurer, 1988, p. 68, 69; Kouwenberg & Murray, 1994, p. 20; Dijkhoff, 1993, p. 88). The *he-* prefix has become partially productive in Papiamentu: non-Dutch-derived verbs that do not conform to the Papiamentu bisyllabic stress-tone pattern also take it to form participles. For instance, the past participle of Pap. *dal* ‘to hit’ (from Spanish *dále!* ‘hit him!’) is *hedal* ‘hit’ (cf. Kouwenberg & Murray, 1994, p. 20; Maurer, 1988, p. 69). The Papiamentu past participle formation strategies are summarized in Table 10.2.

Table 10.2 Papiamentu past participle formation strategies

Iberian and bisyllabic			Dutch and monosyllabic		
Basic	Participle		Basic	Participle	
<i>morde</i>	<i>mordé</i>	‘bite’	<i>fèrf</i>	<i>hefèrf</i>	‘paint’
			<i>wèlder</i>	<i>hewèlder</i>	‘weld’
<i>harka</i>	<i>harká</i>	‘rake’	<i>tren</i>	<i>hetren</i>	‘train’
			<i>dal</i>	<i>hedal</i>	‘hit’

## 10.7 Papiamentu in the Netherlands

As noted previously, ca. 100.000 Antillean immigrants reside in the Netherlands, most of whom speak Papiamentu at home. So far, little is known about Papiamentu as an HL in the Netherlands or about Papiamentu-to-Dutch language shift.

Recent statistics (Extra, 2013) place Papiamentu as the sixth most frequently spoken HL in the Netherlands, following Turkish, Arabic, Berber, English and Hind(ustan)i. Unlike most speakers of these five languages, Antillean immigrants often are competent in Dutch prior to moving to the Netherlands.

Kook et al. (1996) and Muysken et al. (1996) study aspects of intergenerational language shift among Antillean immigrants. By means of bilingual parent-child dialogic reading sessions, they shed some light on language choice and functional language differentiation in Antillean migrant families in the Netherlands – ‘Antillean’ here used as a cover term for people from the ABC Islands. In a typical migrant scenario, generation  $n + 1$  will use the HL less than generation  $n$  and indeed Kook et al. (1996) and Muysken et al. (1996) confirm that this also applies to Papiamentu speaking migrants in the Netherlands. The study furthermore shows that, as is to be expected, the Dutch proficiency of migrant children correlates positively with that of their parents.

A predominant pattern in the speech of the mothers that surfaced in the bilingual parent readings is the insertion of single Dutch words in Papiamentu utterances. If the child interprets these insertions as Papiamentu, they “are likely to become borrowings in the next generation.” (Muysken et al., 1996, p. 485).

By means of a structural (or syntactic) priming experiment, Kootstra & Sahin (2018) examine the influence of Dutch on the argument structure of Papiamentu in the Netherlands. Both standard Papiamentu and Dutch have the choice between a double object construction (*e muhé ta duna e homber e bala / de vrouw geeft de man de bal* ‘the woman gives the man the ball’) and a prepositional dative (*e muhé ta duna e bala na~pa e homber / de vrouw geeft de bal aan de man* ‘the

woman gives the ball to the man'). In Papiamentu there is a very strong preference for the double object construction, whereas there is no clear preference in Dutch. Kootstra & Sahin (2014) show that, when describing a ditransitive event, Papiamentu speakers significantly more often take recourse to prepositional dative constructions after having heard a Dutch prime sentence which also contains a prepositional dative. The study thus suggests that enough contact with Dutch may in the long run result in a change in Papiamentu argument structure.

Interestingly Kootstra and Sahin (2018) show that heritage speakers behave differently from homeland speakers. Both groups are in contact with Dutch, but the role of Dutch is stronger in the Netherlands than on Aruba. The findings show that speakers in the Netherlands use more prepositional dative constructions in the baseline experiments, whereas speakers on Aruba alter their behavior more after hearing a Dutch prime than the speakers in the Netherlands.

In the remainder of this section we illustrate some linguistic features of heritage Papiamentu and Papiamentu-Dutch codeswitching. Our corpus contains (a) interviews with Papiamentu heritage speakers in the Netherlands,<sup>1</sup> (b) free conversations among heritage speakers<sup>2</sup> and (c) bilingual parent child readings (recordings pertaining to Kook et al., 1996 and Muysken et al., 1996). Although we also encountered alternational codeswitching, we will focus below on patterns of insertional codeswitching. As we will see, the patterns of insertion are qualitatively not very different from the patterns of borrowing outlined in Section 10.5.1. That is, nouns are inserted more frequently than verbs, verbs more than adverbs, etc., and Dutch insertions are predictably more frequent in semantic domains such as education.

Consider for instance an utterance whose matrix (i.e. underlying morphosyntactic structure) is strictly Papiamentu. The inserted Dutch elements – three nouns, one verb and one discourse marker – are underscored. The nontransparent use of the preposition *riba* is calqued on the Dutch collocation *reageren op* ('to react on').

*Mita studia geneeskunde, anto hopi biaha nos ta tin rollenspel ...*  
 1SG PRES study medicine so much time 1PL PRES have role.playing.game  
*dus ... ami ta arts i mi tin.ku reageer riba pashent.*  
 so 1SG.FOC PRES doctor and 1SG have.to react on patient  
 'I study **medicine**, so many times we play a **role playing game** ...so... I am a  
**doctor** and I have to **react** to the patient.'

1. The interviews were carried out in 2012 by Obi Heijer of the Radboud Universiteit Nijmegen as part of the larger Multilingual Netherlands project.

2. These were recorded by a team supervised by Marianne Gullberg at the Max Planck Institute for Psycholinguistics in Nijmegen.



Below we illustrate some insertional codeswitching patterns according to word class.

In principle, any Dutch **noun** can be incorporated freely into any Papiamentu matrix clause (Muysken et al., 1996, p. 485; cf. Kowallik & Kramer, 1994), be it in singular, or plural.

*Na Korsouno no tin **mogelijkheid** pesei.*  
 in Curaçao NEG EXIS possibility for.that  
 ‘In Curaçao there is no **possibility** for that.’  
*Tin su **voordelen**.*  
 EXIS 3SG.POSS advantages  
 ‘(It) has its **advantages**.’

Unsurprisingly, in the speech of our informants who are also students at Dutch universities, many inserted nouns relate to the academic realm (cf. for instance (21).

As for plural marking, the following patterns occur in our corpus: Papiamentu plural endings may go on Dutch nouns, even when these are already plural. However, Dutch plural endings may not go on Papiamentu nouns.

*e **muis-nan*** [DET mouse-PL]  
*e **raton-nan*** [DET mouse-PL]  
***muiz-en*** [mouse-PL]  
***muiz-en-nan*** [mouse-PL-PL]  
 \*(*e*) ***raton-en*** [DET mouse-PL]

A similar pattern holds for diminutives. Papiamentu plural endings may go on Dutch nouns, such as *muis* ‘mouse’ below, even if these have a Dutch diminutive (-*je*) ending and a plural (-*s*) ending, but Dutch diminutives may not go on a Papiamentu noun.

***muis-je-s*** [mouse-DIM-PL]  
*e **muis-je-nan*** [det mouse-DIM-PL]  
***muiz-je-s-nan*** [mouse-DIM-PL-PL]  
 \**raton-je(-s)* [mouse-DIM-(PL)]

Like nouns, Dutch verbs are also embedded freely and frequently. This happens typically by taking the stem of the Dutch verb:

*Nan ta egt **begeleid** e **mucha-nan**.*  
 3PL PRES really guide DET child-PL  
 ‘They really guide the children.’

Inserted Dutch verbs can occur in serial verb constructions typical of Papiamentu:

*A base di kiko el a bin trek su konklushon?*

on basis of what 3SG PERF come pull 3SG.POSS conclusion

‘On what basis did he finally draw his conclusion?’

*El a bai verhuis bai Zwolle.*

3SG PERF go move go Zwolle

‘He (went and) moved to Zwolle.’

If the inserted verb is a particle verb (e.g. Dutch *afroندن*, *omscholen*, etc.), Papiamentu speakers typically take the stem of the verb placing the particle post-verbally. In fact, these Dutch particle verbs appear to be inserted more often than Dutch single-morpheme verbs, which we assume relates to the fact that clear Papiamentu equivalents are often unavailable for such particle verbs.

*Poko anja pasá el a school om un tiki.*

few year back 3SG PERF reeducate a little

‘A few years back he went through retraining a bit.’

*Mi ta rond af mi skriptie anto.*

1SG PRES finish-off 1SG thesis thus

‘So I’m finishing my thesis.’

If the particle verb is transitive, depending on which constituent the speaker wants to topicalize, the speaker will either place the particle directly behind the verb and before the object, as in the last example above, or clause finally, behind the object, as in the next example, depending on what is being emphasized.

*Pero gewoon mi ta druk mi mes uit.*

but just 1SG PRES press 1SG self out

‘But I am just expressing myself.’

In the speech of some of our Papiamentu-Dutch bilingual informants the frequency of Dutch-derived discourse markers, adverbs and adverbial expressions was striking, as nicely illustrated in the next examples.

*Gewon eigenlijk per toeval m’a topa nan.*

just actually by.accident 1SG.PERF meet 3PL

‘Actually I just met him by accident.’

*Anto, ja, voor de rest gewon m’a krese na Korsou*

So yes for the rest normal 1SG.PERF grow.up in Curaçao

‘So, yeah, apart from that I simply grew up in Curaçao.’

We furthermore found common Dutch adverbial expressions such as *laat staan* ‘let alone’, *op den duur* ‘in the long run’, *over het algemeen* ‘generally’, *hier en daar*

'here and there' embedded in Papiamentu matrix clauses. Although some Dutch discourse markers (e.g. *gewon*, *eigenlúk*) are also heard in more traditional Papiamentu (cf. Section 10.5.4), they occur at a much higher rate in the discourse of Papiamentu heritage speakers in the Netherlands.

All of our informants quite accurately adhered to standard Papiamentu rules for TMA marking, also when Dutch lexemes were inserted. The only anomaly was the occasional placement of a Dutch-derived discourse marker in between the TMA marker and the verb.

*Anto sindsdien e ta gewon keda na Korsou*  
 so since 3SG PRES just stay in Curaçao  
 'So from then on he is just staying on Curaçao.'  
*Nunka m'a echt puntra mi mama-nan*  
 never 1SG-PERF really ask 1SG mother-PL  
 'I never really asked my mother and her people.'

This placement is ungrammatical in traditional Papiamentu: Pap. *e ta* (\**anto*) *keda na Korsou* 3SG-PRES-(so) he's staying on Curaçao'. In Dutch, on the other hand, discourse markers can be placed freely between any auxiliary and its complement verb.

One informant frequently used Dutch *plus* /pløs/ as a conjunction where traditional Papiamentu would use either *i* or *ku*.

*E ta dekaan plus duna studiebegeleiding*  
 3SG PRES dean plus give study.coaching  
 'He is a dean and coaches students.'  
*E ta duna hulandes plus ingles plus, ehm...*  
 3SG PRES give Dutch plus English plus ehm  
 'She teaches Dutch and English and, ehm...'

The insertion of Dutch prepositions mostly concerns idiomatic prepositional expressions (*per toeval*, *over het algemeen*, *voor de rest*, etc.), more peripheral prepositions such as *tijdens* 'during', and Dutch idiomatic expressions of the type [verb + prepositional phrase] inserted wholesale into Papiamentu matrix clauses.

*Nan a buig nan mes derover.*  
 3PL PERF bend 3PL self over.it  
 'They considered it.' (cf. Dutch *zich (d)erover buigen*)

Although in the foregoing we have focused on insertional codeswitching, our corpus also contains a considerable amount of alternational codeswitching:

*Daarna m'a stroom VWO, maar het lukte me niet.*  
 afterwards 1SG.PERF stream VWO, but it managed me not  
 'Afterwards I proceeded to VWO, but I didn't manage.'

## 10.8 Summary and conclusion

The case of Papiamentu presented in this chapter illustrates the typical features of post-colonial HLs quite well. The key features of changes in the communities where these languages are spoken include:

- Extensive lexical borrowing, including calques and phrases.
- Shift processes leading to intensification of contact-induced structural changes.

Such features can probably be found in different post-colonial HLs more generally, but this probably needs to be studied more systematically, e.g. in the communities with an Indian or Pakistani background in the United Kingdom. It certainly holds for the other major post-colonial HLs in the Netherlands, Sranantongo from Surinam and Moluccan Malay from Indonesia.



## The political dimension of heritage languages

### Endangered languages, language rights, and the preservation of diversity

#### 11.1 Introduction: The politics of diversity management

This chapter is concerned with the political dimension of HLs. From a political perspective, the status and recognition of these languages are part of what may be called *diversity management*: the response of institutional entities, ranging from a state government to a school administration, to language diversity and multilingualism within their boundaries. We also must realize, of course, that no power center in an institutional entity (a school, a country) is entirely separate from its constituencies; in this case the various language communities that are part of that entity. Thus, diversity management is necessarily a political process, often subject to complex negotiations and conditioned by language status and prestige.

Diversity management is currently a much-debated topic, within academia, in the political arena, and in society at large, all over the world. It would carry much too far to engage this topic in anything like its full breadth; here we can only touch on a few pointers. Why focus on HLs in the first place? We have seen in this book that they throw interesting light on the topics of language acquisition, variation, change and contact – all of which are academic concerns. However, there are also more social concerns that carry into the political arena: HLs are often spoken by vulnerable groups and paying attention to an HL can thus be a way of strengthening the position of those groups and ensure access of its members to education and other benefits.

We will begin this chapter by contrasting the HL frame with two other common frames of reference in public discourse: the Babylon frame and the Tsunami frame (Section 11.2). Then we discuss strategies for reversing language shift (11.3) and HL education (11.4). In Section 11.5 language documentation is discussed, and in 11.6 we return to the issue of codeswitching in relation to language maintenance. Section 11.7 concludes by focusing on the question of linguistic human rights.

**Aims of this chapter**

- To introduce the main political notions concerning HLs.
- To introduce the notions of language rights and language revival.
- To discuss the main issues surrounding HL education.
- To present the prospects of language documentation.

## 11.2 Frames of reference

As we all know, diversity management can take various perspectives, which we will describe as frames in the sections that follow. These frames play a crucial role in cultural analysis and in political science.

### 11.2.1 The Babylon frame

In one non-inclusive perspective, language diversity is perceived as a threat. This may be labeled the “Babylon” frame, as in the popular imaginary pictures of the destruction of the tower of Babel. The frame of Babylon involves a tower that is being shattered, with chaos and fleeing people everywhere, crucially involving multilingualism.

Babylon symbolizes confusion and destruction here: all languages are spoken together, and the resulting chaos leads to destruction. The political implications of adopting this frame as a basis for policy is often that a larger organization, often the state, tries to get rid of all languages other than the ones condoned by the powers that be. Reasons given are often that this is supposedly in the interest of national security, but also supposedly in the interest of the advancement of all citizens and of universal education. The French sociolinguist Calvet (1974) has coined the term *glottophagie* ‘lit. language eating’ for this policy. Skuttnab-Kangas (2002, p. 46, 7) uses the terms *linguicism* and *linguistic genocide*. Cobarrubias (1983) distinguishes in this respect between actively attempting to kill a language, letting a language die by depriving it of resources, and unsupported coexistence. Differences between these options are often slight.

The French state, from Robespierre and Napoleon until recently, has tried to impose standard French on all its citizens, to the detriment of languages such as Breton, Basque, Provençal, and Alsatian German. Only more recently a more inclusive language policy recognizing these and other languages as HLs has been put into place. The French example was followed in many countries, notably Turkey in the post-Ottoman period, when Turkish was imposed on all citizens, to the detriment of languages such as Kurdish, Greek, Armenian, and the languages of the Caucasus. Again, Turkey has also changed directions, and now Kurdish is allowed

in several domains. However, the country is still far from presenting itself as the truly multilingual nation that it was in Ottoman times.

### 11.2.2 The Tsunami frame

A second frame regarding the position of HLs is the Tsunami frame, in which the terrible image of this natural disaster is used to symbolize mass immigration. In the Tsunami frame, an originally more homogeneous territory faces an enormous disaster, in the form of a wave. Populist politicians sometimes use this frame, including the actual word ‘Tsunami’ to refer to waves of immigrants.

In this frame, it is not diversity as such that is perceived as problematic, but only ‘new’ diversity that results from migration. “Waves” of migrants bring new languages, and the traditional languages of a country – both dominant languages and regional minority languages – risk being overrun. In this frame, “traditional” HLs have a different status from “new” HLs. This type of model, even though not explicitly associated with the Tsunami frame, is found in many European countries and broadly characterizes the language policies of the European Union. There is tolerance of or even active support for languages such as Sorbian (a Slavic HL spoken in eastern parts of Germany) much more than for Somali. In northern Italy regionalism and appreciation for local varieties may coexist with xenophobia which creates a dislike for new HLs, immigrant languages such as Akan and Somali. At the same time, traditional dialects and local language varieties are cherished. Associated with this approach is the perception that traditional HLs are quaint and contribute to the national cultural heritage, while new HLs threaten this heritage.

### 11.2.3 The Heritage frame

A more inclusive model is one in which both traditional and immigrant HLs are recognized as such. This model, which we may label the ‘Heritage’ frame, is found particularly in countries with a strong tradition of immigration, in which immigration is perceived as somehow constitutive of the national cultures: Australia, Canada, and the United States. In contrast, many European countries do not derive their identity from immigration. Altogether, there is growing worldwide recognition of our world cultural heritage, particularly in countries with an immigration tradition, and this includes of course HLs. The Heritage frame involves tradition, music, traditional costumes, architectural remains, cooking and traditions.

The term of ‘Heritage Language’ itself is interesting in this respect: it suggests the existence of languages that are not HLs. But what languages could that refer to? At the one end of the spectrum, these would include dominant national languages. Thus, in the US, which has a very inclusive definition, all languages but English



are sometimes considered HLs (see Chapter 1 for a discussion on differences with the Canadian approach). Consider two quotes from the “Alliance for the Advancement of Heritage Languages” website:

The Alliance for the Advancement of Heritage Languages (the Alliance) is committed to advancing language development for HL speakers in the United States as part of a larger effort to educate members of our society who can function professionally in English and in other languages.

The following definition of HLs is given on the site:

In the United States, a “heritage language” is any language other than English that is spoken by an individual, a family, or a community. Heritage languages can include immigrant languages, spoken by immigrants arriving in the United States (e.g., Spanish); indigenous languages, spoken by peoples who are native to the Americas (e.g., Navajo); and colonial languages, of the various European groups that first colonized what is now the United States (e.g., French and German). Heritage language speakers have various levels of proficiency and connection to the language and culture.

In countries with less inclusive models, immigrant languages would also not be thought of as HLs, which is ironic given the fact that the term HLs began to be used in Canada where it was especially reserved to refer to immigrant languages and explicitly excluded the indigenous languages.

Examples of HLs in the very broad definition on the US site would be:

*Languages indigenous to the U.S.:* Anishinaabemdaa, Chinuk Wawa, Denaakk’e Athabaskan, IchCinshKiin, Navajo

*Latin American and European Languages:* French, German, Italian, Portuguese, Russian, Spanish, Ukrainian

*East Asian, South Asian, and Pacific Island Languages:* Chinese, Hindi, Ilokano, Japanese, Korean, Persian, Samoan, Tibetan, Tongan, Urdu

A second list that circulates on the internet of HLs even includes languages without current speakers, such as Sanskrit, which have mostly a cultural and religious value:

Cantonese	Choctaw	Croatian	Czech
French	German	Hindi	Korean
Mandarin	Russian	Sanskrit	Tagalog
Tamil	Urdu	Yiddish	

In this definition, languages such as Sanskrit or Latin would also qualify as HL, something which is clearly up to individuals to decide. HL learners are those who

“...have familial or ancestral ties to a particular language and who exert their agency in determining whether or not they are HLLs (HL learners) of that HL (HL) and HC (heritage culture)”, in the definitions by Hornberger & Wang (2008, p. 27).

Although we claimed in Chapter 1 that the dominant language of the country is not a heritage language, at least not in the sense that we were interested in, namely a language in close contact with other language, others could argue that the dominant language can be a heritage language in the definition by Hornberger and Wang depending on the perspective the speakers take on the position of their heritage language. In the US, English is not considered a HL, but in France, French would certainly be considered central to the nation’s cultural heritage, and if France would adopt this terminology, the French would probably not exclude French. Thus the very notion of ‘heritage’ has political dimensions.

### 11.3 Reversing language shift and indigenous language revival

From the perspective of the speaker, an HL is part of a community tradition and a family history. From a broader perspective, however, HLLs are part of world cultural heritage. One of the astonishing sources of richness of humankind is the sheer diversity of languages. In their words, their patterns, and the cognitive systems revealed through these words and patterns, we see great diversity.

This diversity was little known to the general public until recently, although language scholars have been confronted with it at least since the European colonial expansion in the 16th century. Also, many theoretical models in linguistics and psychology have taken the basic uniformity of the human cognitive and linguistic systems as their starting point. It remains hotly debated how to model language diversity within linguistic theory, but it is clear that there are many ways in which languages differ. Nonetheless, languages die at an alarming rate, and smaller communities undergo language shift (see Chapter 3), leading to the loss of diversity on a global scale.

Organizations such as UNESCO, committed to promoting linguistic diversity and multilingualism, have stressed that this very diversity is an essential part of the world’s cultural heritage. Efforts are made towards the maintenance and revival of indigenous languages in different parts of the world, the documentation of language endangerment (e.g. through the publication of language atlases), the promotion of best practices of language revitalization, etc.

As an example, let us look at the Mayan language family, with about six million speakers in Mexico, Guatemala, Belize, Honduras, and El Salvador. Some members of the family may have as few as 30 fluent speakers, while others have over a million (England, 2003). The larger languages in the family include K’iche’,

Yucatec Maya, Q'eqchi', and Mam. The Maya movement in Guatemala aims for cultural reaffirmation of this oppressed group and has made language a central concern in its efforts to preserve and partially revive the Maya heritage of the region. Several important choices have to be made in such revival programs, choices involving both technical linguistic issues and language ideology.

A first issue is localism versus unification. On the one hand, a variety should be adopted that is close to the locally used forms of the language; this would plead for recognizing dialectal variants. On the other hand, revival movements often favor unification, arguing that the fragmentation of the language cluster into a host of different small varieties makes none of them are viable by themselves.

A second issue concerns the relation with the dominant national language. How are borrowings dealt with, and to what extent should they be accepted? Should the orthography resemble that of the national language, or not at all? The illustration below contains some of the symbols of the old Mayan script, but this is not suitable for most contemporary communicative purposes.



A third question concerns the effects of language change. Should the variety that is taught in school reflect older and possibly morphologically more complex forms of the language (which are assumed to be more pure and traditional), or the form of the language as currently spoken at home?

A similar range of issues can be found in immigrant HLs that do not have an official status in the countries of origin. The varieties of Berber in Western Europe are a typical example. With immigration from Morocco and Algeria to France, Belgium, the Netherlands, and Germany, several Berber languages are now spoken in these countries. Although Berber became a constitutional national language of Algeria in 2001 and a constitutionally official language of Morocco in 2011, there is still discussion about which varieties should be recognized as standard, as well as about which writing system should be adopted.

The Berber language family consists of at least:

- a large cluster of northern Berber varieties, including Tarifit (the language of the Rif mountains in Morocco) and Kabyle (spoken in Algeria)
- Tuareg Berber in the south of Morocco and Algeria
- Isolated varieties spoken in Libya and Egypt

Precise subdivisions are a subject of scholarly debate. Within all these clusters, there is considerable variation, and recent attempts at standardization and unification have not yet been fully successful.

There is also the question of orthography. Older sources often represent Berber in Arabic script, while at present the main split is between using a western-style Latin script or the Tifinagh script, which is based on early inscriptions. In Morocco Tifinagh is the official script, but it does not enjoy widespread use. In actual practice, use of Latin script is quite frequent.

Since Berber has been subordinate to varieties of Arabic now for many centuries, it has undergone considerable influence from that language. Thus, another issue is whether ‘pure’ or ‘street’ Berber should be used in revitalization efforts, including in education. The illustrations below indicate (a) the STOP traffic sign in the northern Moroccan town of Nador (photographed in 2003, it has since been removed) in both Arabic and Tifinagh writing; and (b) a correspondence table between Latin, Arabic, and Tifinagh script.



r	ⵓ	ⵝ	dj	ⵢ	x	a	ⵏ	.
s	ⵔ	ⵉ	γ	ⵍ	ⵎ	u	ⵏ	:
ʃ	ⵔ	ⵉ	h	ⵎ	ⵏ	i	ⵏ	ⵉ
t	ⵏ	ⵏ	h	ⵎ	ⵏ	e	-	ⵉ
.	ⵏ	*	y	ⵏ	ⵏ	b	ⵏ	ⵏ
t	ⵏ	E	j	ⵏ	I	c	ⵏ	c
w	ⵏ	ⵏ	k	ⵏ	ⵏ	d	ⵏ	ⵏ
x	ⵏ	X	l	ⵏ	ⵏ	.	ⵏ	v
z	ⵏ	X	m	ⵏ	ⵏ	d	ⵏ	E
z	ⵏ	X	n	ⵏ	ⵏ	f	ⵏ	X
			q	ⵏ	Z	g	ⵏ	X
			e	ⵏ	ⵏ	gw	ⵏ	ⵏ

## 11.4 HL education

HL education thrives best in an additive bilingualism model involving multicompetent learners with complex identities (see also Cummins, 1981, 1989). Students participate in education in the dominant language but in addition are motivated to follow classes in the HL (Brinton et al., 2008). Kupisch and Rothman (2016) show that if HL speaking children are educated in their HL, their competence tends to be the same as that of monolingual children.

Chinese language classes outside China provides a typical example. Chinese HL schools have existed in many countries, and for many years. In the United States the first HL school was established in San Francisco in 1886 (Chao, 1996, cited in Liu, 2013). Now there are well over a thousand Chinese HL schools in the United States, and they have considerable support from the local Chinese communities. It remains to be established whether the input provided by these schools is sufficient for children to acquire Chinese well, or whether other types of exposure to Chinese are also needed and are perhaps more significant.

Several issues need discussion concerning these efforts, treated here in separate sections. How is HL education organized and supported? Which variety of the HL is taught? If there is multilingualism in the home country, is it the dominant language that is taught, or can it be a home vernacular language?

### 11.4.1 Organization and support

A first issue is the organization of HL education and the source for its support. Often the education programs are organized by community organizations. Students get instructed in the dominant language in school during the week and most of the day, but there are special classes on Saturdays or in the late afternoons where the HL is introduced.

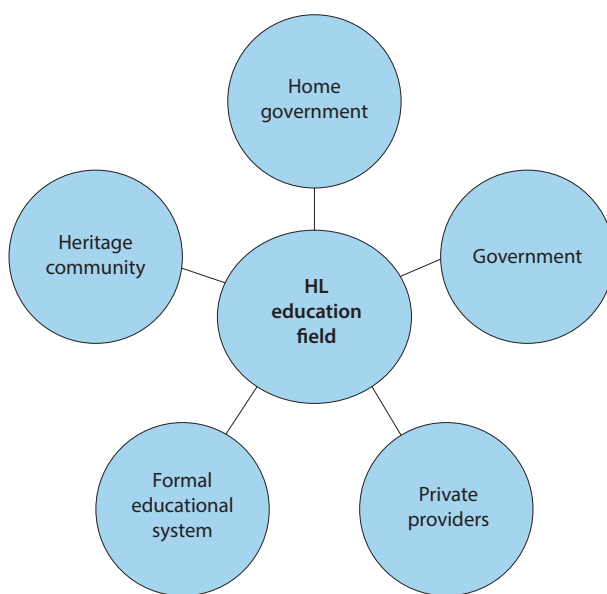
In various countries HLs are a subject in government-sponsored public education. The United States has a long tradition of local HL education programs, and after Cuban immigration in the 1960s this gained impetus, leading to the establishment of bilingual education programs in many states. The enhanced role of Spanish played an important role in this. However, around the turn of the millennium, this started changing, and now there is only very limited access to public HL education. Local governments have increasingly stressed the need for instruction in English. In the UK, there are many what are termed 'community language schools' or 'complementary schools', some of which receive some funding from the home country embassies.

In many countries, HL speakers are organized in local HL community groups. Sakuma (2013) presents a sociolinguistic study of language ideology, culture and ethnicity among Japanese immigrants and their descendants (hereafter, Nikkeis)

in Brazil. They gather at a local Japanese cultural association, searching for what it means to be ‘Japanese’. Nikkeis are a prestigious minority in Brazil. Members of the association are in constant negotiation, trying to strike a balance between the symbolic values of Japanese, the pragmatic need for Portuguese, as well as their own language competencies. Ethnicity, culture and language define and redefine one another as they interact and transform over time.

For some immigrant HL communities, the national government of the country of origin is a source of support. The Turkish politician Tayyip Erdoğan visited Germany several times to give speeches to large immigrant audiences in Cologne and Düsseldorf. While his main purpose was to convince Turkish immigrants to vote (for him) in Turkish elections, he also stressed that these immigrants should retain their Turkish heritage. While the Turkish government has promoted appropriate Islamic religious instruction abroad, it has not explicitly set up a program of language classes. In contrast, several European governments have sponsored such classes abroad for children of Portuguese, Spanish, etc. migrants living in other European countries. Currently, the People’s Republic of China is strengthening the knowledge of Mandarin in Chinese ex-patriate communities throughout the world.

We can schematize the following forces in the field of HL education: the home country government, the national government of the country of migration, the mobilization of the Heritage community, the formal educational system, and possible private providers such as religious organizations, charities or companies.



**Figure 11.1** Parties engaged in the field of HL education (based on the schema in Brecht and Walton, 1994)

### 11.4.2 Varieties of the HL taught

A second important issue is the choice which variety of the HL is to be taught in the special program. When you teach Spanish to HL speakers in a Texas classroom, do you teach standard Spanish, or do you teach the varieties close to what the students may hear at home, or do you teach both? Local Spanish spoken in Texas has some of the non-standard features of Spanish dialects of northern Mexico, but in addition has undergone semantic influence from English, as in *Hablame p'atrás* 'Call me back' (lit. 'Speak+to me to back'). This is comparable to the constructions described by Irizarri van Suchtelen (2016) and discussed in Chapter 8. Also, many speakers will frequently code-switch between Spanish and English. Flores and Toro (2000) argue that the distribution of null subjects in US Spanish is best predicted by the speaker's dialect background (rather than, for instance, by the length of stay in the US and exposure to English).

Loyalty to the local community and attempts to link to the language spoken at home argue for the teaching of a local variety, while interests in promoting links to the outside community (e.g. with homeland in case of immigrants, making it possible for example to study at a university there or do business) would be an argument for using the standardized variant. Another factor may be that the actual variant of the HL that a community speaks may not be perceived by the speakers themselves as prestigious, so they may well select a more standard variety for use in educational programs.

### 11.4.3 Dominant language from home country or home vernacular language?

A third, related, issue concerns the actual HL that is taught in case there is a choice, especially between the dominant language in the home country and a vernacular language.

In Chinese schools, Mandarin has gradually replaced Cantonese as the language of instruction from the 1970s onward. Other Chinese languages, such as Hakka, Fuzhouese, or Taiwanese, are not taught because they are considered to have no value in the professional development of the immigrants, even though many immigrants speak those varieties at home.

The history of HL education in the Netherlands also illustrates this choice. About fifty or sixty years ago, large numbers of workers were recruited from Mediterranean countries such as Morocco and Turkey to work in the expanding industries of countries like Germany, Belgium, and the Netherlands. This recruitment originally only concerned young men, often single, but soon this led

to the settlement of multilingual families. These spoke the vernacular languages of Morocco and Turkey, in addition to some Dutch.

In the 1960s and 1970s, partly under the influence of the bilingual education movement in the United States, this led to various experiments with home language education for the children of the migrant workers. It started out with volunteers teaching the children to read in their vernacular languages, but it soon was made official under the label OETC (*Onderwijs Eigen Taal en Cultuur* = Education Own Language and Culture). From 1970 onward this OETC was carried out outside of class hours, but with the support of the Ministry of Education, and in 1974 it was brought into the regular curriculum. In 1998 a new system was instituted: OALT (*Onderwijs Allochtone Levende Talen* = Education Allochtonous Living Languages), under the responsibility of the municipalities. OALT was abolished on August 1 2004, officially because there was no empirical evidence for its usefulness. From then on, emphasis was placed exclusively on teaching Dutch as a second language. Altogether neither OETC nor OALT were very successful, and it is not clear whether the Ministry ever really tried to make them a success story.

The main point we want to make, however, concerns the languages that were adopted in these programs. They were Moroccan Arabic and Turkish, disregarding the fact that many families with a Moroccan background spoke Tamazight (Berber) at home rather than Moroccan Arabic. Likewise, some of the families with a Turkish background will have spoken a variety of Kurdish or another minority language at home. On the other hand, it is possible that parents preferred education in Arabic and Turkish rather than in Berber and Kurdish.

This is contrast to one of the motivations for home language education provided by UNESCO in its well-known 1953 declaration:

It is axiomatic that the best medium for teaching a child is his mother tongue. Psychologically, it is the system of meaningful signs that in his mind works automatically for expression and understanding. Sociologically, it is a means of identification among the members of the community to which he belongs. Educationally, he learns more quickly through it than through an unfamiliar linguistic medium. (UNESCO, 1953, p. 11)

This indicates that in fact, there may have been another motive behind the original OETC programs, namely to prepare children from migrant backgrounds for a smooth return to the country from which their parents originated. When the families showed no signs of going back, as many Dutch officials had naively expected, OETC and OALT were abandoned.



#### 11.4.4 HL proficiency as a learning resource within the mainstream classroom

While HL education is primarily directed at members of smaller non-dominant communities, we also see efforts to use HLs in the education of speakers of mainstream languages. In California we find many schools where pupils with an Anglo background follow some of their classes in Spanish, so that they effectively are schooled bilingually. Terms used for this model are 'dual immersion', (the official term), 'two-way immersion', 'dual language', 'dual language immersion', and 'dual enrollment'. These terms are not exact synonyms, but nonetheless they are often used interchangeably.

Time-wise, there are 50/50 programs, which have a halfway split between English and the HL (often called 'partner languages') from the beginning, and 90/10 programs, which start with most time spent in the HL and arrive at 50/50 by third or fourth grade.

In most dual language immersion programs both native English and native Spanish speakers are enrolled, with ideally each class made up of equal numbers of both groups. There are also dual language immersion programs for other languages, including Korean, Mandarin, Japanese or French. Advantages cited include that students with an English background acquire fluent Spanish (and may receive a special certificate to prove that), but also the broadening of cultural and socio-economic horizons.

Four main types of dual language programs are distinguished in terms of the student constituency:

- a. In developmental, or maintenance, programs, primarily students who are native speakers of the HL are enrolled. These fall outside of the strict definition of two-way immersion programs.
- b. In contrast, in real two-way (bilingual) immersion programs the numbers of native English speakers and native speakers of the HL are balanced.
- c. Foreign language immersion, language immersion or one-way immersion programs enroll primarily native English speakers.
- d. In HL programs (in the narrow sense) mainly students participate who are dominant in English but whose parents, grandparents, or other ancestors spoke the HL.

In actual practice, these four types of dual language programs are sometimes combined or mixed.

## 11.5 Documentation of heritage varieties and language death

Over twenty years ago, several linguists, including Robins and Uhlenbeck (1991) and Hale et al. (1992) sounded the alarm about the rapid pace with which small languages were disappearing worldwide. An update was given in Krauss (2007). In response, the academic community has intensified the pace at which languages in different parts of the world are beginning to be documented. A few smaller and bigger funding sources have contributed to this effort, ranging from the DOBES program of the Volkswagen Foundation in Germany to the EDLP program sponsored by the Hans Rausing Foundation at SOAS London to the National Science Foundation in the United States. UNESCO has likewise provided support through various Atlas projects.

Altogether, much has been achieved in this respect, although significant work remains to be done. To give one example, Hammarström (2010) “aims to list all known language families that are not yet extinct and all of whose member languages are very poorly documented”. Hammarström concludes that particularly in Papua New Guinea several languages remain undocumented. Often these are spoken in very inaccessible regions, sometimes border zones with limited access, or regions where drug trade related, or political violence has made work difficult.

There have been several studies that document the process of ‘language death’ in small HLs. Dorian (1981), for example, studied the decay of Gaelic in East Sutherland in Scotland. Another well-known study is Schmidt (1985), which focuses on the Dyirbal language in northern Australia. Dyirbal is well known to linguists because R. M. W. Dixon has shown that the language has unique features: it is fully ergative in not only its morphology (referring to a special way of marking the subject) but also in its syntactic arrangements (it has very free word order; Dixon, 1972). Schmidt had 52 speakers from various age groups carry out a translation test (English > Dyirbal) of about 200 sentences, covering many of the key patterns in Dyirbal. In addition, she recorded spontaneous conversations and traditional stories. Her data also included a 500-word translation task (English > Dyirbal), some comprehension tasks, informal interviews about language use, and informal observations. As she had expected, the different generations in the community differed greatly in their command of the language: the elders were fluent, the middle generation knew more of the language than they were ready to admit at first sight (but their spoken production showed that the language had undergone many changes), and the children knew very little. Schmidt concludes (1985, p. 228): “Perhaps the most important factor for Dyirbal’s decline is compulsory education in English schools. Lack of institutional support has severe consequences for Dyirbal. . . . intense contact with white European civilization has resulted in the gradual abandonment of Dyirbal as a viable means of communication.”

Schmidt reports in detail on numerous grammatical, lexical, and phonological changes in the system of Dyrirbal as an HL, ‘Young Dyrirbal’ (YD). Morphological ergativity and free word order are abandoned by the less fluent YD speakers, in favor of English-style marking of subject and object and English Subject-Verb-Object word order. Several allomorphic alternations in the case marking system are dropped. The distinction between alienable and inalienable possession is neutralized. ‘Peripheral’ case endings are collapsed or dropped altogether (1985, p. 229). Overall, clause subordination has also become much less frequent in YD. Schmidt (1985, p. 230) notes that derivational morphology is much more resistant to change than inflectional morphology. Likewise, unmarked forms such as singular forms of pronouns also resist change. These changes in a dying language are very reminiscent of the type of changes in case marking that we find in immigrant HLs, as described in Chapters 4 and 7, and may be described with the notion of metatypy introduced in Chapter 1.

While many younger people still spoke some Dyrirbal in the community when Schmidt conducted her study, Muysken did research on a moribund HL in Bolivia, Uchumataqu, which used to be spoken in Irohito, on the Desaguadero River near Lake Titicaca (Muysken, 2010b). It is now largely replaced by Aymara. Uchumataqu is not the vernacular community language of Irohito anymore, but it is not completely gone either. Many people, remember, know a number of words and fixed expressions, and there is a small circle of Uchumataqu amateurs, who try to recollect as much as possible of the language. In the 1992 census 87 out of 200 people claimed to be speakers of the language. Based on the work done in 2001 two simple *cartillas*, booklets with text for reading have been prepared for the local school, and an hour a week of Uchumataqu instruction was planned (cf. Muysken, 2002). Now the last known effective speaker, Julia Vila, who died in 2004, lies buried in a nameless grave in the small Irohito cemetery, marked with a small wooden cross and three stalks of reed. As far as we know, the community has never been very large, ranging from 7 in the earliest source to as many as 80 in the early 20th century.

To describe the features of the language is somewhat risky, since these have undergone many changes over the years. Nonetheless, the overall characteristics of the language are that it tends to be verb-final with relatively limited derivational morphology, case markers or postpositions for the oblique arguments, little if any participant-marking on the verb, and a semi-obligatory indicative particle placed anywhere in declarative sentences, but often on the verb. The following sentence from Vellard (1949, p. 151) constitutes a representative example (PA = past; IND = indicative):

*wakpa-chul tsini waki chork k'ota-kis huxk-u-chay*  
 all-together egg Guaqui big lake-to leave-PST-IND  
 'All eggs have gone to the big lake of Guaqui.'

Modifiers precede the noun, so the place name Guaqui is simply proposed to the noun it specifies. There is a past tense marker *-u-*, preceding the indicative enclitic *-chay*. This latter particle is a true fixture in any Uchumataqu utterance, almost a shibboleth and recognizable by semi-speakers.

Uchumataqu went through a period of bad fortune, showing the vulnerability of language transmission in a small community. Though the Uru as a group managed to survive the drought of the 1940s and at present are stronger than ever, they did not manage to maintain their language, a loss which is now dearly felt in the community. However, it is not so clear this was felt as a loss at the time.

Dorian (1998, p. 3) writes that 'languages are seldom admired to death but are frequently despised to death.' The question is whether the differences in prestige between Aymara and Uchumataqu contributed to the shift to Aymara in the community. Following a suggestion made by Woolard (1989), it could be that Aymara was used inside of the community more by those wanting to distinguish themselves sociolinguistically from other less socially prominent community members. However, the more prominent members of the community later took on the role of guardians of Uchumataqu. This is certainly the situation at present: the dozen or so adult males most interested in rescuing the language could be viewed like the miniature version of the Irohito Rotary Club (although economically in totally different circumstances from their counterparts in the industrial world). This is also in line with a suggestion made by Woolard (1989, p. 364) about purism as a possible channel for intra-community linguistic self-profiling.

Though in various stages of decay, both Dyrbal and Uchumataqu are HLs from the cultural perspective of the countries they are spoken in, Australia and Bolivia. Their linguistic properties reflect many of the same processes that we see in other HLs discussed in this book.

## 11.6 Codeswitching in HLs and language loss

In Chapter 4 we discussed the phenomenon of codeswitching. In HL settings, it stands to reason that every code-switch to the majority language has the potential to call to attention the fact that the HL is losing ground to that majority language. Actually, switching to that language can easily be perceived as nudging the HL one small step closer to extinction. People certainly don't always have all this in mind when they code-switch; in fact, many researchers note that speakers seem

to be hardly aware of doing it. Nevertheless, there often is some awareness of what is happening to the HL, some concern about its future. Communities differ in the extent to which practicing codeswitching is perceived as a socio-political act, and there is likely to be a correlation between the extent of tension between subgroups within the community and the degree to which codeswitching is perceived as contentious. The literature contains some masterful descriptions of such communities in which tension permeates the linguistic choices people make, and we will review one such study: Hill & Hill (1986) on the use of codeswitching with Spanish in the Mexicano-speaking community.

The Malinche dialect of Mexicano (Nahuatl), spoken in a mountainous area near the Mexican city of Puebla, contains many Spanish elements. It has been dominated by Spanish since the days of the Conquest. Spanish words have been integrated into the Mexicano system, including many function words, some derivational bound morphemes, and many fixed phrases and idioms. Along with all this lexical borrowing, there has also been considerable syntactic convergence (see Chapter 7 for this phenomenon), the synthetic Mexicano syntax (long words with lots of bound morphemes indicating grammatical functions) having become progressively more analytic (lots of separate words), like Spanish. Crucially, pretty much every Spanish content word or function word can appear in Mexicano. It is estimated that about 60% of words in Mexicano speech are of Spanish origin. Many of these words are long-established Spanish loans. One can imagine that if the intense codeswitching between Turkish and Dutch described in previous chapters would continue for another 400 years, Dutch Turkish would also end up with a sizable collection of Dutch-origin words.

Codeswitching between Mexicano and Spanish is frequent. It often seems to mark the pragmatic reasons found to be served by codeswitching the world over. In the following example, the conjunction *pero* (“but”), borrowed into the language long ago, was most likely intended to draw attention to the dramatic nature of the information being conveyed. Borrowed Spanish function words tend to occur especially at narrative peaks, often claimed to be a typical switch site. At less salient points in discourse, their Mexicano equivalents tend to be used (cf. Hill & Hill, 1986, p. 287).

*pero nimitzonilia in ihcuacon hasta onimitonih in ic onicnanquilih en español*  
 ‘but I’m telling you then I was just sweating for replying to him *in Spanish*’

(Hill & Hill, 1986, p. 354)

Interestingly, Spanish elements are also sometimes used deliberately to index modernity. In other words: Spanish is associated with modernity and codeswitching happens because of this association. The important thing to realize is that they can be used for this purpose irrespective of the degree to which they have

become entrenched as well-established loanwords. In a bilingual situation it is clear to speakers from which language a word originates, and by virtue of this every word can potentially be called on to index the values associated with that language. One could object that after 400 years of contact, many Spanish words will have become so adapted in their pronunciation that they sound 'native', and therefore not recognizable as Spanish-origin at all. However, Hill and Hill note that phonological incorporation of Spanish words is variable, but most are not adapted. In addition, all speakers of Mexicano also speak Spanish, so they will generally understand that a Spanish word and a similarly-sounding Mexicano word with the same meaning must be the same word. What all this comes down to is that Mexicano speech is always Spanish-tinged, and that people will be able to recognize how 'hispanicized' it is.

Community members differ in the degree to which they are dominant in Spanish or Mexicano. The crucial bit of information is that knowing better Spanish and knowing better Mexicano each come with value judgments, roughly embracing modernity in the former case and tradition in the latter. This way, everyday language use becomes an ideological battleground. Malinche speakers seem to monitor the extent of Spanish-origin elements they use in their Mexicano discourse. Apparently, increasing the number of Spanish elements can be used to convey a modern identity. Likewise, speakers are also shown sometimes to be avoiding the use of Spanish words and thereby convey local pride or belongingness. Makiyara (2001) demonstrates the same phenomenon surrounding the use of Spanish words in Rapanui, the native language of Easter Island).

In cases where avoidance of the majority language is called for, nearly all loanwords can be kept out of HL discourse, except perhaps those that fill obvious lexical gaps (Hill & Hill, 1986, p. 387; Aikhenvald, 2002, p. 195). Purism thus may act as a brake on the borrowing process, making it harder to adopt a word that clearly belongs to the other language, no matter how useful it would be. One could see this as a valuable tool in the protection of the HL, but in reality, it often comes at a price that could be just as deadly as the overbearing dominance of the majority language. Once people care about the outcome of the maintenance versus shift struggle, language choice often becomes a weapon in the struggle between power and solidarity. In such situations, much of the codeswitching will constitute acts of identity-marking. The price is that this tends to set people up against each other.

There have not been many efforts yet to systematically compare communities in the degree to which this kind of struggle is salient in everyday communication (but see Bhatt & Bolonyai, 2011). Presumably, there is a continuum. On one extreme are HL communities in which people don't seem to care too much about which language people use in in-group conversation; at the other extreme there is fierce competition between the languages and considerable symbolism involved

in every language choice. In Hill & Hill's book on Mexicano, the struggle is in evidence on practically every page.

Indexicality is not limited to bilingualism: in all languages people associate elements with styles or registers (e.g. backwardness, elitism, high or low education, etcetera) and may exploit these associations for communicative effect by judicious use of such elements (imagine, for instance, the reactions if you use *ain't* in formal English discourse). However, since in a monolingual situation most lexemes, including almost all function words, will be shared between all registers, most words cannot be used to index any one register. In HL settings, on the other hand, almost all foreign-origin material can be regimented for this task.

### 11.7 Linguistic human rights and HLs

The final issue we would like to mention concerns linguistic human rights. Do HL speakers have the right, in a formal sense, to keep using and developing their language? Skuttnab-Kangas, Philipsson and Rannut (1994) have developed the notion of linguistic human rights, as a way of overcoming linguistic discrimination. The notion of linguistic human rights already has a long history.

The 1948 Universal Declaration of Human Rights, adopted by the United Nations, has language as one of its categories for equal rights, but it does not explicitly list and elaborate on linguistic rights. It does mention language in several of its articles, however: Article 2 on no discrimination based on language; Article 10 on the right to a fair trial, involving the right to an interpreter; Article 19 on the right to freedom of expression, including the right to choose any language for this; Article 26 on the right to be educated, with reference to the language of instruction.

The Universal Declaration of Linguistic Rights was approved on June 9, 1996 in Barcelona, Spain, by the World Conference on Linguistic Rights, but it has not been formally endorsed by UNESCO, possibly because there were too many uncertainties involved. On 5 March 2012, a condensed and updated version of the declaration was presented as the Girona Manifesto.

Regarding the human rights of HL speakers, sometimes these rights become part of a political struggle, as in the case of the speakers of Hungarian in Transylvania, part of Romania, particularly during the regime of Nicolae Ceausescu. Nationalism is then at the root of an attempt to forbid the use of a minority HL. In other cases, alleged violations of human linguistic rights are invoked to act against a neighboring country, as with the annexation of the Crimea by Russia in March 2014.

There is considerable discussion about the role of schools in supporting linguistic rights. They clearly have a role. Skuttnab-Kangas (2002, p. 46) notes:

“Even if schools cannot save languages, ..., schools can kill them more or less on their own.” Nonetheless, Hornberger (1988, p. 237), who studied bilingual education in a Quechua speaking community not so far from Irohito (where Uchumataqu was spoken) in the province of Puno (Peru), points to the need to embed projects of language revitalization in an overall process of social and political change. She stresses:

..., what is needed for effective maintenance planning and effective use of schools as agents for language maintenance is: autonomy of the speech community in deciding about use of languages in their schools and a societal context in which primary incentives exist for the use of one, two, or multiple languages in that and every other domain.

Thus, enforcing linguistic human rights is quite complicated. Community support varies considerably, as was also noted in Chapter 3.

## 11.8 Conclusion and overview

This chapter concludes our book on HLs. In this last chapter we have zoomed out to the political dimension of HLs, and discussed endangered languages, language rights, and the preservation of diversity. These languages can be studied, and have been, for many different reasons. They offer insights into changing patterns of language use in migrant communities, often influenced by a dominant language. They also can tell us something about language acquisition under different circumstances, and about what goes on in bilingual processing.

The study of heritage languages is part of a large scholarly and cultural effort that attempts to come to grips with our cultural heritage, but from a dynamic rather than a static perspective. HLs are languages that have undergone multiple changes due to the settings in which they are acquired and used.

What are HLs? For individual researchers different dimensions are central to the definition of HLs (Chapter 1). These include (a) not having an official status; (b) speakers having undergone a shift in language dominance; (c) divergent grammars; (d) personal and ethnic or ancestral ties; (e) acquisition at a later age; and (f) status as a community language. In this book we have tried to present various, sometimes very different, perspectives on the complex phenomenon of heritage languages (Chapter 2). There are several older studies that adopt a *diaspora* perspective: the spread of a language across the world, and others adopt the *immigration* perspective. However, in much current research the emphasis is on what individual speakers do, and how and when they acquire the HL, the *speaker* perspective.



HLs should be seen against the background of broader processes of language *maintenance* and *shift* in bilingual communities (Chapter 3). The social environment in which an HL is spoken determines much of what happens to it. We can and should situate HLs in the sociolinguistic framework of *bilingual language use* and *codeswitching* (Chapter 4). People use their HLs and the dominant language both together and apart in everyday life, which leads to interesting cases of language mixing.

A number of different *methods* have been used to gather HL materials, mostly recordings of spontaneous natural communication, questionnaires, and controlled elicitation techniques (Chapter 5). These methods should be internally, externally, and ecologically as valid as possible. Increasingly, digital techniques play a role in gathering and analyzing the data.

An important issue in HL studies is their *variability* (Chapter 6). Many different factors have been found to explain changes in an HL, but an important preliminary question is how to establish the baseline. There are a number of good possibilities. Many tests have been developed to assess linguistic profiles including speakers' proficiency, language attitudes and patterns of language use.

A key question of course is what changes exactly HLs have undergone. What are *typical HL phenomena*, and how do we account for them (Chapter 7)? We argue that both internal and external factors play a role in shaping HLs. Many recent studies focus on acquisition and processing, but characteristics of the other languages spoken by the multilingual speaker are also essential. These may influence the HL through transfer, although this influence is not always easy to establish and sometimes indirect. When there is relatively high maintenance of the community language, and a period of active bilingual usage within the community spanning sometimes several generations, contact-induced language change is often found.

Researchers have adopted different *grammatical frameworks* to study HLs, and the choice of framework implies theoretical choices by the researcher (Chapter 8). These frameworks include generative grammar, variationist linguistics, optimality theory, and usage-based analyses.

It is clear that the nature of *multilingual processing* is central to our understanding of how HL speakers function (Chapter 9), and HL speakers can be seen as a specific type of multilingual. It is nearly impossible for a bilingual to completely switch off one of his/her languages, and there are important effects due to cross-linguistic priming.

Given that intensive and sustained bilingual usage, involving considerable *time depth* and sometimes several generations, is needed for contact-induced language change it is useful to consider HLs which do not only have a migration history, but also an earlier colonial history in common with a dominant language

(Chapter 10). Papiamentu is such a language, and it has undergone many changes, in part due to the dominant language, Dutch.

Heritage language studies are central to our concerns with language in modern societies. They form an important part of the identity of their speakers and need to be viewed from multiple perspectives. This has been our emphasis throughout this book: both acquisition and use are important. We also stress, and therefore our book has as its sub-title 'a language contact perspective', that HLs always involve multilingual language use. Finally, and perhaps most importantly, they testify to the cultural resilience of bilingual communities across the world, communities that often must adapt to the language and culture of the dominant society, but at the same time want to preserve something of their heritage. HLs and heritage cultures are increasingly seen as feeding into and enriching the countries where they are found.



## Technical terms used in this book related to heritage languages

We have tried to define the technical terms used in the literature we discuss in this book as they came along. However, here they are again in alphabetical order. Please bear in mind that sometimes one concept has various interpretations, as they express the points of view of numerous different researchers.

*Access versus representation:* Distinction made by some researchers regarding the question whether a person ‘knows’ a certain pattern or not; in some theoretical accounts one can know a pattern, but experience trouble ‘accessing’ it in language processing.

*Acquisition in a naturalistic setting:* Learning a language through interaction rather than through explicit instruction or self-study.

*Acquisition:* Cover term for any way of learning a language, either consciously or unconsciously.

*Acrolect:* Variety of a language that has high prestige, often used to denote the highest-prestige register in a creole-speaking community.

*Activation threshold hypothesis:* If you use a word or pattern frequently enough you have immediate access to it, but in HLLs sometimes infrequent use leads to the phenomenon that a certain word or pattern is below the activation threshold.

*Additive bilingualism:* Type of bilingualism in which learning a second language does not lead to loss of or less competence in the first language.

*Additive borrowing:* Borrowing of patterns, sounds or words which do not replace already existing elements in the language, but rather are added to them.

*Adult L2 learners:* People learning a second language as adults, often assumed to involve greater difficulty than learning while younger.

*Afro-Iberian:* Cover term for varieties of Portuguese and Spanish that emerged in contact with Africans.

*Age of arrival:* Age at which a person or group of persons arrived in a new speech community, through migration.

*Age of onset:* Age at which a person is first exposed to a new language.

*Age-effects:* Effects on speed and mode of acquiring or processing a language due to the age of a person.

*Age grading*: Differences in language use within a community associated with particular age groups. Shifts in style occur in many speakers of a language community during their lifetime; often when people use age-appropriate language, such as teenage slang or standard language when adults enter the job market.

*Agent phrase*: The *by*-phrase in a passive construction.

*Alternational codeswitching*: Form of codeswitching in which independent chunks from different languages (often clauses or larger phrases) succeed each other.

*Amazon Mechanical Turk*: A website powered by Amazon.com where people can perform tasks, such as participating in an experiment, often for a small payment.

*Analytic*: In an analytic expression separate meanings are expressed by separate words or elements; compare English 'I ate' with Papiamentu *mi a kome* '1sg PAST eat'. In the latter, the past tense is expressed analytically, while in English it is expressed synthetically (see Synthetic).

*Anaphoric use*: Use of a referent pointing back to something said before: 'Did you see John? / Yeah, *he* was in the supermarket ahead of me at the cashier.'

*Ancestral ties*: Ties to a language, place or cultural phenomenon through someone's ancestors, rather than through daily practice.

*Anxiety and insecurity*: Negative feelings that a person may have because of (supposed) lack of command of a specific language, often the case with heritage speakers.

*Attrition*: Process by which a speaker may lose access to a language or specific parts of that language when s/he no longer uses the language daily.

*Basic vocabulary*: The part of the vocabulary of a language that describes basic concepts used frequently. Often the term is used to describe the 100 or 200 word-list proposed by Morris Swadesh.

*Basilect*: Variety of a language that has low prestige, e.g. the everyday register in a creole-speaking community (compare 'acrolect').

*Bilingualism*: Using more than one language in everyday situations.

*Borrowing*: Making a word or pattern from another language part of your own language.

*Brokering*: A situation in which some members of a community (e.g. an immigrant community) act as intermediaries in communication with outsiders. Sometimes we see children functioning as 'brokers'.

*Calque*: Copying an expression from a different language but using native words, like Spanish *rasca cielos* 'scrape skies' for 'sky scraper'.

*Chain migration*: Migration pattern in which one group of immigrants acts as the host for a second group from the same country of origin, and so on.

- Child first language loss*: The phenomenon that the first language a child learns may actually be lost because the child does not use that language any more as an older child or adult.
- Circular migration*: Migration pattern in which people move back and forth between two countries.
- Codeswitching*: The use of more than one language in a single speech event (sentence, turn, pair of turns).
- Cognitive Linguistics*: School in linguistics in which form/meaning units are the building block of grammar and in which language acquisition, structure, variation, and change are explained through the operation of general cognitive principles.
- Colonial language*: Language associated with the colonizers of a territory.
- Community language school*: School program specially geared towards teaching the community language of a minority group (immigrant or indigenous).
- Community language*: The language which members of a specific minority group use among themselves.
- Community norms*: Norms or conventions for language use that seem to hold in a community. It is the question whether there are actual norms within a bilingual community for the use of the HL.
- Competence*: Somewhat abstract knowledge of the rules and patterns of our grammar, separate from use (though it informs use).
- Complementary school*: see community language school.
- Comprehension*: Processing incoming signals in communication and transforming them into meanings.
- Construction Grammar*: Grammar model which focuses on grammatical constructions as the basic building blocks, constructions which can be combined into larger units.
- Constitutional national language*: A national language specifically designated as such in the constitution.
- Contact induced grammaticalization*: Grammaticalization resulting from contact with another language, specifically from calquing.
- Content vocabulary*: The words in a language that refer to specific (often concrete) concepts.
- Contextual embedding*: Embedding of grammatical knowledge in semantic or pragmatic contexts; this may be demanding, especially in some experimental tasks used in psycholinguistic approaches.
- Conventionalization*: In usage-based models the process by which a linguistic unit becomes part of the conventional stock of units of a language.
- Convergence*: Process of language change through which two languages become more like each other.

- Creole*: A language of which we can establish the time of its emergence precisely, as opposed to languages which very gradually split off from other languages.
- Critical period for acquisition*: Assuming that the human capacity for acquiring a language undergoes rapid developmental changes as a child is growing up; some researchers assume that there is a critical period during which it is easiest to acquire a language.
- Critical period for attrition or language loss hypothesis*: The idea here is that the critical period assumed for language development also holds for language loss. The earlier the attrition process starts, the more severe it can be.
- Cross linguistic influence*: As bilinguals use several languages, one of their languages may influence the other one during language processing.
- Cross-language activation*: Since both languages are potentially accessible during language processing, properties of one language may be active and interfere while the other one is being used.
- Cultural and ancestral association*: The assumption that HLLs are culturally linked to an ancestral community and associated with a specific identity.
- Cultural fluency*: Term used to describe thorough knowledge of and capability of participating in a culture.
- Cyclic migration*: Migration pattern in which immigrants periodically return to their country of origin and in which new migrants may enter the immigrant community.
- Declarative knowledge*: see 'procedural knowledge'
- Dialect(al) variation*: Speakers generally master several variants or dialects of the same language, and the variation between these may play a role in the formation of the heritage variety.
- Diaspora*: The spread of a language to different parts of the globe, often leading to splitting up into different language communities.
- Differential Object Marking*: Grammatical pattern in which objects are optionally marked with a special preposition or suffix, depending on different syntactic, semantic, and pragmatic conditions.
- Diminutive suffix*: A particle on the noun to mark that the referent is little or is used in an affectionate way.
- Discourse marker*: A word, particle or clitic added to an utterance to mark a specific discourse effect such as contrast or mitigation.
- Ditransitive construction*: Construction in which a verb such as 'give' or 'tell' has both a direct (Theme) and an indirect (Goal) object.
- Divergence*: Change in a language away from earlier related forms and meanings.
- Diversity management*: The policies of a state, institution, or organization to handle the different language varieties of its members.
- Documentation of heritage varieties*: see language documentation.

- Dominance shift*: A change over an individual's life time affecting which language is the dominant one.
- Dual immersion* (also 'two-way immersion', 'dual language', 'dual language immersion', and 'dual enrolment'): This generally refers to an educational program in which pupils with a native background in the dominant language of the society receive part of their courses in a non-dominant language, e.g. English-speaking pupils take part of their courses in Spanish in California.
- Ecological validity in data collection*: Some guarantee that the type of language data being collected is exactly like naturally occurring language use.
- Entrenchment*: The degree to which a cognitive unit is routinized and automated in the human cognitive system.
- Establishing the baseline*: In HL research, determining with exactitude the variety from which the HL developed.
- Ethical concerns in experimental research*: Making sure, as you are running an experiment, that ethical concerns such as the rights and the privacy of the subjects studied are guaranteed.
- Ethnic affiliation*: The ethnic group that a speaker belongs to.
- Ethnicity*: Properties associated with belonging to a ethnic group.
- Ethnolects broad*: The set of languages or language varieties used by a specific ethnic group, within the context of a larger social setting.
- Ethnolects narrow*: The variety of a dominant language spoken by a specific ethnic group or by several ethnic groups.
- Expats*: Immigrants (either temporary or more permanent) with considerable resources, either financially or in terms of educational background, and close ties to the country of origin.
- Explicitness hypothesis*: The explicitness hypothesis predicts that HL speakers will tend to select structures from the HL in their production which convey the intended meaning explicitly. The heritage speakers might lack confidence that the message will be understood properly, and therefore introduces more overt elements that are supposed to guide the hearer in the processing than monolinguals would do (cf. Polinsky, 2006).
- Explicit knowledge*: The conscious or semi-conscious knowledge of explicit grammatical rules and patterns.
- Exploratory study*: A broad and open research design without specific hypothesis aiming to get a general picture from an understudied domain. Conclusions based on exploratory studies should be taken with great caution.
- External validity*: Some guarantee that the results of the data collection or experiment can be generalized to a larger group of subjects or if possible to the whole population.



- Fieldwork*: Research in a real-life setting, either in the same region as where the researcher works, or in a foreign region or country.
- Foreign adoptees*: Children adopted but born in a different country (and potentially exposed to a different language originally).
- Formal and informal registers*: Ways of speaking or writing that differ in their degree of formality.
- Frequency distribution*: The quantitative distribution of a given item or feature in a larger data set in terms of the number of times it occurs.
- Frequency effects*: Effects on the production, comprehension, or use of a language item due to its frequency.
- Function words*: (Often short) words that have a grammatical function in the sentence, such as auxiliary or article.
- Functional categories*: Grammatical categories that correspond to function words, such as Tense, Person, Gender, Complementizer, Determiner.
- Functional differentiation*: when two or more languages each have a different function in the speech community.
- Fundamental difference hypothesis*: The hypothesis that learning a language as an L2 by adults is fundamentally different from the acquisition of an L1 by a child.
- Generations*: Particularly relevant in immigrant communities, different generations have different degrees of command of the community language and the dominant language.
- Generative models*: Models of language that are based on the ideas developed in generative linguistics, particularly by Noam Chomsky.
- Generic use*: Use of a referent in a general way, such as ‘*Donkeys can be stubborn.*’
- Girona Manifesto*: A manifesto drafted by PEN International in Girona, Spain, to support linguistic rights, particularly those of endangered languages. Signed in 2011, it builds on an earlier declaration drafted in Barcelona in 1996.
- Glossonym*: Name for a language.
- Glottophagie* ‘lit. language eating’: Name for the process through which the presence of a dominant language leads to the disappearance of smaller languages.
- Goal*: Semantic role of the argument towards which an action takes place, such as ‘go to Rome’.
- Grammaticality judgment tests*: Experiments in which speakers are directly asked whether a sentence or structure is grammatically well-formed or not.
- Grammaticalization*: Process by which a particular word or turn of phrase acquires a role in the grammar of a language, e.g. the verb ‘go’ being used as a future tense marker. Often grammaticalization is accompanied by phonological reduction and loss of concrete meaning.
- HALA-test* (relation between language strength and naming times): A test developed by William O’Grady to measure language dominance in terms

of length of naming time for objects. HALA stands for ‘Hawai‘i Assessment of Language Access’.

*Heritage language competence as a learning resource*: Educational program within which the HL is used as a resource in the mainstream classroom.

*Heritage language education*: see community language school.

*Heritage language re-learners*: People who want to learn the HL of their ancestral community as adults.

*Heritage language school*: see community language school.

*Heritage speaker broad*: Speakers who have cultural ties, through their ancestry, to a specific HL, whether they speak it or not.

*Heritage speaker narrow*: Speakers who are now dominant in a different language than the one learned during childhood.

*Higher clausal projections*: If we assume a tree-like structure for the utterance, the higher clausal projections, covering the general pragmatics of the utterance, are the top nodes involving topic, focus, speaker intent, etc.

*Home language education*: Any form of education involving specifically the language spoken at home.

*Home language*: The language spoken most at home.

*Iberian-based*: A Creole language with a lexicon derived from Portuguese and/or Spanish.

*Immigrant language*: The original language of a group of immigrants.

*Immigration perspective*: The perspective on HLs in which the country to which various groups migrate is central.

*Imperfect or incomplete acquisition*: Situation in which a language is not fully acquired, something which must be distinguished from cases where a language is being lost or forgotten.

*Implicit knowledge*: Knowledge of the rules and patterns of a language that manifests itself only in actual use, but of which the language user is not conscious.

*Implicit learning mechanisms*: Learning mechanisms which are not conscious and cannot be verbalized by the learner. The assumption is that heritage speakers may differ from second language learners in that their learning of the HL involved more implicit than explicit leaning, in contrast to L2 learners.

*Incompleteness hypothesis*: A heritage speaker cannot completely acquire the heritage language because of limited input.

*Indeterminacy*: The notion that a single sequence of words can have various structural interpretations, which may lead to grammatical change.

*Indigenous language*: A language traditionally spoken in a given area.

*Insertional codeswitching*: A type of codeswitching in which the utterance has a base or matrix language into which words or chunks from another language are inserted.

- Intake*: That part of the input that a learner manages to process when acquiring a language.
- Integration difficulty hypothesis*: Hypothesis that discourse related domains are more difficult than domains related to narrow syntax because of their integrative nature, since syntactic knowledge needs to be integrated with information from other cognitive domains, such as interaction.
- Interface hypothesis*: Hypothesis suggesting that cross-linguistic influence takes place particularly at those places where two parts of the grammar (e.g. syntax and pragmatics) interact.
- Inter-individual variability*: Variability between individual speakers.
- Internal migration and urbanization*: Social process through which populations leave the countryside and become part of the urban community.
- Internal validity*: Guarantee that the results or the data collected truly represent the properties or variables you were looking for, not something else.
- Intersection between L2 acquisition and L1 loss*: The hypothesis that there is a crossing point in the course of child development at which facility of loss of the first language and ease of acquisition of a second language intersect.
- Intra-individual variability*: Variability within the production and perception of a single speaker.
- Ius sanguinis*: Rights for a person (e.g. citizenship) acquired through being born into a certain lineage.
- Ius soli*: Rights for a person (e.g. citizenship) acquired through being in a certain place.
- Judgment*: Explicit and conscious response by a subject to properties of a stimulus item.
- Koinés*: In a context of language spread, a variety of the language emerges which has lost some of its variability and, possibly, its complexity.
- L1 acquisition*, see acquisition.
- L1 attrition*, see attrition.
- L1 restructuring*: The assumption that speakers may reanalyze parts of their L1 at a later age, presumably under the influence of a dominant second language.
- L1*: The language or languages a person acquires first in life.
- L2*: A language a person acquires later than the first one.
- Language attitudes*: Conscious or unconscious attitudes towards a language.
- Language choice*: The possibility, in a discourse, for a multilingual speaker to choose one or more of a set of different languages or language varieties, depending on the interlocutor, the setting, the topic of conversation, etc.
- Language death*: The process by which a language ceases to be spoken, because it is no longer transmitted to a new generation of speakers. HL speakers may

resemble speakers of a dying language, but of course in the case of many HLs, they are still spoken in the country of origin of the migrant group.

*Language documentation*: Recording language practices, often of a language which is under threat or undergoing rapid changes, like an HL. Documentation is the basis for language description.

*Language ideology*: The beliefs that speakers have about a language, its norms and contexts of use.

*Language loss*: see attrition.

*Language maintenance*: Situation in which a language of a non-dominant community continues being spoken, often in spite of pressure from a dominant language.

*Language policies regarding HLs*: Official policies concerning the use of HLs in education, the media, and health care, etc.

*Language revival*: New opportunities for use of a language which was under threat.

*Language rights*: The rights of use in education, the judicial system, etc. of a non-dominant language.

*Language shift*: Situation in which a language of a non-dominant community is being given up under pressure from a dominant language.

*Less is more hypothesis*: The hypothesis that children, with fewer cognitive resources at their disposal, are paradoxically better at learning a *language* since they by necessity have to start with small bits before building up larger structures.

*Lexical proficiency*: Being able to use and understand the words of a language. Often HL speakers lose lexical proficiency.

*Lexifier language*: In pidgin and creole genesis, the language that provides the lion's share of the vocabulary of the language.

*Lingua franca*: A language jointly used by members of different speech communities in a multilingual setting.

*Linguicism*: Discrimination of someone on the basis of the language spoken by that person.

*Linguistic area*: Part of the world or region where several unrelated or distantly related languages are spoken but where all varieties share certain characteristics that cannot have come from a common ancestor, but must have spread through borrowing.

*Linguistic diversity*: The number of languages and language varieties spoken in a given institution, region, or country.

*Linguistic genocide*: Conscious attempt to make a language community disappear.

*Linguistic human rights*: Human rights related to language rights.

*Literacy*: Degree to which a given individual, group, or speech community knows how to read and write.

*Loan translation*: see calque.

- Locative preposition*: Preposition used to mark the location of an entity.
- London Jamaican*: Variety of Jamaican Creole often mixed with English, spoken in London.
- Long time contacts*: Situation of prolonged bilingualism within a community, sometimes leading to more profound changes in one or both of the languages involved.
- Loss*: see attrition.
- Markedness*: Situation in which a language has features which are not common among languages and often complex and hard to learn or process. Such features often undergo reduction in situations of language contact.
- Matrix*: Basic underlying morpho-syntactic structure of a sentence.
- Measuring fluency*: Techniques for determining how fluent a speaker is in her different languages.
- Memory load*: Degree to which a given item or grammatical process makes high demands on working memory.
- Mesolect*: Variety of a language that has intermediate prestige, e.g. in a creole-speaking community (see ‘acrolect’ and ‘basilect’).
- Metatypy*. A subordinate language changes its typological properties under the influence of a dominant language, as an extreme case of convergence
- Migration patterns*: The ways in which a specific community organizes the migration of its members to a different place or country: cyclic, chain, seasonal, etc.
- Minimal sign/meaning pairings*: The minimal unit which has both a form and a meaning, such as the two elements in *caballo-s* ‘horse-PL’.
- Missing input competence divergence hypothesis*: The hypothesis that heritage learning should not be characterized as incomplete (see incompleteness hypothesis), but simply the natural result of not having access to specific input (for instance, because the child is only exposed to informal registers or to non-standard forms).
- Model of language processing*: Psycholinguistic model of how the mind processes language, both in terms of perception and of production.
- Modularity*: The idea that our human linguistic knowledge is not a single mass, but rather a set of sharply defined, relatively independent components.
- Morphological deficits*: Lack of competence in marking or comprehending particular morphological distinctions in an HL.
- Narrow syntax*: Syntactic rules and patterns in themselves, not directly interacting with pragmatics and semantics.
- Networks*: The social ties that influence a person’s language and other aspects of behavior. Some heritage speakers do not have a dense network of people with whom they use the HL.
- Normal diachronic change*: Changes that may happen in any language.

- Null (subject) pronouns:* The possibility in a language like Spanish to leave out a pronoun e.g. in subject position where a language like French or English does not allow it (Sp *camina*/Fr *il marche*/ Eng '(s)he walks').
- Official language:* The language of a country, region, or institution specified in laws or in the constitution as the one to be used in public.
- Onset of Bilingualism:* The moment at which a person becomes bilingual.
- Optimality Theory:* Grammatical model in which the outcomes are not so much specified as correct or incorrect but as the optimal one, violating the lowest number of constraints.
- Over-acceptance:* Tendency of a speaker to accept all or most sentences offered to her or him as correct.
- Overhearers:* People who only know bits and pieces of a language because they have heard it spoken, in childhood or by their neighbors.
- Over-regularization of patterns:* Tendency of a learner to use a common structure in cases where native speakers don't use it.
- Papiamentu:* The creole language that emerged on Aruba, Bonaire, and Curaçao in the Caribbean and is spoken there as well as by migrants from there, e.g. in the Netherlands.
- Parameter theory:* The hypothesis that the many differences between languages reflect a smaller number of deeper more fundamental structural differences, the so-called parameters, such as the null subject parameter (see Null subject pronouns).
- Particle verb:* Verb that is typically used in combination with a particle, often with a specialized meaning, such as 'give in' in English.
- Passive auxiliary:* Auxiliary verb used to mark a sentence as passive, such as English 'be' or German *werden*.
- Perceptual advantages:* The idea that early acquisition of a language at home produces advantages in perception (e.g. with respect to phonological contrasts).
- Performance:* the actual practice of using grammatical knowledge in actual production and comprehension.
- Phonetic production benefits:* The assumption that heritage speakers are better able to make phonetic contrasts than L2 learners because of early exposure.
- Phonological adaptation:* Process by which the pronunciation of a borrowed word is adapted to that of the 'host' language.
- Phonotactic pattern:* Pattern of allowed consonants and vowels in a syllable, morpheme, or word. Some languages, for instance, prefer 'open' syllables that end in a vowel.
- Picture-matching task (comprehension):* Experimental task in which a speaker subject is confronted with a choice matching pictures to options in descriptive labels or names offered.

- Picture-naming task (production)*: Experimental task in which a speaker subject is asked to provide a label or name for pictures.
- Politeness*: The strategies by which a speaker marks respect for the status of an interlocutor.
- Positive correlation between age and proficiency for L1*: Situation in which someone's age corresponds directly with her proficiency in the first language.
- Post-colonial heritage language*: HL spoken in a (former) colony of the country where heritage speakers have migrated to.
- Post-nominal adjective*: Adjective placed after the noun.
- Poverty of the stimulus*: The idea in generative linguistics that language learners create much richer representations of the language they are exposed to than what is purely present in the input or stimulus.
- Pragmatic salience*: The degree to which a specific word or construction has pragmatic force.
- Priming between languages*: Situation in which the use of a construction or word in one language influences the use of the construction or word in another language.
- Priming in language change*: Situation in which frequent use of a construction or word leads to a change in the language as such.
- Priming in second language acquisition*: Situation in which the earlier use of a construction or word affects the processes of second language acquisition.
- Procedural and declarative knowledge*: Declarative knowledge involves knowing facts, like the word for 'horse' in Spanish, and it is fairly conscious ('How do you say 'horse' in Spanish?' – 'Caballo').
- Procedural knowledge involves knowing how to do something, like pronounce <ll> in Spanish, or putting the words in a sentence in the right order.
- Pro-drop*: see Null (subject) pronouns
- Production advantages heritage accent (different from L2 accent)*: see phonetic production benefits
- Purism*: Tendency for individuals or groups of speakers to consider certain constructions, pronunciations, or words as 'foreign' and to be avoided in their own speech.
- Qualitative differences*: Significant structural differences between an HL and the baseline.
- Quantitative differences*: Differences between an HL and the baseline which do not reflect underlying grammatical differences but rather frequency of use of particular patterns or elements.
- Racialization*: Interpretation of a specific type of variation along lines of race.
- Reanalysis*: Process by which speakers assign a grammatical structure to a string different from how it used to be.

- Receptive bilinguals*: Speakers who can understand but not speak a second language.
- Reduction*: Decrease in meaning content or formal distinctions as the result of attrition, contact, or incomplete learning.
- Register variation*: Variation in language use along the lines of different registers, often defined in terms of degree of formality.
- Register*: Way of speaking or writing depending on implicit or explicit norms of (in)formality.
- Regression hypothesis (Jakobson)*: Hypothesis which predicts that the things that are learned first will be remembered best in a process of attrition or loss.
- Re-learning potential in adulthood*: The idea that heritage speakers who have lost their ancestral language are better able to re-learn it (particularly its phonology) than other L2 learners.
- Relexified/relexification*: Process through which the vocabulary of a language is partly or completely replaced by that of another language, while much of the original syntax, morphology, phonology, and semantics are retained.
- Rememberers*: People who do not actively speak a language, but remember parts of it, often from hearing their grandparents or other older relatives speak it.
- Replacive bilingualism*: Form of bilingualism in which knowledge of a second language has pushed out much of the knowledge of the first language. See also Subtractive bilingualism.
- Representational differences*: Differences between speakers not just in their output and use, but also in their underlying structural representations, i.e. their grammars.
- Restructuring*: Process by which the structures of the language a person speaks are altered in a situation of multilingualism.
- Retelling*: Experimental technique whereby a presented video or book story has to be retold by the subject.
- Returnees*: Migrants who have returned to their country of origin.
- Reversing language shift*: Effort to help speakers maintain their minority language or even increase its use.
- Revival program*: Concerted effort to give a language or language variety new uses and vitality.
- Role input*: The role that exposure to a first or subsequent language has on language development.
- Semantic bleaching*: Process by which a frequently used form loses some of its original literal meaning and becomes more general, grammatical or pragmatic.
- Semantic domain/field*: Circumscribed area of meanings, such as 'kinship,' or 'body parts'.



- Semantic extension*: Process by which a construction or word gains additional meanings.
- Semantic reinterpretation*: Process by which someone assigns a different meaning to a sentence or phrase than the one intended by the original speaker.
- Semi-speakers*: People who know a language only partly.
- Sequential bilingual*: Person who speaks two (or more) languages and has acquired them in sequence.
- Serial verb construction*: Construction (frequent in some West African languages, South-East Asian languages, and creole languages) in which several lexical verbs are combined in a single predicate as in *buy X give Y* 'buy X for Y'.
- Sign languages*: Languages employing signs made by the hands and upper body instead of spoken words, used in Deaf communities.
- Simplification*: Process by which a construction or linguistic form becomes less complex and loses some of its internal distinctions, often as the result of incomplete learning or attrition.
- Simultaneous bilingual*: Person who speaks two (or more) languages and has acquired them at the same time.
- Sociolectal variation*: Situation in which there are linguistic differences within a speech community due to social class distinctions.
- Sociolinguistic pressure*: Pressure upon a speaker to use her language in a specific way because of norms imposed by the community.
- Speaker profiles*: Structured characteristics of (groups of) speakers, which determine their language processing and multilingual proficiency.
- Speech rate*: Rate at which a speaker produces (spontaneous) utterances, often measured in terms of words per minute.
- Stagnated L1 development*: Situation in which further development of the L1 is hindered because of lack of further input or situations in which the language can be used.
- Stochastic variation*: Variation which contains an element of randomness.
- Structure preserving change*: A change in the HL which affects frequency of use but not the fundamental structure of the language.
- Style shift*: Within a conversation, changing to a different variety of the same language, associated with for example register, age group or gender
- Submersion*: Educational program in which speakers are fully taught in a second language, generally the dominant language, without recourse to their home language.
- Subtractive bilingualism*: Type of bilingualism in which the acquisition of a second language leads to diminished proficiency in the first language.
- Superdiversity*: The phenomenon that contemporary societies, particularly those in the industrialized world, are not simply characterized by the presence of

different stable immigrant or non-immigrant minority groups (linguistic and cultural diversity), but by rapidly changing and complex patterns of internal and external migration of groups which are themselves culturally heterogeneous and multilingual.

*Synthetic*: In a synthetic expression separate meanings are combined in a single word or phrase, as in English ‘be’ in ‘So *be* it,’ which is the [third person] [singular] [subjunctive] [present] of the verb ‘be,’ or ‘were’ in ‘if it *were* ...’ would be [third person] [singular] [subjunctive] [past]. Both forms combine various meanings, and are thus highly synthetic. See also Analytic.

*Theme*: Semantic role of an argument directly affected by an action, typically a direct object, as in ‘see a house.’

*Topic*: The sentence topic is what an utterance or statement is about.

*Transfer*: The use of rules or patterns of one language when producing or processing another language.

*Transnational norms*: Norms for a language that come from across a national border; the norms for Turkish in Germany come for a large part from Turkish as spoken in Turkey, for instance.

*Transparency*: When the meaning of a lexical combination, a structure or a pattern is immediately recognizable or interpretable from the way its parts combine.

*Triangulation*: An analysis based on different research methodologies such as conversational speech and experimental data to test if evidence from these methods converges or not

*Truth value judgment tasks*: Experimental tasks in which subjects are asked to rate the truth value of statements in a given experimental task (e.g. involving pictures).

*Ultimate attainment*: The final level of competence that a speaker reaches in, in our case, the HL.

*UNESCO*: The United Nations Educational, Scientific and Cultural Organization was founded in 1945 to promote international collaboration in the fields of education, science, and culture. Its headquarters are in Paris.

*Unitary Linguistic System Hypothesis*: The hypothesis that some bilingual speakers of an HL have a single grammatical system for their different languages.

*Universal constraints*: Constraints on possible linguistics outputs, in Optimality Theory, which are in themselves universal, but of which the effect varies from language to language because of different rankings.

*Universal tendencies*: General tendencies in language processing and language change that are due to universal properties of the language system and of the way humans process this system.

*Unstressed*: Unit (vowel, syllable) not receiving stress in pronunciation.

*Usage-based models*: Models for the study of language in which patterns of use and how these patterns influence language structure are central.

*Variable rules*: Grammatical rules that may or may not apply, depending on independent variables such as speaker characteristics or the structural contact, such as position in the word or sentence.

*Variance among bilinguals*: The fact that not all bilinguals have the same type and level of proficiency in their different languages, may have different attitudes towards their languages, and combine them in different ways.

*Variationist approach*: Approach to language studies in which the variability of language and the possibility of systematically studying the constraints governing this variability are central.

*Vitality*: The chances that a language will continue to be spoken in the future.

*Voice Onset Time (VOT)*: A term from phonetics, defined for plosives as the length of time between the release of a stop consonant and the onset of voicing or vibration of the vocal chords. Changes in VOT can be an indicator of changes in the accents of heritage speakers.

*Vulnerable domains*: Parts, components or modules of the grammar of a language which are particularly susceptible to influence from another language.

*Weaker Language as L1 hypothesis*: The hypothesis that for heritage speakers their L1 is their weaker language.

*Weaker language as L2 hypothesis*: The hypothesis that if a second language is acquired later in life, it remains the weaker language; this in contrast with heritage speakers, where the L1 may be the weaker language.

*Words per minute*: Measurement of fluency in terms of how many words per minute a speaker produces.

*World Englishes*: Varieties of English, whether spoken as L1 or L2, distributed around the world.

# References

- Aalberse, S., & Moro, F. (2014). Stability in Chinese and Malay heritage languages as a source of divergence. In K. Braunmüller, S. Höder, & K. Köhl (Eds.), *Stability and divergence in language contact: Factors and mechanisms* (pp. 141–162). Amsterdam: John Benjamins.
- Aalberse, S., Zhou, Y., & Andringa, S. J. (2017). Extended use of demonstrative pronouns in two generations of Mandarin Chinese speakers in the Netherlands: Evidence of convergence? In E. Blom, L. Cornips, & J. Schaeffer (Eds.), *Cross-linguistic influence in bilingualism: In honor of Aafke Hulk* (pp. 25–48). Amsterdam: John Benjamins.  
<https://doi.org/10.1075/sibil.52.03aal>
- Aalberse, S., & Hulk, A. (2018). Introduction to the special issue – Heritage language studies and early child bilingualism research: Understanding the connection. *International Journal of Bilingualism*, 22(5), 491–496.
- Aalberse, S., Andringa, S., Faber, M., & Lippe, P. (Forthcoming). Definiteness in Wenzhounese Chinese in the Netherlands and in China: Evidence for generational change in two locations. In B. Brehmer, J. Treffers-Daller, & D. Berndt (Eds.), *Lost in transmission: The role of attrition and input in heritage language development*. Amsterdam: John Benjamins.
- Abraham, W. (2011). Spoken syntax in Cimbrian of the linguistic islands in Northern Italy – and what they (do not) betray about language universals and change under areal contact with Italo-Romance. In M. Putnam (Ed.), *Studies on German-language islands* (pp. 233–278). Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.123.10abr>
- Abrahamsson, N., & Hylténstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59, 249–306.  
<https://doi.org/10.1111/j.1467-9922.2009.00507.x>
- Adelaar, K. A., & Prentice, D. J. (1996). Malay: Its history, role and spread. *Atlas of languages of intercultural communication in the Pacific, Asia and the Americas*, 2, 1. Berlin: De Gruyter.  
<https://doi.org/10.1515/9783110819724.2.673>
- Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science*, 17, 814–823. <https://doi.org/10.1111/j.1467-9280.2006.01787.x>
- Adjemian, C. (1983). The transferability of lexical properties. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 250–268). Rowley, MA: Newbury House.
- Aghinotri, R. K. (1987). *Crisis of identity. The Sikhs in England*. New Delhi: Bahri Publications, and Edmonton, CN: Linguistic Research.
- Ahn, S., Chang, C. B., DeKeyser, R., & Lee-Ellis, S. (2017). Age effects in first language attrition: Speech perception by Korean-English bilinguals. *Language Learning*, 67, 694–733.  
<https://doi.org/10.1111/lang.12252>
- Aikhenvald, A. Y. (2002). *Language contact in Amazonia*. Oxford: Oxford University Press.
- Albirini, A., Benmamoun, E., & Chakrani, B. (2013). Gender and number agreement in the oral production of Arabic heritage speakers. *Bilingualism: Language and Cognition*, 16(1), 1–18.  
<https://doi.org/10.1017/S1366728912000132>

- Allen, B., & Salmons, J. C. (2015). Heritage language obstruent phonetics and phonology. American Norwegian and Norwegian-American English. In J. Bondi Johannessen & J. C. Salmons (Eds.), *Germanic heritage languages in North America. Acquisition, attrition and change* (pp. 97–116). Amsterdam: John Benjamins. <https://doi.org/10.1075/silv.18.04.all>
- Alvanoudi, A. (2018). Language contact, borrowing and code switching: A case study of Australian Greek. *Journal of Greek Linguistics*, 18, 3–44. <https://doi.org/10.1163/15699846-01701001>
- Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, 42, 239–273. <https://doi.org/10.1017/S030500091400049X>
- Andersen, R. W. (1974). Nativization and hispanization in the Papiamentu of Curacao: A sociolinguistic study of variation (Unpublished doctoral dissertation). University of Texas at Austin.
- Andersen, R. W. (1989). La adquisición de la morfología verbal. *Lingüística*, 1, 90–142.
- Andersen, R. W. (1991). Development sequences: The emergence of aspect marking in second language acquisition. In T. Huebner & C. Ferguson, (Eds.), *Crosscurrents in second language acquisition and linguistic theory* (pp. 305–323). Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.2.17and>
- Androutsopoulos, J. (2013). Code-switching in computer-mediated communication. In S. C. Herring, D. Stein, & T. Virtanen (Eds.), *Pragmatics of computer-mediated communication* (pp. 667–694). Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110214468.667>
- Anttila, A. (1997). Deriving variation from grammar. In F. Hinskens, R. van Hout, & L. Wetzels (Eds.), *Variation, change and phonological theory* (pp. 35–68). Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.146.04ant>
- Appel, R., & Muysken, P. (1987). *Language contact and bilingualism*. London: Edward Arnold. [2005 University of Amsterdam Press].
- Au, T. K. F., Knightly, L. M., Jun, S. A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science*, 13(3), 238–243. <https://doi.org/10.1111/1467-9280.00444>
- Auer, P. (1999). From code-switching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism*, 3(4), 309–332. <https://doi.org/10.1177/13670069990030040101>
- Ayduk, O., Mendoza-Denton, R., Mischel, W., Downey, G., Peake, P. K., & Rodriguez, M. (2000). Regulating the interpersonal self: Strategic self-regulation for coping with rejection sensitivity. *Journal of Personality and Social Psychology*, 79(5), 776–792. <https://doi.org/10.1037/0022-3514.79.5.776>
- Backus, A. M., & Türker-Van der Heiden, E. (2009). Synchronic use of Turkish *yapmak* and the diachronic changes it leads to. Paper presented at 12th International Conference on Minority Languages, Tartu, Estonia, 28–30 May.
- Backus, A. M., & Dorleijn, M. (2008). Loan translations versus code-switching. In B. E. Bullock, & A. J. Toribio (Eds.), *The Cambridge handbook of linguistic code-switching* (pp. 75–93). Cambridge: Cambridge University Press.
- Backus, A. M. (1992). *Patterns of language mixing. A study in Turkish-Dutch bilingualism*. Wiesbaden: Otto Harrassowitz.
- Backus, A. M. (1996). *Two in one. Bilingual speech of Turkish immigrants in the Netherlands*. Tilburg: Tilburg University Press.
- Backus, A. M. (1998). Turks-Nederlandse codewisseling. Universele en taalspecifieke aspecten van taalcontact. *Nederlandse Taalkunde*, 3, 2–17.

- Backus, A. M. (2005). Codeswitching and language change: One thing leads to another? *International Journal of Bilingualism*, 9(3/4), 307–340.  
<https://doi.org/10.1177/13670069050090030101>
- Backus, A. M. (2013). Turkish as an immigrant language in Europe. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism and multilingualism* (2nd ed., pp. 770–790). Oxford: Blackwell.
- Backus, A. M., Dođruöz, A. S., & Heine, B. (2011). Salient stages in contact-induced grammatical change: Evidence from synchronic vs. diachronic contact situations. *Language Sciences*, 33, 738–752. <https://doi.org/10.1016/j.langsci.2011.02.004>
- Bard, E. Gurman, Robertson, D., & Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language*, 72(1), 32–68. <https://doi.org/10.2307/416793>
- Barz, R. K., & Siegel, J. (1988). *Language transplanted: The development of overseas Hindi*. Wiesbaden: Otto Harrassowitz.
- Bayram, F., Rothman, J., Iverson, M., Miller, D., Puig Mayenco, E., Kupisch, T., & Westergaard, M. (2017). Differences in use without deficiencies in competence: Passives in the Turkish and German of Turkish heritage speakers in Germany. *International Journal of Bilingual Education and Bilingualism*, 1–21. <https://doi.org/10.1080/13670050.2017.1324403>
- Bayram, F., Kupisch, T., Pascual y Cabo, D. & Rothman, J. (In press). Terminology matters on theoretical grounds too!: Coherent grammars cannot be incomplete. *Studies in Second Language Acquisition*.
- Benmamoun, E., Montrul, S., & Polinsky, M. (2013a). Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics*, 39(3–4), 129–181.
- Benmamoun, E., Montrul, S., & Polinsky, M. (2013b). Defining an “ideal” heritage speaker: Theoretical and methodological challenges Reply to peer commentaries. *Theoretical Linguistics*, 39(3–4), 259–294.
- Benson, M. (1960). American-Russian speech. *American Speech*, 35, 163–74.  
<https://doi.org/10.2307/453882>
- Bentahila, A., & Davies, E. E. (1995). Patterns of code-switching and patterns of language contact. *Lingua*, 96(2), 75–93. [https://doi.org/10.1016/0024-3841\(94\)00035-K](https://doi.org/10.1016/0024-3841(94)00035-K)
- Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2012). Effects of phonological feedback on the selection of syntax: Evidence from between-language syntactic priming. *Bilingualism: Language and Cognition*, 15, 503–516. <https://doi.org/10.1017/S1366728911000162>
- Bhatia, T. K., & Ritchie, W. C. (Eds.). (2008). *The handbook of bilingualism*. Oxford: Wiley-Blackwell.
- Bhatt, R., & Bolonyai, A. (2011). Code-switching and the optimal grammar of bilingual language use. *Bilingualism: Language and Cognition* 14(4), 1–25.  
<https://doi.org/10.1017/S1366728910000295>
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition* 12(1), 3–11. <https://doi.org/10.1017/S1366728908003477>
- Birdsong, D., Gertken, L., & Amengual, M. (Eds.) (2012). *Bilingual language profile: An easy-to-use instrument to assess bilingualism*. Austin, TX: COERLL, University of Texas at Austin. <<https://sites.la.utexas.edu/bilingual/>>
- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second language acquisition. *Journal of Memory and Language*, 44, 235–249.  
<https://doi.org/10.1006/jmla.2000.2750>
- Birmingham, J. C. (1970). The Papiamentu language of Curaçao (Unpublished doctoral dissertation). University of Virginia, Charlottesville.

- Blom, J.-P., & Gumperz, J. J. (1972). Social meaning in linguistic structures: Code switching in Northern Norway. In J. J. Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics* (pp. 407–434). New York, NY: Holt, Rinehart, & Winston.
- Blom, W. B. T., Polišíenská, D., & Weerman, F. (2007). Effects of age on the acquisition of agreement inflection. *Morphology*, 16(2), 313–336. <https://doi.org/10.1007/s11525-007-9110-1>
- Blutner, R., Hoop, H. de, & Hendriks, P. (2006). *Optimal communication*. Stanford, CA: CSLI.
- Boas, H. C. (2009). *The life and death of Texas German* (Publication of the American Dialect Society). Durham, NC: Duke University Press.
- Bobb, S. C., & Wodniecka, Z. (2013). Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. *Journal of Cognitive Psychology*, 25(5), 568–585. <https://doi.org/10.1080/20445911.2013.792822>
- Bock, J. K. (1995). Producing agreement. *Current Directions in Psychological Science*, 8, 56–61. <https://doi.org/10.1111/1467-8721.ep10771165>
- Boersma, P., & Weenink, D. (2002). *Praat* [Computer program]. <<http://www.fon.hum.uva.nl/praat/>>
- Boersma, P., & Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, 32, 45–86. <https://doi.org/10.1162/00243890154586>
- Bohman, T. M., Bedore, L. M., Peña, E. D., Mendez-Perez, A., & Gillam, R. B. (2010). What you hear and what you say: Language performance in Spanish–English bilinguals. *International Journal of Bilingual Education and Bilingualism*, 13(3), 325–344. <https://doi.org/10.1080/13670050903342019>
- Bondi Johannessen, J. (2018). Factors of variation, maintenance and change in Scandinavian heritage languages. *International Journal of Bilingualism*, 22(4), 447–465.
- Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19, 447–465. <https://doi.org/10.1017/S0272263197004026>
- Bos, C. (1977). Prestaties en problemen van Molukse leerlingen in het Mavo-onderwijs. *Mens en Maatschappij*, 52(3), 324–341.
- Bosker, H. R., Quené, H., Sanders, T., & de Jong, N. H. (2014). The perception of fluency in native and non-native speech. *Language Learning*, 64, 579–614. <https://doi.org/10.1111/lang.12067>
- Bot, K. de. (1992). A bilingual production model: Levelt's Speaking model adapted. *Applied Linguistics*, 13(1), 1–24. <https://doi.org/10.1093/applin/13.1.1>
- Bot, K. de. (2004). The multilingual lexicon: Modeling selection and control. *The International Journal of Multilingualism*, 1, 17–32. <https://doi.org/10.1080/14790710408668176>
- Bot, K. de, Broersma, M., & Isurin, L. (2009). Sources of triggering in code switching. In L. Isurin, D. Winford, & K. de Bot (Eds.), *Multidisciplinary approaches to code switching* (pp. 85–102). Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.41.07.bot>
- Bot, K. de, Lowie, W., & Verspoor, M. (2005). *Second language acquisition: An advanced resource book*. London: Routledge. <https://doi.org/10.4324/9780203446416>
- Bousquette, J., Putnam, M. T., Salmons, J., Frey, B., & Nützel, D. (2016). Multiple grammars, dominance, and optimization. In G. Legendre, M. T. Putnam, H. de Swart, & E. Zaroukian (Eds.), *Optimality theoretic syntax, semantics and pragmatics: From uni- to bidirectional optimization* (pp. 158–176). Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198757115.003.0007>

- Brecht, R. D., & Walton, A. R. (1994). National strategic planning in the less commonly taught languages [includes Asian languages]. In R. D. Lambert (Ed.), *Foreign language policy: An agenda for change* (pp. 190–212). Thousand Oaks, CA.: Sage.
- Brenneker, P. (1961). *Curaçaoensia. Folkloristische aantekeningen over Curaçao*. Curaçao: Boekhandel St. Augustinus.
- Bresnan, J. (2000). Optimal syntax. In J. Dekkers, F. van der Leeuw, & J. van de Weijer (Eds.) *Optimality theory: Phonology, syntax and acquisition* (pp. 334–385). Oxford: Oxford University Press.
- Bridges, K., & Hoff, E. (2014). Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Applied Psycholinguistics*, 35(2), 225–241. <https://doi.org/10.1017/S0142716412000379>
- Brown, C. M., & Hagoort, P. (1999). *The neurocognition of language*. Oxford: Oxford University Press.
- Brown, J., & Putnam, M. (2015). Functional convergence and extension in contact: Syntactic and semantic attributes of the progressive aspect in Pennsylvania Dutch. In J. C. Salmons & J. B. Johannessen (eds.), *Germanic heritage languages in North America: Acquisition and change* (pp. 135–160). Amsterdam: John Benjamins. <https://doi.org/10.1075/silv.18.06bro>
- Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.
- Bultena, S. (2013). Bilingual processing of cognates and language switches in sentence context (Unpublished doctoral dissertation). Radboud University, Nijmegen.
- Bylund, E., Abrahamsson, N., & Hylstenstam, K. (2010). The role of language aptitude in first language attrition: The case of prepubescent attriters. *Applied Linguistics*, 31, 443–464. <https://doi.org/10.1093/applin/amp059>
- Bylund, E., & Díaz, M. (2012). The effects of heritage language instruction on first language proficiency. A psycholinguistic perspective. *International Journal of Bilingual Education and Bilingualism*, 5, 593–609. <https://doi.org/10.1080/13670050.2012.676620>
- Bylund, E., & Jarvis, S. (2011). L2 effects on L1 event conceptualization. *Bilingualism: Language and Cognition*, 14, 47–59. <https://doi.org/10.1017/S1366728910000180>
- Bylund, E., & Ramírez-Galán, P. (2014). Language aptitude in first language attrition: A study on late Spanish–Swedish bilinguals. *Applied Linguistics*, 37(5), 621–638. <https://doi.org/10.1093/applin/amu055>
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511750526>
- Calvet, J. L. (1974). *Linguistique et colonialisme, petit traité de glottophagie*. Paris: Payot.
- Canagarajah, A. S. (2008). Language shift and the family: Questions from the Sri Lankan Tamil diaspora. *Journal of Sociolinguistics*, 12, 1–34. <https://doi.org/10.1111/j.1467-9841.2008.00361.x>
- Carlock, E., & Wölck, W. (1981). A method for isolating diagnostic linguistic variables: The Buffalo ethnolects experiment. In D. Sankoff & H. Cedergren (Eds.), *Variation omnibus* (pp. 17–24). Edmonton, Alberta: Linguistic Research.
- Carreira, M. (2004). Seeking explanatory adequacy: A dual approach to understanding the term “heritage language learner.”. *Heritage Language Journal*, 2(1), 1–25.
- Carroll, J., Carton, S., & Wilds, C. (1959). An investigation of “cloze” items in the measurement of achievement in foreign languages. *ERIC ED 021 513*.



- Cavanaugh, J. R. (2013). Language ideologies and language attitudes. A linguistic anthropological perspective. In P. Auer, J. Caro Reina, & G. Kaufmann (Eds.), *Language variation – European perspectives IV: Selected papers from the Sixth International Conference on Language Variation in Europe (ICLaVE 6), Freiburg, June 2011* (pp. 45–56). Amsterdam: John Benjamins. <https://doi.org/10.1075/silv.14.03cav>
- Chao, T. H. (1996). Chinese heritage community language schools in the United States. In *Schools in the United States* (pp. 7–13). Washington, DC: National Foreign Language Center.
- Chau, W. (2011). The influence of Dutch on Cantonese in the Netherlands. Rural versus urban areas (Unpublished MA thesis). Leiden University.
- Chen, H. (2012). Acquisition of Mandarin aspectual marking in child Heritage speakers in the Netherlands (Unpublished MA thesis). Radboud University, Nijmegen.
- Chen, J., & Shirai, Y. (2010). The development of aspectual marking in Mandarin Chinese. *Applied Psycholinguistics*, 31, 1–28. <https://doi.org/10.1017/S0142716409990257>
- Cheshire, J. (1999). Taming the vernacular: Some repercussions for the study of syntactic variation and spoken grammar = Estandarizando la lengua vernácula: Algunas repercusiones para el estudio de la variación sintáctica y la gramática del habla. *Cuadernos de filología inglesa*, 8(1), 59–80.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.
- Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, 1147, 192–208. <https://doi.org/10.1016/j.brainres.2007.01.137>
- Clahsen, H., Meisel, J., & Pienemann, M. (1983). *Deutsch als Zweitsprache: Der Spracherwerb ausländischer Arbeiter*. Tübingen: Narr.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620539>
- Clyne, M. G. (1967). *Transference and triggering: Observations on the language assimilation of postwar German-speaking migrants in Australia*. Den Haag: Martinus Nijhoff.
- Clyne, M. G. (1970). Bilingual speech phenomena (with special reference to German-English bilinguals in Victoria). *Kivung*, 3(2), 99–111.
- Clyne, M. G. (1982). *Multilingual Australia*. Melbourne: River Seine.
- Clyne, M. G. (2000). Lingua Franca and ethnolects in Europe and beyond. *Sociolinguistica*, 14, 83–89. <https://doi.org/10.1515/9783110245196.83>
- Clyne, M. G. (2003). *Dynamics of language contact. English and immigrant languages*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511606526>
- Clyne, M. G. (Ed.). (1976). *Australia talks: Essays on Australian immigrant and aboriginal languages* (Pacific Linguistics Series D, No. 23). Canberra: Australian National University.
- Clyne, M. G., & Kipp, S. (1999). *Pluricentric languages in an immigrant context: Spanish, Chinese, Arabic*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110805444>
- Cobarrubias, J. (1983). Ethical issues in status planning. In J. Cobarrubias & J. Fishman (Eds.), *Progress in language planning: International perspectives* (pp. 41–85). Berlin: Mouton de Gruyter.
- Cohen Henriquez, P. (1934). De taal van Curaçao. *Natuur en Mensch*, 54(4–5), 31–34.
- Cohen, R. (1997). *Global diasporas: An introduction*. Seattle, WA: Washington University Press. <https://doi.org/10.4324/9780203228920>
- Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly-proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50, 491–511. <https://doi.org/10.1016/j.jml.2004.02.002>

- Cremer, M., & Schoonen, R. (2013). The role of accessibility of semantic word knowledge in monolingual and bilingual fifth-grade reading. *Applied Psycholinguistics*, 34(6), 1195–1217. <https://doi.org/10.1017/S0142716412000203>
- Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3–49). Los Angeles, CA: Evaluation Dissemination and Assessment Center, California State University.
- Cummins, J. (1989). *Empowering minority students*. Sacramento, CA: California Association for Bilingualism Education.
- Cummins, J. (2005). A proposal for action: Strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *Modern Language Journal*, 89, 585–592.
- Dal Negro, S. (2004). *The decay of a language: The case of a German dialect in the Italian Alps*. Frankfurt: Peter Lang.
- Danesi, M. (1984). Canadian Italian: A case in point of how language adapts to environment. *Polyphony*, 7, 110–113.
- Danesi, M. (1985). Ethnic languages and acculturation: The case of Italo-Canadians. *Canadian Ethnic Studies*, 17, 98–103.
- De Houwer, A. (1999). Environmental factors in early bilingual development: The role of parental beliefs and attitudes. In G. Extra & L. Verhoeven (Eds.), *Bilingualism and migration* (pp. 75–95). Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110807820.75>
- De Houwer, A. (2009). *Bilingual first language acquisition*. Clevedon, UK: Multilingual Matters.
- De Houwer, A., & Bornstein, M. H. (2016). Bilingual mothers' language choice in child-directed speech: continuity and change. *Journal of Multilingual and Multicultural Development*, 37(7), 680–693. <https://doi.org/10.1080/01434632.2015.1127929>
- Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Phil. Trans. R. Soc. B.*, 369, 20120394.
- Deprez, C. (1996). Politique linguistique familiale: Le rôle des femmes. In C. Juillard & L.-J. Calvet (Eds.), *Politiques linguistiques, mythes et réalités (les). Actes des premières Journées scientifiques du réseau "Sociolinguistique et dynamique des langues" de l'AUF* (pp. 155–161). Beyrouth: AUF.
- Deuchar, M. (2005). Congruence and Welsh–English code-switching. *Bilingualism: Language and Cognition*, 8(3), 255–269. <https://doi.org/10.1017/S1366728905002294>
- Diaz, B., Baus, C., Escera, C., Costa, A., & Sebastian-Galles, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proc. Natl Acad. Sci. USA*, 105, 16083–16088. <https://doi.org/10.1073/pnas.0805022105>
- Dick, F., Bates, E., Wulfeck, B., Utman, J. A., Dronkers, N., & Gernsbacher, M. A. (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review*, 108(4), 759–88. <https://doi.org/10.1037/0033-295X.108.4.759>
- Dijkhoff, M. (1993). Papiamentu word formation, a case study of complex nouns and their relation to phrases and clauses (Unpublished doctoral dissertation). University of Amsterdam.
- Dijkhoff, M., Kouwenberg, S., & Tjon Sie Fat, P. (2006). The Dutch-speaking Caribbean. In U. Ammon, N. Dittmar, K. J. Mattheier, & P. Trudgill (Eds.), *International handbook of sociolinguistics* (pp. 2105–2114). Berlin: De Gruyter.

- Dijkstra, A., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5, 175–197. <https://doi.org/10.1017/S1366728902003012>
- Dijkstra, A. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll, & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 178–201). Oxford: Oxford University Press.
- Dijkstra, A., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, 62, 284–301. <https://doi.org/10.1016/j.jml.2009.12.003>
- Dijkstra, A., van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, 1, 51–66. <https://doi.org/10.1017/S1366728998000121>
- Dixon, R. M. W. (1972). *The Dyirbal language of North Queensland*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139084987>
- Doğruöz, A. S., & Backus, A. M. (2009). Innovative constructions in Dutch Turkish: An assessment of ongoing contact-induced change. *Bilingualism: Language and Cognition*, 12, 41–63. <https://doi.org/10.1017/S1366728908003441>
- Doğruöz, A. S. (2007). Synchronic variation and diachronic change in Dutch Turkish: A corpus based analysis (Unpublished doctoral dissertation). Tilburg University.
- Domínguez, L. (2013). *Understanding interfaces: second language acquisition and native language attrition of Spanish subject realization and word order variation*. (Language Acquisition and Language Disorders; No. 55). Amsterdam: John Benjamins.
- Domínguez, L., Hicks, G., & Slabakova, R. (In press). Terminology use in generative acquisition research. The case of ‘incomplete acquisition’ in heritage language grammars. *Studies in Second Language Acquisition*.
- Dorian, N. C. (1981). *Language death: The life cycle of a Scottish Gaelic dialect*. Philadelphia, PA: University of Pennsylvania Press. <https://doi.org/10.9783/9781512815580>
- Dorian, N. C. (1998). Western language ideologies and small language prospects. In L. A. Grenoble & L. J. Whaley (Eds.), *Endangered languages. Current issues and future prospects* (pp. 3–21). Cambridge: Cambridge University Press.
- Dorian, N. C. (Ed.) (1989). *Investigating obsolescence: Studies in language contraction and death*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620997>
- Dorleijn, M., & Nortier, J. (2009). Code-switching and the internet. In B. E. Bullock, & A. J. Toribio (Eds.), *The Cambridge handbook of linguistic code switching*, (pp. 127–141). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511576331.009>
- Dudek, J. B. (1925). The Czech language in America. *American Mercury*, 5, 202–07.
- Duranti, A. (1994). *From grammar to politics: Linguistic anthropology in a Western Samoan Village*. Berkeley, CA: University of California Press.
- Durgunoglu, A., & Verhoeven, L. (1998). *Literacy development in a multilingual context*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12, 453–76. <https://doi.org/10.1111/j.1467-9841.2008.00374.x>
- Eckkrammer, E. M. (2004). Passive voice in Papiamentu. A corpus-based study on dialectal variability. In G. Escure & A. Schwegler (Eds.), *Creoles, contact, and language change: Linguistic and social implications* (pp. 209–219). Amsterdam: John Benjamins. <https://doi.org/10.1075/cll.27.1oeck>

- England, N. C. (2003). Mayan language revival and revitalization politics: Linguists and linguistic ideologies. *American Anthropologist*, 105(4), 733–743.  
<https://doi.org/10.1525/aa.2003.105.4.733>
- Ernestus, M. (2000). *Voice assimilation and segment reduction in casual Dutch: a corpus-based study of the phonology-phonetics interface*. (Unpublished doctoral dissertation). VU Amsterdam.
- Espinosa, A. (1930). *Estudios sobre el Español de Nuevo Méjico. I. Fonética*. Traducción y notas de Amado Alonso y Angel Rosenblat. Buenos Aires: Boletín de Dialectología Hispánica.
- Espinosa, A. (1946). *Estudios sobre el Español de Nuevo Méjico. II. Morfología*. Traducción y notas de Amado Alonso y Angel Rosenblat. Buenos Aires: Boletín de Dialectología Hispánica.
- Estigarribia, B. (2015). Guaraní-Spanish mixing in Paraguay: Is jopara a third language, a language variety, or true codeswitching? *Journal of Language Contact*, 8(2), 183–222.  
<https://doi.org/10.1163/19552629-00802002>
- Evans, N. (2010). *Dying words: Endangered languages and what they have to tell us*. Walden, MA: Wiley-Blackwell.
- Evans, N., & Levinson, S. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32, 429–492.  
<https://doi.org/10.1017/S0140525X0999094X>
- Evans Braziel, J., & Mannur, A. (Eds.). (2003). *Theorizing diaspora: A reader*. Malden, MA: Blackwell.
- Eversteijn, N. (2011). “All at once”. Language choice and codeswitching by Turkish-Dutch teenagers (Unpublished doctoral dissertation). Tilburg University.
- Extra, G. (2002). The other languages of multicultural Europe: Perceptions, facts and educational policies. In S. Baker (Ed.), *Language policy: Lessons from global models* (pp. 130–151). Monterey, CA: Monterey Institute of International Studies.
- Extra, G. (2013). Mapping increasing linguistic diversity in multicultural Europe and abroad. In D. M. Singleton, J. A. Fishman, L. Aronin, & M. Ó Laoire (Eds.), *Current multilingualism: A new linguistic dispensation* (pp. 139–162). Berlin: De Gruyter Mouton.  
<https://doi.org/10.1515/9781614512813.139>
- Extra, G., & Yagmur, K. (2010). Language proficiency and socio-cultural orientation of Turkish and Moroccan youngsters in the Netherlands. *Language and Education*, 24(2), 117–132.  
<https://doi.org/10.1080/09500780903096561>
- Extra, G., & Verhoeven, L. (Eds.). (1993). *Community languages in the Netherlands*. Lisse: Swets & Zeitlinger.
- Extra, G., & Verhoeven, L. (1998). Immigrant minority groups and immigrant minority languages in Europe. In G. Extra & L. Verhoeven (Eds.), *Bilingualism and migration* (pp. 3–28). Berlin: Mouton de Gruyter.
- Extra, G., Aarts, R., van der Avoird, T., Broeder, P., & Yagmur, K. (2002). *De andere talen van Nederland. Thuis en op school*. Bussum: Coutinho.
- Eybers, G. J. (1916). *Testament Nobo di nos Señor y Salvador Jesu-Christo*. Amsterdam: Nederlands Bijbel Genootschap.
- Fenyvesi, A. (1995/1996). The case of American Hungarian case: Morphological change in McKeesport, PA. *Acta Linguistica Hungarica*, 43, 381–404.
- Fenyvesi, A. (Ed.). (2005). *Hungarian language contact outside Hungary*. Amsterdam: John Benjamins. <https://doi.org/10.1075/impact.20>
- Ferguson, C. A. (1959). Diglossia. *Word*, 15, 325–340.  
<https://doi.org/10.1080/00437956.1959.11659702>

- Fernández, E. M., Souza, R. A., & Carando, A. (2017). Bilingual innovations: Experimental evidence offers clues regarding the psycholinguistics of language change. *Bilingualism: Language and Cognition*, 20, 2, 251–268.
- Ferreira, F. (2003). The misinterpretation of non-canonical sentences. *Cognitive Psychology*, 47, 164–203. [https://doi.org/10.1016/S0010-0285\(03\)00005-7](https://doi.org/10.1016/S0010-0285(03)00005-7)
- Fishman, J. A. (1972). *The sociology of language; An interdisciplinary social science approach to language in society*. Rowley, MA: Newbury House.
- Fishman, J. A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon, UK: Multilingual Matters.
- Fishman, J. A. (2001a). 300-plus years of heritage language education in the United States. In J. Kreeft Peyton, D. Ranard, & S. McGinnis (Eds.), *Heritage languages in America. Preserving a national resource* (pp. 81–97). Washington, DC: Center for Applied Linguistics & Delta Systems.
- Fishman, J. A. (Ed.) (2001b). *Can threatened languages be saved?* Clevedon, UK: Multilingual Matters.
- Flecken, M. (2010). Event conceptualization in language production of early bilinguals (Unpublished doctoral dissertation). Radboud University, Nijmegen.
- Flege, J. E., MacKay, I. R., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*, 23, 567–598. <https://doi.org/10.1017/S0142716402004046>
- Flores, C., & Rato, A. (2016). Global accent in the Portuguese speech of heritage returnees. *Journal of Heritage Language*, 13(2), 161–183.
- Flores, C., Santos, A., Jesus, A., & Marques, R. (2017). Age and input effects in the acquisition of mood in Heritage Portuguese. *Journal of Child Language*, 44(4), 795–828. <https://doi.org/10.1017/S0305000916000222>
- Flores, N., & Toro, J. (2000). The persistence of dialect features under conditions of contact and leveling. *Southwest Journal of Linguistics*, 19(2), 31–42.
- Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology*. Cambridge, MA: The MIT Press.
- Fokker, A. A. (1914). Het Papiamentoe of basterd-spaans der West Indiese eilanden. *Tijdschrift voor Nederlandse Taal en Letterkunde*, 33, 54–79.
- Fouse, G. C. (2002). *The story of Papiamentu: a study in slavery and language*. Lanham, MD: University Press of America, Incorporated.
- Frazier, L., & Clifton Jr., C. (1996). *Construal*. Cambridge, MA: The MIT Press.
- Fricke, M., & Kootstra, G. J. (2016). Primed codeswitching in spontaneous bilingual dialogue. *Journal of Memory and Language*, 91, 181–201. <https://doi.org/10.1016/j.jml.2016.04.003>
- Gal, S. (1979). *Language shift*. New York, NY: Academic Press.
- Garcia, O. (2009). *Bilingual education in the 21st Century: A global perspective*. Oxford: Wiley-Blackwell.
- Garrett, P. B. (2005). What a language is good for: Language socialization, language shift, and the persistence of code-specific genres in St. Lucia. *Language in Society*, 34(3), 327–361. <https://doi.org/10.1017/S0047404505050128>
- Gaskell, M. G. (Ed.). (2007). *The Oxford handbook of psycholinguistics*. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198568971.001.0001>
- Gass, S. (1987). The resolution of conflicts among competing systems: A bidirectional perspective. *Applied Psycholinguistics*, 8, 329–50. <https://doi.org/10.1017/S0142716400000369>

- Gathercole, V. C. Mueller, & Thomas, E. M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition*, 12, 213–237. <https://doi.org/10.1017/S1366728909004015>
- Geense, P., & Tsui, T. C. (2001). Chinesees. In G. Extra & J. J. de Ruiter (Eds.), *Babylon aan de Noordzee. Nieuwe talen in Nederland* (pp. 78–100). Amsterdam: Bulaaq.
- Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, 22(3), 611–63. <https://doi.org/10.1017/S0305000900009971>
- Gershkoff-Stowe, L. (2002). Object naming, vocabulary growth, and the development of word retrieval abilities. *Journal of Memory and Language*, 46, 665–687. <https://doi.org/10.1006/jmla.2001.2830>
- Gertken, L. M., Amengual, M., Birdsong, D., & Birdsong, D. (2014). Assessing language dominance with the Bilingual Language Profile. Measuring L2 proficiency: Perspectives from SLA. *Second Language Acquisition*, 78, 208–225.
- Gibson, E., Piantadosi, S. T. & Fedorenko, K. (2011). Using Mechanical Turk to obtain and analyze English acceptability judgments. *Language and Linguistics Compass*, 5(8), 509–524. <https://doi.org/10.1111/j.1749-818X.2011.00295.x>
- Gijsberts, M., & J. Dagevos. (2005). *Uit elkaars buurt. De invloed van etnische concentratie op integratie en beeldvorming*. The Hague: SCP.
- Gilbert, G. G. (1965). English Loanwords in the German of Fredericksburg, Texas. *American Speech*, 40(2), 102–112. <https://doi.org/10.2307/453716>
- Giles, H., & P. Johnson. (1987). Ethnolinguistic identity theory: A social psychological approach to language maintenance. *International Journal of the Sociology of Language*, 68, 69–100.
- Giles, H., Bourhis, R. Y., & Taylor, D. M. (1977). Towards a theory of language in ethnic group relations. In H. Giles (Ed.), *Language, ethnicity and intergroup relations* (pp. 307–348). London: Academic Press.
- Givón, T. (1979). Prolegomena to any sane creology. In Ian F. Hancock (Ed.), *Readings in Creole studies* (pp. 3–35). Amsterdam: John Benjamins. <https://doi.org/10.1075/ssl5.2.01giv>
- Golato, A. (2002). German compliment responses. *Journal of Pragmatics*, 34(5), 547–571. [https://doi.org/10.1016/S0378-2166\(01\)00040-6](https://doi.org/10.1016/S0378-2166(01)00040-6)
- Gollan, T. H., & Goldrick, M. (2016). Grammatical constraints on language switching: Language control is not just executive control. *Journal of Memory and Language*, 90, 177–199.
- Gollan, T. H., & Goldrick, M. (2018). A switch is not a switch: Syntactically-driven bilingual language control. *J. Exp. Psychol. Learn. Mem. Cogn.*, 44(1), 143–156. <https://doi.org/10.1037/xlm0000462>
- Gollan, T. H., Montoya, R. I., Cera, C. & Sandoval, T. C. (2008). More use almost always a means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58(3), 787–814. <https://doi.org/10.1016/j.jml.2007.07.001>
- Gollan, T. H., Starr, J., & Ferreira, V. S. (2015). More than use it or lose it: The number-of-speakers effect on heritage language proficiency. *Psychonomic Bulletin & Review*, 22(1), 147–155. <https://doi.org/10.3758/s13423-014-0649-7>
- Gonzo, S., & Saltarelli, M. (1983). Pidginization and linguistic change in emigrant languages. In R. W. Andersen (Eds.), *Pidginization and creolization as language acquisition* (pp. 181–197). Rowley, MA: Newbury House.
- Green, D. W. (2011). Language control in different contexts: The behavioral ecology of bilingual speakers. *Frontiers in Psychology*, 2, article 103. <https://doi.org/10.3389/fpsyg.2011.00103>

- Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of cognitive psychology*.
- Green, E. (1961). English and Yiddish in Detroit. A survey and analysis of reciprocal influences in bilinguals' pronunciation, grammar, and dictionary (Unpublished doctoral dissertation). University of Michigan, Ann Arbor.
- Grimshaw, J. (1997). Projection, heads, and optimality. *Linguistic Inquiry*, 28, 373–422.
- Grin, F. (2003). *Language policy evaluation and the European Charter for Regional or Minority Languages*. London: Palgrave Macmillan. <https://doi.org/10.1057/9780230502666>
- Grodzinsky, Y., & Reinhart, T. (1993). The innateness of binding and coreference. *Linguistic Inquiry*, 24(1), 69–101.
- Groot, A. M. B. de. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. New York, NY: Psychology Press. <https://doi.org/10.4324/9780203841228>
- Grosjean, F. (2001). The bilingual's language modes. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Malden, MA: Blackwell.
- Grosjean, F., & Ping Li. (2013). *The psycholinguistics of bilingualism*. Oxford: Wiley-Blackwell.
- Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press.
- Guanglun, M. (2014). Learning Chinese as a heritage language in Australia and beyond: The role of capital. *Language and Education*, 28(5), 477–492. <https://doi.org/10.1080/09500782.2014.908905>
- Guijarro-Fuentes, P., & Marinis, T. (2011). Voicing language dominance. Acquiring Spanish by British English/Spanish bilingual children. In K. Potowski & J. Rothman (Eds.), *Bilingual youth. Spanish in English-speaking societies* (pp. 227–248). Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.42.14gui>
- Gullberg, M., Indefrey, P. & Muysken, P. (2009). Research techniques for the study of codeswitching. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge handbook of linguistic code-switching* (pp. 21–39). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511576331.003>
- Gumperz, J. J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511611834>
- Gutiérrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*, 24, 267–288. <https://doi.org/10.1017/S0142716403000158>
- Gynan, S. (1998). Migration patterns and language maintenance in Paraguay. *Journal of Sociolinguistics*, 2(2), 59–70. <https://doi.org/10.1111/1467-9481.00043>
- Hale, K., Krauss, M., Watahomigie, L. J., Yamamoto, A. Y., Craig, C., LaVerne Masayesva, J., & England, N. C. (1992). Endangered languages. *Language*, 68, 1–42.
- Halmari, H. (1997). *Government and codeswitching. Explaining American Finnish*. Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.12>
- Hammarström, H. (2010). The status of the least documented language families in the world. *Language Documentation and Conservation*, 4, 177–212.
- Hancock, A. (2014). Chinese complementary schools in Scotland and the continua of biliteracy. In A. Hancock & X. L. Curdt-Christiansen (Eds.), *Learning Chinese in diasporic communities: Many pathways to being Chinese* (pp. 59–80). Amsterdam: John Benjamins. <https://doi.org/10.1075/aals.12.04han>

- Hanulová, J., Davidson, D. J., & Indefrey, P. (2011). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Language and Cognitive Processes*, 26, 902–934.  
<https://doi.org/10.1080/01690965.2010.509946>
- Harrington, M. (1987). Processing transfer: Language-specific processing strategies as a source of interlanguage variation. *Applied Psycholinguistics*, 8, 351–377.  
<https://doi.org/10.1017/S0142716400000370>
- Harrison, B. (2000). Passing on the language: Heritage language diversity in Canada. <<http://www.statcan.gc.ca/pub/11-008-x/2000002/article/5165-eng.pdf>> (August 4, 2015).
- Hartling, A. S. (2016). The ubiquitous Spanish *se* – Monoclausal reciprocal constructions in heritage Danish in Argentina. Presented at Sociolinguistics Symposium 21, Murcia, Spain.
- Hartog, J. (1968). *Curaçao, from colonial dependence to autonomy*. Aruba: De Wit.
- Hartsuiker, R. J., & Pickering, M. J. (2008). Language integration in bilingual sentence production. *Acta Psychologica*, 128, 479–489. <https://doi.org/10.1016/j.actpsy.2007.08.005>
- Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, 15, 409–414. <https://doi.org/10.1111/j.0956-7976.2004.00693.x>
- Haspelmath, M. (2008). A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery*, 6(1), 40–63. <https://doi.org/10.1349/PS1.1537-0852.A.331>
- Hatzidaki, A., Branigan, H. P., & Pickering, M. J. (2011). Co-activation of syntax in bilingual language production. *Cognitive Psychology*, 62, 123–150.  
<https://doi.org/10.1016/j.cogpsych.2010.10.002>
- Haugen, E. (1953). *The Norwegian language in America: A study in bilingual behavior, I: The bilingual community. II: The American dialects of Norwegian*. Philadelphia, PA: University of Pennsylvania Press. <https://doi.org/10.9783/9781512820522>
- Haugen, E. (1956). *Bilingualism in the Americas: A bibliography and research guide* (Publication 26 of the American Dialect Society). Birmingham, AL: University of Alabama Press.
- He, A. W. (2010). The heart of heritage: Sociocultural dimensions of heritage language learning. *Annual Review of Applied Linguistics*, 30, 66–82.  
<https://doi.org/10.1017/S0267190510000073>
- He, A. W., & Yun, X. (Eds.). (2008). *Chinese as a Heritage Language: Fostering rooted world citizenry*. Honolulu, HI: NFLRC/University of Hawaii Press.
- Heine, B., & Kuteva, T. (2003). On contact-induced grammaticalization. *Studies in Language*, 27(3), 529–572. <https://doi.org/10.1075/sl.27.3.04hei>
- Hell, J. G. van, & Dijkstra, A. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin and Review*, 9, 780–789.  
<https://doi.org/10.3758/BF03196335>
- Henzl, V. (1982). American Czech: A comparative study of linguistic modifications in immigrant and young children's speech. In R. Sussex (Ed.) *Slavic languages in emigré communities* (pp. 33–46). Edmonton, Alberta: Linguistic Research.
- Hernandez Chavez, E., Cohen, A. D., & Beltrano, A. F. (1975). *El lenguaje de los Chicanos: Regional and social characteristics of language used by Mexican Americans*. Arlington, VA: Center for Applied Linguistics.
- Hesseling, D. C. (1933). Een Spaans boek over het Papiaments. *Tijdschrift voor Nederlandse Taal- en Letterkunde*, 52, 40–69.
- Hill, J., & Hill, K. (1986). *Speaking Mexicano. Dynamics of syncretic language in Central Mexico*. Tucson, AZ: University of Arizona Press.



- Hlavac, J. (2003). *Second-generation speech: Lexicon, code-switching and morpho-syntax of Croatian-English bilinguals*. Bern: Peter Lang.
- Hlavac, J. (2012). Psycholinguistic, metalinguistic and socio-psychological accounts of code-switching: A comparative analysis of their incidence in a large Croatian-English sample. *Suvremena lingvistika/Contemporary Linguistics*, 73, 47–71.
- Hoff, E., Rumiche, R., Burrige, A., Ribot, K. M., & Welsh, S. N. (2014). Expressive Vocabulary Development in Children from Bilingual and Monolingual Homes: A Longitudinal Study from Two to Four Years. *Early childhood research quarterly*, 29(4), 433–444. <https://doi.org/10.1016/j.ecresq.2014.04.012>
- Holm, J. A. (1988). *Pidgins and Creoles* (2 Vols). Cambridge: Cambridge University Press.
- Hornberger, N. H. (1988). *Bilingual education and language maintenance: A southern Peruvian case*. Dordrecht: Foris. <https://doi.org/10.1515/9783110849158>
- Hornberger, N. H., & Wang, S. C. (2008). Who are our heritage language learners? Identity and biliteracy in heritage language education in the United States. In D. M. Brinton, O. Kagan, & S. Bauckus (Eds.), *Heritage language education: A new field emerging* (pp. 3–35). New York, NY: Routledge.
- Hyams, N. M. (1986). *Language acquisition and the Theory of Parameters*. Dordrecht: Reidel. <https://doi.org/10.1007/978-94-009-4638-5>
- Hyltenstam, K., Bylund, E., Abrahamsson, N., & Park, H.-S. (2009). Dominant language replacement: The case of international adoptees. *Bilingualism: Language and Cognition*, 12(2), 121–140. <https://doi.org/10.1017/S1366728908004008>
- Indefrey, P., Şahin, H., & Gullberg, M. (2017). The expression of spatial relationships in Turkish-Dutch bilinguals. *Bilingualism: Language and Cognition*, 20(3), 473–493. <https://doi.org/10.1017/S1366728915000875>
- Irizarri van Suchtelen, P. (2016). Spanish as a heritage language in the Netherlands. A cognitive linguistic exploration (Unpublished doctoral dissertation). Radboud University, Nijmegen.
- Isac, D., & Reiss, C. (2013). *I-language: An introduction to linguistics as cognitive science* (2nd ed.). Oxford University Press.
- Jackendoff, R. S. (1997). *The architecture of the language faculty*. Cambridge, MA: The MIT Press.
- Jacobs, B. (2009). The Upper Guinea origins of Papiamentu. Linguistic and historical evidence. *Diachronica*, 26(3), 319–379. <https://doi.org/10.1075/dia.26.3.02jac>
- Jacobs, B. (2011). The origin and originality of passivization in Papiamentu. *Journal of Portuguese Linguistics*, 10(2), 31–56. <https://doi.org/10.5334/jpl.96>
- Jacobs, B. (2014). The phonological and morphosyntactic integration of Dutch loans in Papiamentu. Manuscript.
- Jacobs, B. with M. van der Wal. (2015). The discovery, nature, and implications of a Papiamentu text fragment from 1783. *Journal of Pidgin and Creole Languages*, 30(1), 44–62. <https://doi.org/10.1075/jpcl.30.1.02jac>
- Jacobson, R. (1977). The social implications of intra-sentential code-switching. New directions in Chicano scholarship (special issue). *New Scholar*, 6, 227–256.
- Jakobson, R. (1941). *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala, Sweden: Almqvist & Wiksell.
- Jansen, B., Lalleman, J., & Muysken, P. (1981). The alternation hypothesis: Acquisition of Dutch word order by Turkish and Moroccan foreign workers. *Language Learning: A Journal of Applied Linguistics*, 31, 315–336. <https://doi.org/10.1111/j.1467-1770.1981.tb01387.x>
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York, NY: Routledge. <https://doi.org/10.4324/9780203935927>

- Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the US. *Applied Psycholinguistics*, 24(1), 131–163.  
<https://doi.org/10.1017/S0142716403000079>
- Jia, G., Aaronson, D., & Wu, Y. (2002). Long-term language attainment of bilingual immigrants: Predictive variables and language group differences. *Applied Psycholinguistics*, 23, 599–621.  
<https://doi.org/10.1017/S0142716402004058>
- Johanson, L. (1992). *Strukturelle Faktoren in türkischen Sprachkontakten*. (Sitzungsberichte der Wissenschaftlichen Gesellschaft an der J. W. Goethe-Universität Frankfurt am Main, 29:5). Stuttgart: Steiner.
- Johanson, L. (2002). Contact-induced linguistic change in a codecopying framework. In M. C. Jones, & E. Esch (Eds.), *Language change: The interplay of internal, external and extra-linguistic factors* (Contributions to the sociology of language, 86) (pp. 285–313). Berlin: Mouton de Gruyter.
- Johanson, L. (2008). Remodeling grammar. Copying, conventionalization, grammaticalization. In P. Siemund & N. Kintana (Eds.), *Language contact and contact languages* (pp. 61–79). (Hamburg Studies on Multilingualism 7.) Amsterdam: Benjamins.
- Johnson, J. S. & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60–99. [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0)
- Jonker, E. F. (2009). Toen de Amboneesjes kwamen. Achtergrond, theorie en praktijk van het (bicultureel) onderwijs aan drie generaties Molukse kinderen in Nederland 1951–2008. In A. E. M. G. Minnaert, H. C. Lutje Spelberg, & H. T. A. Amsing (Eds.), *Het pedagogisch quotiënt. Pedagogische kwaliteit in opvoeding, hulpverlening en educatie* (pp. 229–260). Houten: Bohn, Stafleu, Van Loghum. [https://doi.org/10.1007/978-90-313-7977-4\\_12](https://doi.org/10.1007/978-90-313-7977-4_12)
- Jørgensen, J. N. (2003). Bilingualism and minority languages. *International Journal of the Sociology of Language*, 159, 73–92.
- Jørgensen, J. N. (2008). Polylingual languaging around and among children and adolescents, I. *International Journal of Multilingualism*, 5(3), 161–176.  
<https://doi.org/10.1080/14790710802387562>
- Joubert, S. M. (1999). *Handwoordenboek Nederlands-Papiaments*. Zutphen: Koninklijke Wöhrmann.
- Jurafsky, D. (2003). Probabilistic modeling in psycholinguistics: linguistic comprehension and production. In R. Bod, J. Hay, & S. Jannedy (Eds.) *Probability Theory in linguistics* (pp. 39–96). Cambridge, MA: The MIT Press.
- Kagan, O., & Friedman, D. (2003). Using the OPI to place heritage speakers of Russian. *Foreign Language Annals*, 36(4), 536–545. <https://doi.org/10.1111/j.1944-9720.2003.tb02143.x>
- Kallmeyer, W., & Keim, I. (2003). Linguistic variation and the construction of social identity in a German-Turkish setting. A case study of an immigrant youth group in Mannheim, Germany. In J. Androutsopoulos & A. Georgakopoulou (Eds.), *Discourse constructions of youth identities* (pp. 29–46). Amsterdam: John Benjamins.  
<https://doi.org/10.1075/pbns.110.03kal>
- Kasstan, J. R., Auer, A., & Salmons, J. (2018). Heritage-language speakers: Theoretical and empirical challenges on sociolinguistic attitudes and prestige. *International Journal of Bilingualism*, 22(4), 387–394. <https://doi.org/10.1177/1367006918762151>
- Kehoe, M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition*, 7(1), 71–88.  
<https://doi.org/10.1017/S1366728904001282>

- Kehoe, M., Lleó, C., & Rakow, M. (2011). Speech rhythm in the pronunciation of German and Spanish monolingual and German-Spanish bilingual 3-year-olds. *Linguistische Berichte*, 227, 323–351.
- Keijzer, M. (2010a). *First language acquisition and first language attrition: Parallels and divergences*. Saarbrücken: Lambert Academic Publishing.
- Keijzer, M. (2010b). The regression hypothesis as a theoretical framework for first language attrition. *Bilingualism: Language and Cognition*, 13(1), 9–18.  
<https://doi.org/10.1017/S1366728909990356>
- Keim, I., & Cindark, I. (2002). Der Umgang mit dem negativen Face in zwei jugendlichen Migrant Innengruppen: Gespielt aggressive vs. ironische Formen der Kritik. In D. Hartung & A. Shethar (Eds.), *Kulturen und ihre Sprachen. Die Wahrnehmung anders Sprechender und ihr Selbstverständnis* (pp. 57–86). Berlin: Trafo.
- Kester, E.-P., & Fun, J. (2012). Language use, language attitudes and identity among Aruban students in the Netherlands. In N. Faraclas, R. Severing, C. Weijer, & L. Echteld (Eds.), *Multiple cultures and citizenships* (pp. 231–248). Curaçao: University of Curaçao /Fundashon pa Planifikashon di Idioma.
- Kester, E.-P., & Hortencia, T. E. (2010). Language use, language attitudes and identity among Curaçaoan students in the Netherlands. In N. Faraclas, R. Severing, C. Weijer, & E. Echteld (Eds.), *Crossing shifting boundaries: Language and changing political status in Aruba, Bonaire and Curaçao* (pp. 49–66). Curaçao: University of Curaçao /Fundashon pa Planifikashon di Idioma.
- King, K. A. (2016). Language policy, multilingual encounters, and transnational families. *Journal of Multilingual and Multicultural Development*, 37(7), 726–733.  
<https://doi.org/10.1080/01434632.2015.1127927>
- Klatter, J. A. K., & Kroon, S. (Eds.). (1997). *Dutch overseas: Studies in maintenance and loss of Dutch as an immigrant language*. Tilburg: Tilburg University Press.
- Klein, F. (1980). A quantitative study of syntactic and pragmatic indications of change in the Spanish of bilinguals in the US. In W. Labov (Ed.), *Locating language in time and space* (pp. 69–82). London: Academia Press.
- Klein, W., & Perdue, C. (1992). *Utterance structure: Developing grammars again*. Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.5>
- Klein, W., & Dittmar, N. (1979). *Developing grammars*. Berlin: Springer.  
<https://doi.org/10.1007/978-3-642-67385-6>
- Klooster, W., & Oostindie, G. (2011). *Curaçao in the age of revolutions, 1795–1800*. Leiden: KITLV Press.
- Knightly, L. M., Jun, S. A., Oh, J. S., & Au, T. K. F. (2003). Production benefits of childhood overhearing. *Journal of the Acoustical Society of America*, 114(1), 465–474.  
<https://doi.org/10.1121/1.1577560>
- Kontra, M. (1998). *The sociolinguistics of Hungarian outside Hungary project (Final Report to the Research Support Scheme)*. Budapest: Linguistics Institute, Hungarian Academy of Sciences.
- Kontra, M. (2006). Hungarian in- and outside Hungary. In U. Ammon, N. Dittmar, K. Mattheier, & P. Trudgill (Eds.), *Sociolinguistics/Soziolinguistik* (2nd completely revised and extended ed., Volume 3, pp. 1811–1818). Berlin: Walter de Gruyter.
- Kook, H., & Narain, G. (1993). Papiamentu. In G. Extra & L. Verhoeven (Eds.) *Community languages in the Netherlands* (pp. 69–92). Lisse: Swets & Zeitlinger.
- Koontz Garboden, A. (2004). Language contact and Spanish aspectual expression: A formal analysis. *Lingua*, 114, 1291–1330. <https://doi.org/10.1016/j.lingua.2003.07.008>

- Koornneef, A., Avrutin, S., Wijnen, F., & Reuland, E. (2011). Tracking the preference for bound-variable dependencies in ambiguous ellipses and only-structures. In J. Runner (Ed.), *Experiments at the interfaces* (pp. 67–100). Bingley: Emerald.
- Kootstra, G. J., Dijkstra, A., & Starren, M. In J. D. Wright (Ed.), *International encyclopedia of social and behavioral sciences*. (2015). Second language acquisition. (2nd ed., Vol. 21, pp. 349–359). Oxford: Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.53025-6>
- Kootstra, G. J., van Hell, J. G., & Dijkstra, A. (2009). Two speakers, one dialogue. An interactive alignment perspective on code-switching in bilingual speakers. In L. Isurin, D. Winford, & K. de Bot (Eds.), *Multidisciplinary approaches to code switching* (pp. 129–159). Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.41.09koo>
- Kootstra, G. J., van Hell, J. G., & Dijkstra, A. (2012). Priming of code-switches in sentences: The role of lexical repetition, cognates, and language proficiency. *Bilingualism: Language and Cognition*, 15, 797–819. <https://doi.org/10.1017/S136672891100068X>
- Kootstra, G. J., & Doedens, W. J. (2016). Cross-language interaction and convergence in bilingual sentence production: Immediate and long-lasting effects of structural priming and verb bias between languages. *Bilingualism: Language and Cognition*, 19(4), 710–732. <https://doi.org/10.1017/S1366728916000420>
- Kootstra, G. J. & Rossi, E. (2017). Moving beyond the priming of single-language sentences: A proposal for a comprehensive model to account for linguistic representation in bilinguals. *Behavioral and Brain Sciences*, 40, e297.
- Kootstra, G. J., & Şahin, H. (2018). Cross-language structural priming as a potential mechanism of contact-induced language change. *Language*, 94(4), 902–930.
- Kootstra, G. J., van Hell, J. G., & Dijkstra, A. (2010). Syntactic alignment and shared word order in code-switched sentence production: Evidence from bilingual monologue and dialogue. *Journal of Memory and Language*, 63, 210–231. <https://doi.org/10.1016/j.jml.2010.03.006>
- Kopeliovich, S. (2010). Family language policy: A case study of a Russian-Hebrew bilingual family – Toward a theoretical framework. *Diaspora, Indigenous, and Minority Education*, 4(3), 162–178. <https://doi.org/10.1080/15595692.2010.490731>
- Kouwenberg, S., & Murray, E. (1994). *Papiamentu*. Munich: Lincom.
- Kouwenberg, S. (1994). *A grammar of Berbice Dutch Creole*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110885705>
- Kovács, M. (2001). *Code-switching and language shift in Australian Finnish in comparison with Australian Hungarian*. Åbo: Åbo Akademi University Press.
- Kowalik, S., & Kramer, J. (1994). Influencias del neerlandés en el papiamento. In J. Lüdtke & M. Perl (Eds.), *Sprache und Kultur in der Karibik* (pp. 147–165). Tübingen: Niemeyer.
- Kramer, J. (2004). *Die iberoromanische Kreolsprache Papiamentu*. Hamburg: Helmut Buske.
- Krauss, M. E. (2007). Mass language extinction and documentation: The race against time. In O. Miyaoka, O. Sakiyama, & M. Krauss (Eds.), *Vanishing languages of the Pacific Rim* (pp. 3–24). Oxford: Oxford University Press.
- Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, 9, 119–135. <https://doi.org/10.1017/S1366728906002483>
- Kroll, J. F., & Groot, A. M. B. de. (2005). *Handbook of bilingualism: Psycholinguistic approaches*. Oxford: Oxford University Press.
- Kroll, J. F., Bogulski, C. A., McClain, R. (2012). Psycholinguistic perspectives on second language learning and bilingualism: The course and consequence of cross-language competition. *Linguistic Approaches to Bilingualism*, 2, 1–24. <https://doi.org/10.1075/lab.2.1.01kro>

- Kroll, J. F., Bobb, S., & Hoshino, N. (2014). Two Languages in Mind. *Current directions in psychological science*, 23, 159–163.
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5, 831–843. <https://doi.org/10.1038/nrn1533>
- Kuhn, J. (2001). Generation and parsing in optimality theoretic syntax – Issues in the formalization of OT-LFG. In P. Sells (Ed.) *Formal and empirical issues in optimality-theoretic syntax* (pp. 313–366). Stanford, CA: CSLI.
- Kupisch, T. (2013). A new term for a better distinction? A view from the higher end of the proficiency scale. *Theoretical Linguistics*, 39(3–4), 203–214.
- Kupisch, T. (2014). Adjective placement in simultaneous bilinguals (German–Italian) and the concept of cross-linguistic overcorrection. *Bilingualism: Language and Cognition*, 17(1), 222–233. <https://doi.org/10.1017/S1366728913000382>
- Kupisch, T. (2018). 2L1 simultaneous bilinguals as heritage speakers in morphosyntax and phonology. Introduction: Recent developments in early bilingualism. *Bilingualism*, 21(4), 653–655.
- Kupisch, T., Barton, D., Hailer, K., Kostogryz, E., Lein, T., Stangen, I., & van de Weijer, J. (2015). Foreign accent in adult simultaneous bilinguals. *Heritage Language Journal*, 11(2), 123–150.
- Kupisch, T., & Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22(5), 564–582.
- Kupisch, T., & van de Weijer, J. (2016). The role of the childhood environment for language dominance: A study of adult simultaneous bilingual speakers of German and French. In J. Treffers-Daller & C. Silva-Corvalán (Eds.), *Operationalising and measuring language dominance* (pp. 174–194). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781107375345.009>
- Kusters, W. (2003). Linguistic complexity: The influence of social change on verbal inflection. Utrecht: LOT Dissertation Series.
- La Morgia, F. (2011). Who is afraid of multilingualism? Evaluating the linguistic impact of migration in Ireland. In M. Darmody, N. Tyrrell, & S. Song (Eds.), *The changing faces of Ireland: Exploring the lives of immigrant and ethnic minority children* (pp. 3–16). Rotterdam: SensePublishers. [https://doi.org/10.1007/978-94-6091-475-1\\_1](https://doi.org/10.1007/978-94-6091-475-1_1)
- Labov, W. (1972a). Some principles of linguistic methodology. *Language in Society*, 1, 97–120. <https://doi.org/10.1017/S0047404500006576>
- Labov, W. (1972b). *Sociolinguistic Patterns*. Philadelphia, PA: University of Pennsylvania Press.
- Labov, W. (1989). The child as linguistic historian. *Language Variation and Change*, 1, 85–97. <https://doi.org/10.1017/S09543394500000120>
- Labov, W. (1994). *Principles of linguistic change, Vol. 1: Internal factors*. Oxford: Basil Blackwell.
- Laleko, O., & Kawamura, H. (2011). Grammatical restructuring in intergenerational language shift: A case of heritage Japanese. Presented at *The 56th Annual Conference of the International Linguistics Association* 15–17 April, Rutgers University, New Brunswick, NJ.
- Laleko, O. (2010). The syntax-pragmatics interface in language loss: Covert restructuring of aspect in heritage Russian (Unpublished doctoral dissertation). University of Minnesota, Minneapolis.
- Laleko, O., & Polinsky, M. (2016). Between syntax and discourse: Topic and case marking in heritage speakers and L2 learners of Japanese and Korean. *Linguistic Approaches to Bilingualism*, 6(4), 396–439. <https://doi.org/10.1075/lab.14018.lal>

- Lambert, W. E. (1975). Culture and language as factors in learning and education. In A. Wolfgang (Ed.), *Education of immigrant students*. Toronto: Ontario Institute for Studies in Education.
- Lambert, W. E. (1977). The effects of bilingualism on the individual: Cognitive and socio-cultural consequences. In P. A. Hornby (Ed.), *Bilingualism: Psychological, social and educational implications* (pp. 15–28). New York, NY: Academic Press.
- Lambert, W. E. (1981). Bilingualism and language acquisition. In H. Winitz (Ed.), *Native language and foreign language acquisition* (pp. 9–22). New York, NY: New York Academy of Science.
- Langacker, R. W. (1987). *Foundations of cognitive grammar, Vol. I: Theoretical prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. New York, NY: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195331967.001.0001>
- Lanza, E., & Ailin Svendsen, B. (2007). Tell me who your friends are and I *might* be able to tell you what language(s) you speak: Social network analysis, multilingualism and identity. *International Journal of Bilingualism*, 11(3), 275–300. <https://doi.org/10.1177/13670069070110030201>
- Lanza, E. (1997). *Language mixing in infant bilingualism. A sociolinguistic perspective*. Oxford: Clarendon Press.
- Legendre, G., Putnam, M. T., de Swart, H., & Zaroukian, E. (Eds.) (2016). *Optimality-Theoretic Syntax, Semantics, and Pragmatics: From Uni- to Bidirectional Optimization* (pp. 300–324). Oxford: Oxford University Press.
- Lehtinen, M. K. T. (1966). An analysis of a Finnish-English bilingual corpus (Unpublished doctoral dissertation). Indiana University, Bloomington. Distributed by University Microfilms, Ann Arbor, MI.
- Lein, T., Kupisch, T., & van de Weijer, J. (2016). Voice onset time and global foreign accent in German–French simultaneous bilinguals during adulthood. *International Journal of Bilingualism*, 20(6), 732–74. <https://doi.org/10.1177/1367006915589424>
- Lenz, R. (1928). *El papiamento: La lengua criolla de Curazao*. Santiago de Chile: Balcells & Cia.
- Lestrade, S. (2010). The space of case (Unpublished doctoral dissertation). Radboud University, Nijmegen.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38; discussion 38–75. <https://doi.org/10.1017/S0140525X99001776>
- Levine, G. (2000). *First-language acquisition in the immigrant situation: Yiddish in the United States* (Linguistische Arbeiten 426). Tübingen: Max Niemeyer.
- Li, D. (1999). The functions and status of English in Hong Kong: A post-1997 update. *English World-Wide*, 20(1), 67–110. <https://doi.org/10.1075/eww.20.1.03li>
- Li, Wei. (1994). *Three generations, two languages, one family. Language choice and language shift in a Chinese community in Britain*. Clevedon, UK: Multilingual Matters.
- Li, Wei. (2007). Dimensions of bilingualism. In Li Wei (Ed.), *The bilingualism reader* (2nd ed., pp. 3–22). London: Routledge.
- Li, Wei. (2011). Moment Analysis and translanguaging space: Discursive construction of identities by multilingual Chinese youth in Britain. *Journal of Pragmatics*, 43(5), 1222–1235. <https://doi.org/10.1016/j.pragma.2010.07.035>

- Li, Wei & W. Green, D. W. (2014). A control process model of code-switching. *Language and Cognitive Processes*, 29. <https://doi.org/10.1080/23273798.2014.882515>
- Li, Wei, & Hua, Z. (2010). Voices from the diaspora: Changing hierarchies and dynamics of Chinese multilingualism. *International Journal of the Sociology of Language*, 205, 155–171.
- Lindsey, K. L. (2013). How HALA and the BLP can be used for an earlier detection of language attrition in the Chuvash community. 3rd Annual International Conference of Language Documentation & Conservation. Manoa, Hawaii.
- Linguistic Minorities Project. (1985). *The other languages of England*. London: Routledge.
- Liu, H., Bates, E., & Li, P. (1992). Sentence interpretation in bilingual speakers of English and Chinese. *Applied Psycholinguistics*, 13, 451–484. <https://doi.org/10.1017/S0142716400005762>
- Liu, Na. (2013). The linguistic vitality of Chinese in the United States. *Heritage Language Journal*, 10, 1–11.
- Lleó, C., & Kehoe, M. (Eds.). (2002). On the interaction of phonological systems in child bilingual acquisition. *International Journal of bilingualism*, 6(3), 233–237. <https://doi.org/10.1177/13670069020060030101>
- Lleó, C. (2016). Bilingualism and child phonology. *Oxford Handbooks Online*, May. <https://doi.org/10.1093/oxfordhb/9780199935345.01353>
- Lo Bianco, J. (2008). Policy activity for heritage languages: Connections with representation and citizenship. In D. M. Brinton, O. Kagan, & S. Bauckus (Eds.), *Heritage language education: A new field emerging* (pp. 53–69). New York, NY: Routledge.
- Lobo, V. R. (2013). *L2 teaching at a very early age: A study of Dutch schools*. Utrecht: LOT.
- Loebell, H., & Bock, J. K. (2003). Structural priming across languages. *Linguistics*, 41, 791–824. <https://doi.org/10.1515/ling.2003.026>
- Los, B. L. J., & Starren, M. (2013). A typological switch in early Modern English – and the beginning of one in Dutch? *Leuvense Bijdragen. Tijdschrift voor Germaanse Filologie*, 98, 98–126.
- Louden, M. L. (2016). *Pennsylvania Dutch. The story of an American language*. Baltimore, MD: Johns Hopkins University Press.
- Lynch, A. (2013). Observaciones sobre comunidad y (dis)continuidad en el estudio sociolingüístico del español en Estados Unidos. In D. Dumitrescu & G. Piña-Rosales (Eds.), *El español en los Estados Unidos: E pluribus unum? Enfoques multidisciplinares* (pp. 67–83). New York, NY: Academia Norteamericana de la Lengua Española.
- Lynch, A. (2014). The first decade of the Heritage Language Journal: A retrospective view of research on heritage languages. *Heritage Language Journal*, 11(3), 224–242.
- Lyra, F. (1962). English and Polish in Contact (Unpublished doctoral dissertation). Indiana University, Bloomington. Distributed by University Microfilms, Ann Arbor, MI.
- Maduro, A. J. (1953). *Ensayo pa yega na un ortografia uniforme pa nos Papiamentu*. Curaçao: Scherpenhevel.
- Mahon, M., & Crutchley, A. (2006). Performance of typically-developing school-age children with English as an additional language on the British Picture Vocabulary Scales II. *Child Language Teaching and Therapy*, 22(3), 333–351. <https://doi.org/10.1191/0265659006ct311xx>
- Makihara, M. (2001). Modern Rapa Nui adaptation of Spanish elements. *Oceanic Linguistics*, 40, 191–223. <https://doi.org/10.1353/ol.2001.0021>
- Manosuthikit, A. (2013). Language ideologies and practices of a Burmese community in the US: A critical perspective on multilingualism (Unpublished doctoral dissertation). University of Wisconsin, Madison.

- Maratsos, M. (2000). More overregularisations after all: New data and discussion on Marcus, Pinker, Ullman, Hollander, Rosen & Xu. *Journal of Child Language*, 27, 183–212. <https://doi.org/10.1017/S0305000999004067>
- Martinus, E. F. (1990). Papiamentu: Road to emancipation. In C. Hagège (Ed.), *Language reform: History and future* (Vol. 5, pp. 127–147). Hamburg: Helmut Buske.
- Martinus, E. F. (1996). The kiss of a slave: Papiamentu's West-African connections (Unpublished doctoral dissertation). Universiteit van Amsterdam.
- Maschler, Y. (1998). On the transition from code-switching to a mixed code. In P. Auer (Ed.), *Code-switching in conversation* (pp. 125–149). London: Routledge.
- Matras, Y. (1998). Utterance modifiers and universals of grammatical borrowing. *Linguistics*, 36(2), 281–331. <https://doi.org/10.1515/ling.1998.36.2.281>
- Matras, Y. (2009). *Language contact*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511809873>
- Maurer, P. (1998). *Les modifications temporelles et modales du verbe dans le Papiamentu de Curaçao (Antilles Néerlandaises)*. Hamburg: Helmut Buske.
- Mayer, M. (1969). *Frog, where are you? (A boy, a dog, and a frog)*. New York, NY: Dial Books.
- McDonald, J. L. (1987). Sentence interpretation in bilingual speakers of English and Dutch. *Applied Psycholinguistics*, 8, 379–413. <https://doi.org/10.1017/S0142716400000382>
- Meara, P., Milton, J., & Lorenzo-Dus, N. (2003). *Swansea Language Aptitude Tests (LAT) v2.0*. Swansea: Lognostics.
- Meara, P. (2005). *LLAMA language aptitude tests: The manual*. Swansea: Lognostics.
- Meisel, J. M. (1989). Early differentiation of languages in bilingual children. *Bilingualism across the lifespan: Aspects of acquisition, maturity and loss* (pp. 13–40). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511611780.003>
- Meisel, J. (2014). Heritage language learners: Incomplete acquisition of grammar in early childhood. In A. Enrique-Arias, M. J. Gutiérrez, A. Landa, & F. Ocampo (Eds.), *Perspectives in the study of Spanish language variation* (pp. 429–454). Santiago de Compostela: Universidade de Santiago de Compostela.
- Mencken, H. L. (1936) [1947]. *The American language* (4th ed.). New York, NY: Knopf.
- Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, 40, 25–40. <https://doi.org/10.1006/jmla.1998.2602>
- Meuter, R. (2005). Language selection in bilinguals: Mechanisms and processes. In J. F. Kroll, A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 349–370). Oxford: Oxford University Press.
- Meuter, R. (2009). Language selection and performance optimisation in multilinguals. In A. Pavlenko (Ed.), *Multidisciplinary approaches to code switching* (pp. 27–51). Clevedon, UK: Multilingual Matters. <https://doi.org/10.1075/sibil.41.05meu>
- Milroy, L., & Li, Wei. (1995). A social network perspective on code-switching and language choice: The example of the Tyneside Chinese community. In L. Milroy & P. Muysken (Eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching* (pp. 136–157). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620867.007>
- Montrul, S. (2004a). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition*, 7(2), 125–142. <https://doi.org/10.1017/S1366728904001464>



- Montrul, S. (2004b). *The acquisition of Spanish. Morphosyntactic development in monolingual and bilingual L1 acquisition and in adult L2 acquisition*. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.37>
- Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.39>
- Montrul, S. (2010). Dominant language transfer in adult second language learners and heritage speakers. *Second Language Research*, 26(3), 293–327. <https://doi.org/10.1177/0267658310365768>
- Montrul, S. (2011). Morphological errors in Spanish second language learners and heritage speakers. *Studies in Second Language Acquisition*, 33(2), 163–192. <https://doi.org/10.1017/S0272263110000720>
- Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139030502>
- Montrul, S., & Potowski, K. (2007). Command of gender agreement in school-age Spanish bilingual children. *International Journal of Bilingualism*, 11(3), 301–328. <https://doi.org/10.1177/13670069070110030301>
- Montrul, S., & Bowles, M. (2009). Back to basics: Incomplete knowledge of Differential Object Marking in Spanish heritage speakers. *Bilingualism: Language and Cognition*, 12(3), 363–383. <https://doi.org/10.1017/S1366728909990071>
- Montrul, S., Bhatt, R., & Bhatia, A. (2012). Erosion of case and agreement in Hindi heritage Speakers. *Linguistic Approaches to Bilingualism*, 2, 141–176. <https://doi.org/10.1075/lab.2.2.02mon>
- Montrul, S., & Sánchez-Walker, N. (2013). Differential Object Marking in child and adult Spanish heritage speakers. *Language Acquisition*, 20, 109–132. <https://doi.org/10.1080/10489223.2013.766741>
- Moro, F. (2016). Dynamics of Ambon Malay: Comparing Ambon and the Netherlands (Unpublished doctoral dissertation). Radboud University, Nijmegen.
- Moro, F., & Klamer, M. (2015). Give-constructions in Heritage Ambon Malay in the Netherlands. *Journal of Language Contact*, 8(2), 263–198. <https://doi.org/10.1163/19552629-00802004>
- Mota, C. A. (1997). The maintenance of Portuguese as a heritage language in Winnipeg (Unpublished MA thesis). University of Manitoba, Winnipeg.
- Müller, N., & Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, 4(1), 1–21. <https://doi.org/10.1017/S1366728901000116>
- Muysken, P. (1997). Older studies on migrant languages in the United States: A model for European research? In A. Melis & P. Weber (Eds.), *Festschrift 20 jaar Onderzoekscentrum voor Meertaligheid* (pp. 238–247). Vrije Universiteit Brussel.
- Muysken, P. (2000). *Bilingual speech. A typology of code-mixing*. Cambridge: Cambridge University Press.
- Muysken, P. (2002). Uchumataqu: Research in progress on the Bolivian Altiplano. *International Journal on Multicultural Societies*, 4, 235–237.
- Muysken, P. (2008). *Functional Categories*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511755026>
- Muysken, P. (2010a). Scenarios for language contact. In R. Hickey (Ed.), *Handbook of language contact* (pp. 265–281). Oxford: Blackwell. <https://doi.org/10.1002/9781444318159.ch13>

- Muysken, P. (2010b). The demise and attempted revival of Uchumataqu (Uru). In J. A. Flores Farfán, & F. Ramallo (Eds.), *New perspectives on endangered languages. Bridging gaps between sociolinguistics, documentation and language revitalization* (pp. 93–118). Amsterdam: John Benjamins. <https://doi.org/10.1075/clu.1.06muy>
- Muysken, P. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition*, 16(4), 709–730. <https://doi.org/10.1017/S1366728912000727>
- Muysken, P. (2016). From Colombo to Athens: An areal perspective on bilingual compound verbs. *Languages. Mixed verbs and linguistic creativity in bi/multilingual communities. Languages*, 1(1), 2. <https://doi.org/10.3390/languages1010002>
- Muysken, P., Kook, H., & P. Vedder. (1996). Papiamentu/Dutch code-switching in bilingual parent-child reading. *Applied Psycholinguistics*, 17, 485–505. <https://doi.org/10.1017/S0142716400008213>
- Myers-Scotton, C. (1993). *Duelling languages: Grammatical structure in codeswitching*. Oxford: Clarendon Press.
- Myers-Scotton, C. (2006). Natural code-switching knocks on the laboratory door. *Bilingualism: Language and Cognition*, 9, 203–212. <https://doi.org/10.1017/S1366728906002549>
- Nagy, N., & Reynolds, B. (1997). Optimality theory and variable word-final deletion in Faetar. *Language Variation and Change*, 9(1), 37–56. <https://doi.org/10.1017/S0954394500001782>
- Nagy, N. (2011). A multilingual corpus to explore geographic variation. *Rassegna Italiana di Linguistica Applicata*, 43(1–2), 65–84.
- Nagy, N. (2015). A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua*, 164, 309–327. <https://doi.org/10.1016/j.lingua.2014.04.012>
- Nagy, N., Chociey, J., & Hoffman, M. F. (2014). Analyzing ethnic orientation in the quantitative sociolinguistic paradigm. *Language & Communication*, 35, 9–26. <https://doi.org/10.1016/j.langcom.2013.11.002>
- Nichols, J. M. (1992). *Linguistic diversity in space and time*. Chicago, IL: University of Chicago Press. <https://doi.org/10.7208/chicago/9780226580593.001.0001>
- Nortier, J. (1990). *Dutch-Moroccan Arabic code-switching among young Moroccans in the Netherlands*, Dordrecht: Foris.
- Nortier, J. (2012). “The more languages, the more English?": A Dutch perspective. In A. De Houwer & A. Wilton (Eds.), *English in Europe today. Sociocultural and educational perspectives* (pp. 113–132). Amsterdam: John Benjamins.
- O’Grady, W. (2011). Interfaces and processing. *Linguistic Approaches to Bilingualism*, 1, 63–66. <https://doi.org/10.1075/lab.1.1.08ogr>
- O’Grady, W., Schafer, A., Perla, J., Lee, On-Soon, & Wieting, J. (2009). A psycholinguistic tool for the assessment of language loss: The HALA project. *Language Documentation & Conservation*, 3(1), 100–112.
- O’Grady, W., Kwak, H.-Y., Lee, O.-S., & Lee, M. (2011). An emergentist perspective on Heritage Language acquisition. *Studies in Second Language Acquisition*, 33(2), 223–245. <https://doi.org/10.1017/S0272263110000744>
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524537>
- Oh, J., S. Jun, Knightly, L., & Au, T. (2003). Holding on to childhood language memory. *Cognition*, 86(3), B53–64. [https://doi.org/10.1016/S0010-0277\(02\)00175-0](https://doi.org/10.1016/S0010-0277(02)00175-0)
- Oller, D. K., & Eilers, R. E. (Eds.). (2002). *Language and literacy in bilingual children*. Clevedon, UK: Multilingual Matters.

- Onar Valk, P. (2015). Transformation in Dutch Turkish subordination?: Converging evidence of change regarding finiteness and word order in complex clauses (Unpublished doctoral dissertation). Tilburg University.
- Oostindie, G. (1995). Taal, geschiedenis en Antilliaanse identiteiten. In H. E. Coomans, M. Coomans-Eustatia, & P. Prins (Eds.) *Caraïbische cadens. Liber amicorum opgedragen aan de Gevolmachtigde Minister van de Nederlandse Antillen Edsel A.V. [Papy] Jesurun* (pp. 148–153). Bloemendaal: Stichting Libri Antilliani.
- Oostindie, G. (2008). Migration and its legacies in the Dutch colonial world. In G. Oostindie (Ed.), *Dutch colonialism, migration and cultural heritage* (pp. 1–23). Leiden: KITLV Press. [https://doi.org/10.1163/9789004253889\\_002](https://doi.org/10.1163/9789004253889_002)
- Orellana, M. F. Faulstich, & D'warte, J. (2010). A different kind of head start: Response to the early reading panel report. *Educational Researcher*, 39(4), 295–300. <https://doi.org/10.3102/0013189X10369829>
- Oriyama, K. (2017). Reforming Australian policy for Chinese, Indonesian, Japanese, and Korean heritage languages: Examples from the Japanese community. In O. Kagan, M. Carreira, & C. Hitchins Chik (Eds.), *The Routledge handbook of heritage language education: From innovation to program building* (pp. 265–281). London: Routledge.
- Ortega, L. (2014). Ways forward for a bi/multilingual turn in SLA. In S. May (ed.), *The multilingual turn: Implications for SLA, TESOL, and bilingual education* (pp. 32–53). New York, NY: Routledge.
- Ortega, L. (2016). Multi-competence in second language acquisition: Inroads into the mainstream? In V. Cook & L. Wei (Eds.), *The Cambridge handbook of linguistic multi-competence* (pp. 50–76). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781107425965.003>
- Ortega, L., & Byrnes, H. (2010). The longitudinal study of advanced L2 capacities. *Modern Language Journal*, 94(1), 157–158. <https://doi.org/10.1111/j.1540-4781.2009.00998.x>
- Osch, B. van, Hulk, A., Sleeman, P., & Aalberse, S. (2015). Heritage speakers' knowledge of Spanish mood: Interfaces or variability. Presented at Bilingualism in the Hispanic and Lusophone World (BHL), Leiden University, 15 January.
- Osch, B. van, & Sleeman, P. (2018). Spanish heritage speakers in the Netherlands: Linguistic patterns in the judgment and production of mood. *International Journal of Bilingualism*, 22(5), 513–529.
- Osch, B. van, Hulk, A., Sleeman, P., & Irizarri van Suchtelen, P. (2014). Gender agreement in interface contexts in the oral production of heritage speakers of Spanish in the Netherlands. In A. Auer & B. Köhnlein (Eds.), *Linguistics in the Netherlands 2014* (Vol. 31., pp. 93–106). Amsterdam: John Benjamins.
- Pap, L. (1949). *Portuguese American speech. An outline of speech conditions among Portuguese immigrants in New England and elsewhere in the United States*. New York, NY: Columbia University, King's Crown Press. Distributed by University Microfilms. Ann Arbor, MI, 1979.
- Paradis, C. & Prunet, J.-F. (1991). *Phonetics and phonology, 2: The special status of coronals: Internal and external evidence*. New York, NY: Academic Press.
- Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children. *Studies in Second Language Acquisition*, 18(1), 1–25. <https://doi.org/10.1017/S0272263100014662>
- Paradis, M. (2007). L1 attrition features predicted by a neurolinguistic theory of bilingualism. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language attrition. Theoretical perspectives* (pp. 121–133). Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.33.09par>

- Pauwels, A. (2005). Maintaining the community language in Australia: Challenges and roles for families. *International Journal of Bilingual Education and Bilingualism*, 8(2–3), 24–131.
- Pavlenko, A., & Jarvis, S. (2002). Bidirectional transfer. *Applied Linguistics*, 23(2), 190–214. <https://doi.org/10.1093/applin/23.2.190>
- Peeters, D., Dijkstra, A., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of Memory and Language*, 68, 315–332. <https://doi.org/10.1016/j.jml.2012.12.003>
- Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam: John Benjamins. <https://doi.org/10.1075/swll.11.14per>
- Pfaff, C. (1979). Constraints on language-mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language*, 55, 291–318. <https://doi.org/10.2307/412586>
- Pfaff, C. (1994). Early bilingual development of Turkish children in Berlin. In G. Extra & L. Verhoeven (Eds.), *The cross-linguistic study of bilingual development* (pp. 75–97). Amsterdam: North-Holland.
- Phillips, C. (2013). Some arguments and non-arguments for reductionist accounts of syntactic phenomena. *Language and Cognitive Processes*, 28, 156–187. <https://doi.org/10.1080/01690965.2010.530960>
- Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *PNAS*, 108(9), 3526–3529. <https://doi.org/10.1073/pnas.1012551108>
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–190; discussion 190–226. <https://doi.org/10.1017/S0140525X04000056>
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic persistence in written language production. *Journal of Memory & Language*, 39, 633–651. <https://doi.org/10.1006/jmla.1998.2592>
- Pierce, M., Boas, H. C., & Roesch, K. A. (2015). The history of front rounded vowels in New Braunfels German. In J. Bondi Johannessen & J. C. Salmons (Eds.), *Germanic heritage languages in North America* (pp. 117–132). Amsterdam: John Benjamins. <https://doi.org/10.1075/silv.18.05pie>
- Pires, A., & Rothman, J. (2009). Disentangling sources of incomplete acquisition: An explanation for competence divergence across heritage grammars. *International Journal of Bilingualism*, 13(2), 11–238. <https://doi.org/10.1177/1367006909339806>
- Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence 2-year-olds' bilingual proficiency. *Child Development*, 82, 1834–1849. <https://doi.org/10.1111/j.1467-8624.2011.01660.x>
- Polinsky, M. (2006). Incomplete acquisition: American Russian. *Journal of Slavic Linguistics*, 14, 191–262.
- Polinsky, M. (2008a). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal*, 6(1), 40–71.
- Polinsky, M. (2008b). Heritage language narratives. In D. M. Brinton, O. Kagan & S. Baucus (Eds.), *Heritage language education: A new field emerging* (pp. 149–164). London: Routledge.
- Polinsky, M. (2008c). Relative clauses in heritage Russian: Fossilization or divergent grammar? *Formal Approaches to Slavic Linguistics*, 16, 333–358. Ann Arbor, MI: Michigan Slavic Publications.

- Polinsky, M. (2011). Reanalysis in adult heritage language. *Studies in Second Language Acquisition*, 33(2), 305–328. <https://doi.org/10.1017/S027226311000077X>
- Polinsky, M. (2018a). *Heritage languages and their speakers*. Cambridge: Cambridge University Press.
- Polinsky, M. (2018b). Bilingual children and adult heritage speakers: The range of comparison. *International Journal of Bilingualism*, 22(5), 547–563.
- Polinsky, M., & Kagan, O. (2007). Heritage languages: In the ‘wild’ and in the classroom. *Language and Linguistics Compass*, 1(5), 368–395. <https://doi.org/10.1111/j.1749-818X.2007.00022.x>
- Poplack, S. (1980). Sometimes I’ll start a sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a typology of code-switching. *Linguistics*, 18, 581–618. <https://doi.org/10.1515/ling.1980.18.7-8.581>
- Poplack, S. (2017). *Borrowing. Loanwords in the speech community and in the Grammar*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780190256388.003.0004>
- Poplack, S., Sankoff, D., & Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, 26(1), 47–104. <https://doi.org/10.1515/ling.1988.26.1.47>
- Poplack, S., Zentz, L., & Dion, N. (2012). Phrase-final prepositions in Quebec French: An empirical study of contact, code-switching and resistance to convergence. *Bilingualism: Language and Cognition*, 15(2), 203–225. <https://doi.org/10.1017/S1366728911000204>
- Postma, J. (1974). The origin of African slaves: The Dutch activities on the Guinea coast, 1675–1795. In S. L. Engerman & E. D. Genovese (Eds.), *Race and slavery in the Western hemisphere: Quantitative studies* (pp. 33–49). Princeton, NJ: Princeton University Press.
- Poullisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, 15, 36–57. <https://doi.org/10.1093/applin/15.1.36>
- Prada Pérez de, A., & Pascual y Cabo, D. (2012). Interface heritage speech across proficiencies: Unaccusativity, focus, and subject position in Spanish. In K. Geeslin & M. Díaz-Campos (Eds.), *Selected proceedings from the Hispanic Linguistics Symposium 2010* (pp. 308–318). Somerville, MA: Cascadilla Proceedings Project.
- Prince, A., & Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar*. Oxford: Blackwell. <https://doi.org/10.1002/9780470759400>
- Prince, J. Dyneley. (1913). A text in Jersey Dutch. *Tijdschrift voor Nederlandsche Taal en Letterkunde*, 32, 306–312.
- Putman, J. J. (1849). *Proeve eener Hollandsche spraakkunst, ten gebruike der algemeene armen-school*. Curaçao: Author’s edition.
- Putnam, M. (2019). The (in)stability of grammars. *Studies in Second Language Acquisition*, 41(2), 275–278.
- Putnam, M., & Sánchez, L. (2013). What’s so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*, 3(4), 476–506. <https://doi.org/10.1075/lab.3.4.04put>
- Putnam, M., Kupisch, T., Pascual y Cabo, D. (2018). Heritage grammars across the lifespan. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 251–280). Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.54.12put>
- Putnam, M., Perez-Cortes, S., & Sánchez, L. (2019). Language attrition and the Feature Reassembly Hypothesis. In M. Schmid & B. Köpke (Eds.), *Oxford handbook of language attrition* (pp. 18–24). Oxford: Oxford University Press.

- Putte, F. van, & van Putte-de Windt, I. (2005). *Dikshonario Papiamentu-Hulandes; Hulandes-Papiamentu*. Zutphen: Walburg Pers.
- Putte, F. van. (1999). *Dede pikiña ku su bisiaña: Papiamentu-Nederlands en de onverwerkt verleden tijd*. Zutphen: Walburg Pers.
- Queen, R. (2012). Turkish German bilinguals and their intonation: Triangulating evidence about contact-induced language change. *Language*, 88(4), 791–816. <https://doi.org/10.1353/lan.2012.0078>
- Quint, N. (2000). *Le capverdien: Origines et devenir d'une languemétisse*. Paris: l'Harmattan.
- Ratzlaff, B. (2008). *Dikshonario Papiamentu – Ingles Dikshonario*. Bonaire: TWA Dictionary Foundation.
- Rehbein, J., Herkenrath, A., & Karakoc, B. (2009a). Corpus Rehbein-SKOBI (sprachliche konnektivität bei bilingual türkisch-deutsch aufwachsenden kindern und jugendlichen).
- Rehbein, J., Herkenrath, A., & Karakoç, B. (2009b). Turkish in Germany on contact-induced language change of an immigrant language in the multilingual landscape of europe. *STUF-Language Typology and Universals Sprachtypologie und Universalienforschung*, 62(3), 171–204.
- Reinhart, T. (1983). *Anaphora and semantic interpretation*. London: Croom Helm.
- Reinhart, T. (2006). *Interface strategies: Reference-set computation*. Cambridge, MA: The MIT Press.
- Reuland, E. J. (2011). *Anaphora and language design*. Cambridge, MA: The MIT Press.
- Rickford, J., & Price, M. (2013). Girlz II women: Age-grading, language change and stylistic variation. *Journal of Sociolinguistics*, 17(2), 143–179. <https://doi.org/10.1111/josl.12017>
- Riehl, C. (2015). Language attrition, language contact and the concept of relic variety: The case of Barossa German. *International Journal of the Sociology of Language*, 236, 261–293.
- Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of non-native speaker conversations. *Discourse Processes*, 14(4), 423–441. <https://doi.org/10.1080/01638539109544795>
- Ritter, A. (2014). Bilingual speech in Russian-German language contact: Findings based on an analysis of natural conversations in Russian-speaking immigrant families in Germany. In H. Paulasto, L. Meriläinen, H. Riionheimo, & M. Kok (Eds.), *Language contacts at the crossroads of disciplines* (pp. 251–279). Newcastle upon Tyne: Cambridge Scholars.
- Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science*, 8, 162–172. <https://doi.org/10.1111/j.1467-7687.2005.00403.x>
- Robertson, I. (1983). The Dutch linguistic legacy and the Guyana/Venezuela border question. *Boletín de Estudios Latinoamericanos y del Caribe*, 34, 75–97.
- Robins, R. H., & Uhlenbeck, E. M. (Eds.). (1991). *Endangered languages*. Oxford: Berg Publishers.
- Rooij, V. de. (1996). Cohesion through contrast: Discourse structure in Shaba Swahili/French conversations (Unpublished doctoral dissertation). University of Amsterdam.
- Rothman, J. (2007). Heritage speaker competence differences, language change and input type: Inflected infinitives in Heritage Brazilian Portuguese. *International Journal of Bilingualism*, 11(4), 359–389. <https://doi.org/10.1177/13670069070110040201>
- Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, 13(2), 155–163. <https://doi.org/10.1177/1367006909339814>
- Rubin, J. (1962). Bilingualism in Paraguay. *Anthropological Linguistics*, 4, 52–58.

- Rupert, L. (2012). *Creolization and contraband: Curaçao in the early modern Atlantic world*. Athens: University of Georgia Press.
- Sabino, R. (2012). *Language contact in the Danish West Indies: Giving Jack his jacket*. Leiden: Brill.
- Şahin, H. (2015). Cross-linguistic influences: Dutch in contact with Papiamentu & Turkish (Unpublished doctoral dissertation). Radboud University, Nijmegen.
- Sakuma, T. (2013). Performing “Japanese” in Brazil: Language ideologies in a Japanese Brazilian cultural association in Sao Paulo, Brazil (Unpublished doctoral dissertation). University of Texas at Austin.
- Salmons, J. C. (1990). Bilingual Discourse Marking: Codeswitching, borrowing and convergence in some German-American Dialects. *Linguistics* 28, 453–480.
- Sanchez, T. (2005). Constraints on structural borrowing in a multilingual contact situation (Unpublished doctoral dissertation). University of Pennsylvania.
- Sarhimaa, A. (1999). *Syntactic transfer, contact-induced change, and the evolution of bilingual mixed codes: Focus on Karelian-Russian language alternation* (Studia Fennica Linguistica 9). Helsinki: Finnish Literature Society.
- Sauer, M. (2010). Mediennutzungsmotive türkeistämmiger Migranten in Deutschland. *Publizistik*, 55, 55–76. <https://doi.org/10.1007/s11616-010-0074-5>
- Schaufeli, A. (1991). Turkish in an immigrant setting, A comparative study of the first language of monolingual and bilingual immigrant children (Unpublished doctoral dissertation). University of Amsterdam.
- Scheele, A. F., Leseman, P. P. M., & Mayo, A. Y. (2010). The home language environment of mono- and bilingual children and their language proficiency. *Applied Psycholinguistics*, 31, 117–140. <https://doi.org/10.1017/S0142716409990191>
- Schmid, M. S. (2002). *First language attrition, use and maintenance: The case of German Jews in Anglophone countries*. Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.24>
- Schmid, M. S. (2007). *Language attrition: Theoretical perspectives*. Amsterdam: John Benjamins.
- Schmid, M. S. (2009). On L1 attrition and the linguistic system. *Eurosla Yearbook*, 9, 212–244. <https://doi.org/10.1075/eurosla.9.11sch>
- Schmid, M. S. (2011). *Language attrition*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511852046>
- Schmid, M. S. (2012). The impact of age and exposure on forgetting and retention of the birth language in international adoptees: A perspective from Holocaust survivors. *Linguistic Approaches to Bilingualism*, 2(2), 177–208. <https://doi.org/10.1075/lab.2.2.03sch>
- Schmid, M. S., & Dusseldorp, E. (2010). Innovative and quantitative methods for bilingualism research. *Second Language Research*, 26(1), 5–11. <https://doi.org/10.1177/0267658309337616>
- Schmidt, A. (1985). *Young people’s Dyirbal: An example of language death from Australia*. Cambridge: Cambridge University Press.
- Schoenmakers-Klein Gunnewiek, M. (1997). Dutch language loss and the Conceptual Hypothesis. In J. Klatter-Folmer & S. Kroon (Eds.), *Dutch overseas: Studies in maintenance and loss of Dutch as an immigrant language* (pp. 99–119). Tilburg: Tilburg University Press.
- Schoenmakers-Klein Gunnewiek, M. (1998). *Taalverlies door taalcontact? Een onderzoek bij Portugese migranten* (Unpublished doctoral dissertation). Tilburg University.
- Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, 56, 153–171. <https://doi.org/10.1016/j.jml.2006.10.002>

- Schüpbach, D. (2009). Language transmission revisited: Family type, linguistic environment and language attitudes. *International Journal of Bilingual Education and Bilingualism*, 12(1), 15–30. <https://doi.org/10.1080/13670050802149499>
- Schwartz, B. D. (2006). What's left in early L2 architecture. *Second Language*, 5, 3–26.
- Schwartz, M., & Verschik, A. (2013). Achieving success in family language policy: Parents, children and educators in interaction. In M. Schwartz & A. Verschik (Eds.), *Successful family language policy. Parents, children and educators in interaction* (pp. 1–23). Dordrecht: Springer. [https://doi.org/10.1007/978-94-007-7753-8\\_1](https://doi.org/10.1007/978-94-007-7753-8_1)
- Schwartz, R. G., & Terrell, B. T. (1983). The role of input frequency in early lexical acquisition. *Journal of Child Language*, 9, 57–64.
- Seaman, P. D. (1972). *Modern Greek and American English*. The Hague: Mouton.
- Sebba, M. (1993). *London Jamaican: Language systems in interaction*. London: Longman.
- Sevinç, Y., & Backus, A. (2017). Anxiety, language use and linguistic competence in an immigrant context: a vicious circle?. *International Journal of Bilingual Education and Bilingualism*. <https://doi.org/10.1080/13670050.2017.1306021>
- Sherkina-Lieber, M. (2015). Tense, aspect, and agreement in heritage Labrador Inuttitut: Do receptive bilinguals understand functional morphology? *Linguistic Approaches to Bilingualism*, 5(1), 30–61. <https://doi.org/10.1075/lab.5.1.02sh>
- Shi, Menghui. (2011). A study of mood particle in the Suzhou Dialect in Qing dynasty. *The Research of Wu Dialect*, 6, 256–269.
- Shin, S. J. (2002). Birth order and the language experience of bilingual children. *Tesol Quarterly* 36(1), 103–113. <https://doi.org/10.2307/3588366>
- Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.
- Silva-Corvalán, C. (2014). *Bilingual language acquisition: Spanish and English in the first six years*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139162531>
- Silva-Corvalán, C. (2018). Simultaneous bilingualism: Early developments, incomplete later outcomes?. *International Journal of Bilingualism*, 22(5), 497–512.
- Silva-Corvalán, C., & Treffers-Daller, J. (Eds.). (2016). *Language dominance in bilinguals: Issues of measurement and operationalization*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781107375345>
- Skutnabb-Kangas, T. (2002). When languages disappear, are bilingual education and human rights a cure? Two scenarios. In Li Wei, J.-M. Dewaele, & A. Housen (Eds.), *Opportunities and challenges of bilingualism* (pp. 45–67). Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110852004.45>
- Skutnabb-Kangas, T., Phillipson, R., & Rannut, M. (1994). *Linguistic human rights: Overcoming linguistic discrimination*. Berlin: Walter de Gruyter.
- Smits, C. (1996). Disintegration of inflection: The case of Iowa Dutch (Unpublished doctoral dissertation). Vrije Universiteit, Amsterdam.
- Smits, E., Martensen, H., Dijkstra, A., & Sandra, D. (2006). Naming interlingual homographs: Variable competition and the role of the decision system. *Bilingualism: Language and Cognition*, 9, 281–297. <https://doi.org/10.1017/S136672890600263X>
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1–33. <https://doi.org/10.1075/lab.1.1.01sor>
- Spolsky, B. (2009). *Language management*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511626470>



- Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, 43, 155–167.  
<https://doi.org/10.3758/s13428-010-0039-7>
- Stansfield, C., & Hansen, J. (1983). Field dependence-independence as a variable in second language Cloze Test performance. *Tesol Quarterly*, 17(1), 29–38.  
<https://doi.org/10.2307/3586422>
- Steensig, J. (2001). Some notes on the use of conversation analysis in the study of bilingual interaction. In J. N. Jørgensen (Ed.), *Multilingual behavior in youth groups: Scandinavian studies in the simultaneous use of two or more languages in group conversations among children and adolescents* (pp. 55–79). Copenhagen: Copenhagen Studies of Bilingualism.
- Swadesh, M. (1971). *The origin and diversification of language*, J. Sherzer (Ed.), *Post mortem*. Chicago, IL: Aldine.
- Tahitu, B. (1989). *Melaju sini, het Maleis van Molukse jongeren in Nederland* (Unpublished doctoral dissertation). Leiden University:
- Tarone, E., & Bigelow, M. (2005). Impact of literacy on oral language processing: Implications for SLA research. *Annual Review of Applied Linguistics*, 25, 77–97.  
<https://doi.org/10.1017/S0267190505000048>
- Tarone, E., Bigelow, M., & Hansen, K. (2009). *Literacy and second language oracy*. Oxford: Oxford University Press.
- Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley, CA: University of California Press.
- Toribio, A. J. (2004). Convergence as an optimization strategy in bilingual speech: Evidence from code-switching. *Bilingualism: Language and Cognition*, 7(2), 165–173.  
<https://doi.org/10.1017/S1366728904001476>
- Torres Cacoullous, R., & Travis, C. E. (2015). Gauging convergence on the ground: Code-switching in the community. *International Journal of Bilingualism*, 19(4), 365–386.  
<https://doi.org/10.1177/1367006913516046>
- Torres Cacoullous, R. (2000). *Grammaticization, synchronic variation, and language contact. A study of Spanish progressive -ndo constructions*. Amsterdam: John Benjamins.  
<https://doi.org/10.1075/slcs.52>
- Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: The MIT Press.
- Treffers-Daller, J. (1994). *Mixing two languages: French-Dutch contact in a comparative perspective*. Berlin: Walter de Gruyter. <https://doi.org/10.1515/9783110882230>
- Treffers-Daller, J. (2011). Operationalising and measuring language dominance. *International Journal of Bilingualism*, 15(2), 147–163. <https://doi.org/10.1177/1367006910381186>
- Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Tsimpli, I. M., & Sorace, A. (2006). Differentiating interfaces: L2 performance in syntax/semantics and syntax/discourse phenomena. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *BUCLD 30, Proceedings of the 30th annual Boston University Conference on Language Development* (pp. 653–664). Somerville, MA: Cascadilla Press.
- Türker, E. (2000). *Turkish-Norwegian codeswitching. Evidence from intermediate and second generation Turkish immigrants in Norway*. Oslo: Unipub Forlag.
- UCLA Steering Committee. (2001). *Heritage language research priorities conference report*. Los Angeles, CA: University of California. Retrieved from <<http://www.cal.org/heritage/involved/hlprioritiesconf00.pdf>> (December 26, 2007).

- Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, 4, 105–122. <https://doi.org/10.1017/S1366728901000220>
- UNESCO. (1953). *The use of vernacular languages in education*. Paris: UNESCO.
- Unsworth, S. (2013). Current issues in multilingual first language acquisition. *Annual Review of Applied Linguistics*, 33, 21–50. <https://doi.org/10.1017/S0267190513000044>
- Valdés, G. (1995). The teaching of minority languages as academic subjects: Pedagogical and theoretical challenges. *Modern Language Journal*, 79(3), 299–328. <https://doi.org/10.1111/j.1540-4781.1995.tb01106.x>
- Valdés, G. (2000). *Spanish for native speakers: AATSP professional development series. Handbook for teachers K-16* (Vol. 1). New York, NY: Harcourt College Publishers.
- Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language*, 64, 88–107. <https://doi.org/10.1016/j.jml.2010.08.006>
- Van Coetsem, F. (1988). *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris. <https://doi.org/10.1515/9783110884869>
- Van Coetsem, F. (2000). *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Van Deusen-Scholl, N. (1998). Heritage language instruction: Issues and challenges. *AILA Newsletter*, 1, 12–14.
- Van Deusen-Scholl, N. (2003). Toward a definition of heritage language: Sociopolitical and pedagogical considerations. *Journal of Language, Identity, and Education*, 2(3), 211–30. [https://doi.org/10.1207/S15327701JLIE0203\\_4](https://doi.org/10.1207/S15327701JLIE0203_4)
- Vasilyeva, M., Waterfall, H., Gámez, P. B., Gómez, L. E., Bowers, E., & Shimpi, P. (2010). Cross-linguistic syntactic priming in bilingual children. *Journal of Child Language*, 37, 1047–1064.
- Vazsonyi, A. (1995). *Tul a kecegardan/Beyond castle garden. An American Hungarian dictionary of the Calumet region*, ed. by M. Kontra. Budapest: Teleki László Alapítvány.
- Vedder, P., Muysken, P., & Kook, H. (1996). Language choice and functional differentiation of languages in bilingual parent-child reading. *Applied Psycholinguistics*, 17, 461–484. <https://doi.org/10.1017/S0142716400008201>
- Veenman, J. (1994). *The social integration of Moluccans*. Lelystad: Koninklijke Vermande.
- Veenman, J. (2001). *Molukse jongeren in Nederland. Integratie met de rem erop*. Assen: Koninklijke Van Gorcum.
- Vellard, J. (1949). Contribution à l'étude des indiens Urus ou Kot'suns. *Travaux de l'Institut Français des Études Andines I* (pp. 145–209). Paris-Lima.
- Verhallen, M., & Schoonen, R. (1998). Lexical knowledge in L1 and L2 of third and fifth graders. *Applied Linguistics*, 19(4), 452–470. <https://doi.org/10.1093/applin/19.4.452>
- Verhoef, K., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition*, 110(1), 84–99. <https://doi.org/10.1016/j.cognition.2008.10.013>
- Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, 4(4), 314–330. [https://doi.org/10.1207/S1532799XSSR0404\\_4](https://doi.org/10.1207/S1532799XSSR0404_4)
- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to acquisition and frequency of input. *Applied Psycholinguistics*, 22(2), 217–234. <https://doi.org/10.1017/S0142716401002041>
- Vermes, G., & Boutet, J. (Eds.). (1987). *France, pays multilingue. Pratiques de langues en France*. Paris: l'Harmattan.

- Verschik, A. (2014). Bare participle forms in the speech of Lithuanian Yiddish heritage speakers: Multiple causation. *International Journal of the Sociology of Language*, 226, 213–235.
- Versteegh, K. (2014). *The Arabic language* (2nd ed.). Edinburgh: Edinburgh University Press.
- Vertovec, S. (2001). Transnationalism and identity. *Journal of Ethnic and Migration Studies*, 27(4), 573–582. <https://doi.org/10.1080/13691830120090386>
- Vertovec, S. (2007). Super-diversity and its implications. *Ethnic and Racial Studies*, 29(6), 1024–54. <https://doi.org/10.1080/01419870701599465>
- Vertovec, S. (2010). Towards post-multiculturalism? Changing communities, conditions and contexts of diversity. *International Social Science Journal*, 199, 83–95. <https://doi.org/10.1111/j.1468-2451.2010.01749.x>
- Vertovec, S., & Cohen, R. (Eds.). (1999). *Migration, diasporas and transnationalism*. Cheltenham: Edward Elgar Publishing.
- Villierius, S. (2019). Development of Surinamese Javanese: Language contact and change in a multilingual context (Unpublished doctoral dissertation). Radboud University, Nijmegen.
- Villierius, S., Moro, F., & Klamer, M. (To appear). Encoding transfer events in Surinamese Javanese. *Journal of Language Contact*.
- Weber, K., & Indefrey, P. (2009). Shared syntactic information in German-English bilinguals: An fMRI repetition suppression study. *Neuroimage*, 46, 1164–1172. <https://doi.org/10.1016/j.neuroimage.2009.03.040>
- Werker, J. F., Gilbert, J. H. V., Humphreys, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 52, 349–355. <https://doi.org/10.2307/1129249>
- Wertheim, S. (2003). Linguistic purism, language shift, and contact-induced change in Tatar (Unpublished doctoral dissertation). University of California, Berkeley.
- Wong Fillmore, L. (1991). When learning a second language means losing the first. *Early Childhood Research Quarterly*, 6(3), 323–346. [https://doi.org/10.1016/S0885-2006\(05\)80059-6](https://doi.org/10.1016/S0885-2006(05)80059-6)
- Wong-Fillmore, L. (2011). An ecological perspective on intergenerational language transmission. In M. E. Romero-Little, S. J. Ortiz, & T. L. McCarty (Eds.), *Indigenous languages across the generations – Strengthening families and communities* (pp. 19–48). Tempe, AZ: Arizona State University Press.
- Wood, R. E. (1970). Papiamentu: Dutch contributions (Unpublished doctoral dissertation). Indiana University.
- Wood, R. E. (1971). The English loanwords in Papiamentu. *Nieuwe West-Indische Gids*, 48, 173–189. <https://doi.org/10.1163/22134360-90002217>
- Wood, R. E. (1972). Dutch syntactic loans in Papiamentu. *Revue des Langues Vivantes*, 38(6), 635–647.
- Woolard, K. A. (1989). Language convergence and language death as social processes. In N. Dorian (Ed.) *Investigating obsolescence: Studies in language contraction and death* (pp. 355–367). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620997.028>
- Yagmur, K., & Kroon, S. (2003). Ethnolinguistic vitality perceptions and language revitalisation in Bashkortostan. *Journal of Multilingual and Multicultural Development*, 24(4), 319–336. <https://doi.org/10.1080/01434630308666504>
- Yagmur, K., & Akinci, M.-A. (2003). Language use, choice, maintenance and ethnolinguistic vitality of Turkish speakers in France: Intergenerational differences. *International Journal of the Sociology of Language*, 164, 107–128.

- Yeni-Komshian, G., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in first and second languages of Korean–English bilinguals. *Bilingualism: Language and Cognition*, 3, 131–149. <https://doi.org/10.1017/S1366728900000225>
- Yilmaz, G. (2013). Bilingual language development among the first generation Turkish immigrants in the Netherlands (Unpublished doctoral dissertation). University of Groningen.
- Yip, V., & S. Matthews. (2007). *The Bilingual Child: Early Development and Language Contact*. Cambridge: Cambridge University Press.
- Zabrodskaja, A. (2009). Russian–Estonian language contacts: Grammatical aspects of language use and change (Unpublished doctoral dissertation). Tallin University, Estonia.
- Zeeuw, M. de, Verhoeven, L., & Schreuder, R. (2012). Morphological family size effects in young first and second language learners: Evidence of cross-language semantic activation in visual word recognition. *Language Learning*, 62(1), 68–92. <https://doi.org/10.1111/j.1467-9922.2011.00691.x>
- Zeeuw, M. de, Schreuder, R., & Verhoeven, L. (2013). Processing of regular and irregular past-tense verb forms in first and second language reading acquisition. *Language Learning*, 63(4), 740–765. <https://doi.org/10.1111/lang.12023>
- Zeeuw, M. de, Schreuder, R., & Verhoeven, L. (2015). Lexical processing of nominal compounds in first- and second-language learners across primary grades. *Writing Systems Research*, 7(2), 133–156. <https://doi.org/10.1080/17586801.2014.926806>
- Zenner, E., Speelman, D., & Geeraerts, D. (2013). Macro and micro perspectives on the distribution of English in Dutch. A quantitative usage-based analysis of job ads. *Linguistics*, 51(5), 1019–1064. <https://doi.org/10.1515/ling-2013-0036>
- Zipf, G. K. (1949). *Human behavior and the principle of least effort. An introduction to human ecology*. Cambridge, MA: Addison-Wesley Press.



# Language index

## A

Anishinaabemdaa 228  
Arabic 21, 29, 41, 54, 155, 218,  
231, 235  
Egyptian 155  
Lebanese 29  
Moroccan 21, 41, 235  
Palestine 155  
Aymara 63, 238–239

## B

Berber 16, 218, 230–231, 235  
Breton 6, 61, 226

## C

Chinese 1, 2, 8, 16–17, 29,  
31, 40–41, 53, 113, 114, 116,  
122–123, 128–130, 133, 150,  
155, 157–159, 200, 228,  
232–234  
Cantonese 157, 159  
Mandarin 1, 133, 150  
Chinuk Wawa 228  
Choctaw 228  
Creole 21, 24, 27–28, 54, 136,  
203–206, 213, 247–248, 250,  
253, 255–257, 260  
Berbice Dutch 27, 206  
London Jamaican 22, 203  
Croatian 40, 95, 228  
Czech 34, 228

## D

Danish 14, 145–146, 203  
Argentina 146  
Denaakkè Athabascan 228  
Dutch 2, 3, 4, 7, 13, 16–17,  
24–8, 30, 32, 39, 40, 48, 53,  
55, 57, 61, 64, 70–71, 73–80,  
83–84, 90, 95, 98–99, 101,  
102–103, 106, 123, 128–129,  
143–146, 156–157, 159, 163,

169, 179–180, 190, 193, 198,  
201, 203–222, 235, 240, 245  
Iowa 30  
New Jersey 30  
Pennsylvania 30  
Dyirbal 60–1, 145, 237–239

## F

Finnish 13, 30, 35, 37–38, 44,  
83, 95  
French 4, 6, 23, 32, 33, 76,  
84, 118, 120, 126–128, 142,  
146, 169, 190, 198, 206, 226,  
228–229, 236, 257  
Frisian 3, 46, 61

## G

German 3, 6, 10, 13, 23, 30–32,  
39, 40, 44, 55, 65, 71, 76, 84,  
94–95, 97–98, 102, 107, 118,  
120, 127, 142, 145, 157, 158,  
181, 200, 205, 226, 228, 257  
Pennsylvania 6, 44  
Texas 30, 76, 97  
Greek 35, 38–40, 226

## H

Hindi 7, 8, 21, 29, 31, 153, 203,  
228  
Hungarian 29, 30, 53, 145, 242

## I

IchCinshKiin 228  
Ilokano 228  
Italian 1, 11, 23, 31–34, 40–41,  
100, 126–128, 142, 170–171,  
228

## J

Japanese 21, 29, 31, 48, 56, 57,  
165, 167, 228, 232–233, 236

## K

Korean 21, 121, 129, 133, 137,  
143, 153, 159, 165, 167, 200,  
228, 236

## M

Maya 230  
Mexicano, *see* Nahuatl

## N

Nahuatl 61, 70, 240–242  
Navajo 228  
Norwegian 3, 14, 23, 34, 50,  
70, 130, 142, 196

## P

Papiamentu 21, 26, 103, 145,  
201–223, 245, 248, 257  
Persian 228  
Portuguese 24, 27, 32–33,  
35–36, 54, 126, 133, 144,  
148, 169, 205, 210, 228, 233,  
247, 253

## Q

Quechua 61, 63, 243

## R

Russian 13, 24, 31, 34, 47, 54,  
74, 76, 80, 81, 83, 134, 136,  
142, 144, 159, 170–171, 228

## S

Samoan 228  
Sanskrit 228  
Sorbian 46, 61, 181, 227  
Spanish 4, 5, 7, 10, 16–17, 21,  
23–24, 31–2, 34, 36, 40–41,  
52, 56, 63, 95, 102, 117, 121,  
123, 127–130, 133, 138–139,  
146, 151–152, 155–156, 159,  
164, 173, 176–177, 190–192,

194, 205–207, 209–211, 217,  
228, 232–234, 236, 240–241,  
247–248, 251, 253, 257–258

Andean 63

Chilean 129

New Mexico 34

Swedish 30, 121, 127

American 30–31

## T

Tagalog 228

Tamil 58, 228

Tariana 85

Tibetan 228

Tongan 228

Turkish 9, 13–14, 23, 41,

49–51, 53, 57, 60–61, 64–65,

70–80, 83–84, 90, 93–95,

99, 101, 106, 127, 14–15,

157–158, 179, 180–181, 193,

203, 211, 218, 226, 233, 235,

240, 261

## U

Ukrainian 31, 47, 81, 142,

170–171, 228

Urdu 228

## X

Xhosa 61

## Y

Yakut 61

Yiddish 34, 80, 81, 228

# Subject index

## A

Access 50, 83, 89, 91, 96, 104, 109, 126–127, 136–137, 148, 163–164, 198, 200, 209, 225, 232, 237, 247–248, 253, 256

Acquisition 1, 3, 8–11, 15–17, 19, 21, 33, 40, 53, 87, 89, 111, 112, 116–124, 126, 129, 132, 134, 146–150, 157, 159, 161–163, 171–172, 179, 183–184, 194, 196, 198–199, 201, 225, 243–245, 247, 249–250, 252–254, 257–258, 260

first language 8, 117, 149, 198, 254

second language 19, 40, 87, 119, 134, 157, 162, 198, 254, 258

Acrolect 247, 248, 256

Activation threshold hypothesis 149, 247

Adult L2 learners, *see* Adult second language learning

Adult second language learning 9, 158, 247

Afro-Iberian 205, 247

Age of onset 2, 8–10, 111–112, 117, 119, 121, 138–139, 179, 247

Age-effects 247

Agent phrase 214, 248

Ancestral ties 2–3, 9–11, 42, 45, 47, 61–62, 229, 243, 248

Annotation 88, 95, 97, 109

Anxiety 79, 87

Attrition 1, 6, 9, 12–14, 18–20, 39, 53, 61, 63, 79, 90, 119, 123, 125, 127, 134, 136, 147–150, 153, 155, 159, 164, 171, 200, 229, 239, 247–250, 252, 254, 255, 256, 254–256, 259–260

## B

Baseline 20, 61, 108–109, 111–113, 116–119, 130–132, 138, 142–144, 147–148, 154, 164, 219, 244, 251, 258

Basic vocabulary 70, 72, 205, 248

Basilect 248, 256

Bilingualism

- additive 16, 232, 247
- replacive 259
- simultaneous 121
- subtractive 16, 129, 259, 260

Borrowing 11–13, 31, 61, 68, 74, 76, 81, 136, 143, 157, 210, 211–212, 219, 223, 240–241, 247–248, 255

cultural 70–71

Brokering 53, 248

By-phrase 215, 248

## C

California 56, 236, 251

Calque 201, 208, 248, 255

Caribbean 21, 26–27, 129, 203, 206, 213, 257

Chain migration 248

Child first language loss 249

Clitic 250

Codeswitching 12–14, 18–19, 30, 34, 37, 39, 40–41, 43, 49, 52, 55, 63–86, 90, 93–94, 98, 100, 102, 124, 143, 151, 183, 188, 192, 195–198, 200–201, 219–222, 225, 239, 240–241, 244, 248–249, 253

alternational 219, 222

insertional 14, 69, 72, 81, 219, 220, 222

Coding 154

Colonial language 26, 206, 249

Community language 3, 11, 69, 122, 125, 232, 238, 243, 244, 249, 252, 253

school 249, 253

Community norms 10, 20, 249

Complementary school 249

Comprehension 21, 159, 164, 173, 183, 185–187, 192, 194, 200, 237, 249, 252, 257

Constitutional national language 230, 249

Constraint 151, 154, 173, 175–178, 198

Contact-induced grammaticalization 158

Contact-induced language change 40–41, 196, 244

Content vocabulary, *see* Content word

Content word 70, 72, 210, 240, 249

Convergence 12–13, 15, 39, 155–158, 164, 196, 240, 249, 256

Conversation 19, 49, 52–57, 64, 76, 78, 81–82, 87–88, 90, 92–96, 98–102, 104, 106, 168, 17, 196, 218, 237, 241, 254

Copying 248

Creole 21, 24, 27–28, 54, 136, 203–206, 213, 247–248, 250, 253, 255–257, 260

Creole genesis 255

Critical period

- for acquisition 19, 163, 199, 250
- for attrition 19, 147, 250

Cross linguistic influence 250

Cross-language activation 187, 188, 190–191, 194–197, 250



- Cultural ties 7–8, 10, 30, 46  
 Cultural fluency 250  
 Cyclic migration 250
- D**  
 Declarative knowledge 199  
 Dialect(al) variation 250  
 Diaspora 18, 23, 24, 27–30, 39, 41, 243, 250  
 Diminutive 211, 215, 220, 250  
 Discourse 19, 45, 55, 61, 68–69, 72, 75–76, 81, 90, 98, 130, 152, 165–168, 185, 196–197, 201, 210, 212, 219, 221–222, 225, 240–242, 250, 254  
 Discourse marker 219, 222, 250  
 Divergence 6, 15, 57, 126, 147–148, 250, 256  
 Diversity management 22, 225–226, 250  
 Documentation 22, 89, 225–226, 229, 237, 250, 255  
 Dominance 3–5, 10–11, 47, 112, 133, 137–138, 162, 183, 196, 199, 201, 241, 243, 251–252  
 Dual  
   grammatical category 155, 159  
   language enrolment 251  
   language experiment 196–197  
   language immersion 236, 251
- E**  
 Ecological validity 20, 88, 90, 94, 98, 101–104, 189, 251  
 Education 9, 22, 41, 50, 58, 73, 113, 127, 133, 206, 208–209, 211, 219, 225–226, 231, 232–237, 242–243, 253, 255, 261  
 Entrenchment 74, 179, 251  
 Ethical concerns 20, 108, 251  
 Ethnic affiliation 251  
 Ethnicity 8, 34, 63, 83, 92, 232–233, 251  
 Ethnolects 10, 11, 32, 34, 45, 63, 251  
 Expats 251
- Experiment 88, 89, 101, 103, 107, 196, 197, 198, 218, 248, 251  
 Explicit instruction 247  
 Explicit knowledge 251  
 External factor 141, 146, 155, 158, 244
- F**  
 Fieldwork 37, 59, 90, 252  
 First language 6, 16, 26, 79, 117, 119, 122, 129, 162, 198, 200, 247, 249, 254, 258–260  
 Fluency 56, 133, 134, 135, 250, 256, 262  
 Foreign adoptees 252  
 Formal approach 9, 35, 37, 48, 51, 57, 91–92, 97–98, 126–127, 131, 148, 155, 165, 170, 173–174, 178, 181, 208, 233, 242, 252, 259  
 Frequency 12, 38, 73, 105–106, 125, 131, 137, 148–150, 153–154, 158, 171, 179–181, 185, 187–188, 191, 194, 221, 252, 258, 260  
 Frequency distribution 252  
 Function words 179, 210, 215, 240, 242, 252  
 Functional differentiation 252  
 Functions of codeswitching 67, 86  
 Fundamental difference hypothesis 252
- G**  
 Generation 9, 39, 46, 53, 60–61, 78, 93, 113–116, 119, 122, 131–132, 142, 146, 158, 170–171, 237, 244, 252  
 Generative linguistics 103, 161, 162, 163, 164, 252, 258  
 Glossonym 205, 252  
 Glottophagie 226, 252  
 Grammaticality judgment tests 252
- H**  
 HALA-test 252  
 Heritage language competence 253  
 Heritage language education 253
- Heritage language re-learners 253  
 Heritage speaker, definition 1–11  
 Home language 5, 6, 8, 16, 40, 50, 112, 121, 235, 253, 260  
 Home language education 50, 235, 253
- I**  
 Iberian-based creoles 205, 210, 253  
 Identity 7–8, 42, 44, 50–51, 54–55, 62–63, 65, 67, 80, 84, 86, 122, 130–131, 138, 170–171, 196, 209, 227, 241, 245, 250  
 Immigrant 5–6, 13, 18, 21–23, 27, 30–32, 34–35, 37–39, 41, 49–54, 57–58, 60–61, 65, 79–80, 84, 93–95, 99, 101–102, 109, 119, 123, 125, 128–129, 135, 203, 227–228, 230, 233, 238, 248–250, 252–253, 261  
 Immigration 18, 23, 24, 30, 41, 44, 47–49, 60, 95, 106, 208, 227, 230, 232, 243, 253  
 Imperfect acquisition 53, 89, 146–149, 171–172  
 Implicit knowledge 253  
 Implicit learning mechanisms 253  
 Incomplete acquisition, *see* Imperfect acquisition  
 Incompleteness hypothesis 253, 256  
 Indeterminacy 151, 163–164, 176–177, 253  
 Indigenous language 6, 22, 122, 229, 253  
 Indigenous language revival 22, 229  
 Informal setting 3, 26, 48, 51, 64, 87, 90–92, 94–95, 97–98, 126, 130, 148, 155, 179, 237, 252, 256  
 Insecurity 87  
 Intake 17, 127, 254  
 Inter-individual variability 254  
 Interaction 9, 20, 40, 44, 51, 90, 131, 151, 160–161, 169,

- 172, 181, 195, 197–198, 201, 247, 254
- Interface hypothesis 151, 152, 164, 166, 254
- Internal factor 141, 146, 155, 160
- Internal migration 254
- Internal validity 88, 89–90, 105, 254
- Intra-individual variability 254
- Ius sanguinis 23, 254
- Ius soli 23, 254
- J**
- Judgment 19, 87, 103–107, 118, 125, 127, 142, 165, 200, 252, 254, 261
- Judicial system 255
- K**
- Koiné 27–28, 254
- L**
- Language change 13, 17, 21, 40, 61, 67, 81, 86, 111–112, 131, 161–162, 164, 172, 196, 201, 230, 249, 256, 258, 261
- L1 acquisition, *see* Acquisition, first language
- L1 attrition 149, 254
- L1 restructuring 254
- L2 acquisition, *see* Acquisition, second language
- Language
- attitudes 62, 100, 138, 244, 254
  - choice 18, 43–45, 51–57, 63–64, 67, 78, 83–84, 86, 91, 99–100, 108, 122, 171, 196, 218, 241–242, 254
  - death 56, 60, 237, 254
  - description. 255
  - documentation 89, 225, 226, 250, 255
  - ideology 54, 85, 128, 230, 232, 255
  - loss 1, 6, 61, 119, 239, 249–250, 255
  - maintenance 12, 18, 34, 43–45, 51, 54, 59–60, 62, 65, 79–80, 138, 149, 225, 243–244, 255
  - policy 2, 22, 54, 57, 226–227, 255
  - processing 21, 42, 164, 183–185, 187, 190, 193, 195–196, 200–201, 247, 250, 256, 260–261
  - revival 22, 226, 229, 255
  - rights 225, 226, 243, 255
  - shift 18, 22, 43–45, 52–53, 57–60, 62–63, 65, 78–81, 218, 225, 229, 255, 259
  - use 9–10, 17–18, 23, 30–31, 33, 40–44, 49, 53–58, 67, 69, 82, 85–86, 90, 92, 94, 99–100, 106, 112, 118, 123–125, 128, 130–131, 137–138, 148–150, 172, 183–184, 186–189, 192, 194–196, 200–201, 204–206, 237, 241, 243–245, 248–249, 251, 243, 259
- Less is more hypothesis 255
- Lexical gap 71–72, 74, 241
- Lexical proficiency 135, 136, 255
- Lexifier language 21, 203, 206, 255
- Lingua franca 26, 45, 47, 50, 85, 122, 255
- Linguicism 226, 255
- Linguistic diversity 91, 229, 255
- Linguistic genocide 226, 255
- Linguistic human rights 225, 242–243, 255
- Literacy, degree of 127
- Loan, *see* borrowing
- Loan translation 143, 155, 255
- Locative preposition 213, 256
- Loss, *see* attrition
- M**
- Matrix 69, 101, 151–152, 219–220, 222, 253, 256
- Matrix language 69, 151, 253
- Measuring fluency 256
- Mesolect 256
- Migration 22, 24, 47–48, 98, 114, 200, 203–204, 227, 233, 244, 247–250, 254, 256, 261
- Migration patterns 256
- Missing-input competence divergence hypothesis 148
- Mixed language 84
- Morphological deficits 256
- N**
- Naturalistic setting 1–3, 8–10, 247
- Netherlands 3–4, 7, 9–10, 16–17, 21, 26, 28–29, 41, 46, 48–49, 53, 55, 70–71, 79, 93, 97, 99–100, 103, 114, 122–123, 128, 138, 143–145, 147, 155–156, 159, 179–180, 203–206, 208, 218–219, 219n1, 222–223, 230, 234, 257
- Networks 10, 24, 55, 65, 84, 128, 170, 185, 187, 188, 189, 256
- Normal diachronic change 256
- O**
- Official language 2–4, 26, 207, 230, 257
- Onset of bilingualism 119, 257
- Optimality Theory 153, 161, 169, 172, 244, 257, 261
- Over-acceptance 257
- Overgeneralization 151–152, 216
- Overhearer 52, 252
- P**
- Particle 36, 76, 144, 158, 168, 213–214, 221, 238–239, 250, 257
- Particle verb 144, 221, 257
- Passive auxiliary 206, 207, 214, 215, 257
- Passive construction 248
- Perceptual advantages 257
- Phonetics 31, 97, 262
- Phonetic production benefits 257, 258
- Phonological adaptation 211, 257
- Phonotactic pattern 257
- Picture-matching task 257
- Picture-naming task 258
- Pidgin 255
- Politeness 52, 258

- Post-colonial heritage language 258
- Post-nominal adjective 258
- Pragmatic salience 212, 258
- Prestige 6, 58, 128, 209, 225, 239, 247–248, 256
- Priming 21, 81, 83, 179, 191–192, 198, 218, 258, 264
- Procedural knowledge 199
- Production 21, 102, 107, 112, 118, 127, 142, 152–153, 161, 164, 173, 183–187, 192–195, 198–200, 237, 251–252, 254, 256–258
- Production advantages heritage accent (different from L2 accent) 258
- Proficiency 3, 5–8, 16–17, 42, 45, 51, 53, 57, 62, 72–73, 79, 87, 108, 111–112, 115, 118–118, 121–124, 127–129, 131–138, 157, 183, 191, 194, 198–200, 209, 218, 228, 236, 244, 255, 258, 260, 262
- Pronunciation 19, 34, 67, 121, 154, 190, 199, 200, 241, 257, 261
- Purism 19, 98, 239, 241, 258
- Q**
- Qualitative differences 9, 124, 258
- Quantitative differences 173, 175, 258
- Questionnaires 19, 33, 52, 56, 87, 99, 137–138, 244
- R**
- Racialization 258
- Reanalysis 258
- Receptive bilingual, *see* Overhearer
- Recordings 19, 20, 35, 88, 90, 92–95, 101–102, 104–105, 168, 170, 189, 219, 244
- Reduction 26, 148, 154, 179, 252, 256, 259
- Register 6, 20, 26, 57, 112, 126, 130, 148, 170, 242, 247–248, 259–260
- Register variation 112, 259
- Regression hypothesis 149, 259
- Relexified/relexification 259
- Rememberers 259
- Representation 127, 185, 186, 247
- Representational differences 259
- Restructuring 254, 259
- Retelling 165, 259
- Returnees 158, 259
- Reversing language shift 22, 225, 229, 259
- Revival program 259
- Role input 259
- S**
- Self-study 247
- Semantic domain/field 259
- Semantic extensions 156
- Semantic specificity 72
- Semi-speakers 6, 239, 260
- Sequential bilinguals 15, 17, 40, 119, 120, 121
- Serial verb construction 260
- Sign languages 260
- Simplification 26, 44, 157, 260
- Social factors 44, 48, 58, 67, 84
- Sociolectal variation 206, 260
- Sociolinguistic pressure 260
- Speaker profiles 7, 260
- Speakers' perspective 23, 41
- Speech community 6, 20, 55, 60, 69, 82–83, 90, 126, 128, 130, 169–170, 179, 243, 247, 252, 255, 260
- Speech rate 20, 21, 134, 187, 192, 194, 195, 260
- Stagnated L1 development 260
- Stress 6, 10, 147, 162, 217, 245, 261
- Submersion 208, 260
- Superdiversity 129, 260
- Swadesh list 136, 210
- Syllable 216, 217, 257, 261
- T**
- Transcription 19, 88, 93, 95, 96, 109
- Transfer 63, 79, 103, 146, 155, 157, 171, 183, 192, 195, 198, 201, 244, 261
- Transnational norms 261
- Transparency 148, 151, 153, 159, 261
- Truth value judgment tasks 261
- Two-way immersion 236, 251
- U**
- Ultimate attainment 124, 261
- Unitary Linguistic System Hypothesis 261
- Universal tendencies 210, 261
- Unstressed 205, 261
- Urbanization 254
- Usage-based models 178, 179, 249, 262
- V**
- Variance among bilinguals 262
- Variation 15, 20, 49, 54, 63, 81, 92, 94, 106, 108, 111–115, 117–119, 124, 128, 130–131, 134–135, 138, 141, 147, 161, 163–164, 169, 173–176, 178–179, 181, 184, 205–206, 208, 225, 231, 249, 250, 258–260
- Variationist approach 21, 262
- Vitality 45–46, 49, 56, 57–58, 137, 259, 262
- Vowel 36, 169, 170, 173, 257, 261
- Vulnerable domains 262
- W**
- Weaker Language as L1 hypothesis 262
- Weaker language as L2 hypothesis 262
- World Englishes 262
- Writing system 230–231
- Written data 91, 97–98

Heritage languages, such as the Turkish varieties spoken in Berlin or the Spanish used in Los Angeles, are non-dominant languages, often with little prestige. Their speakers also speak the dominant language of the country they live in. Often heritage languages undergo changes due to their special status. They have received a lot of scholarly attention and provide a link between academic concerns and educational issues. This book takes a language contact perspective: we consider heritage languages from the perspective of their history, their structural properties, and their interaction with other surrounding languages.

ISBN 978 90 272 0471 4



JOHN BENJAMINS PUBLISHING COMPANY