

**Dynamics of Ambon Malay:
Comparing Ambon and the Netherlands**

Published by

LOT

Trans 10

3512 JK Utrecht

The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl

<http://www.lotschool.nl>

Cover illustration: *Migrazioni* by Roberta Venanzi. All rights reserved.

ISBN: 978-94-6093-204-5

NUR 616

Copyright © 2016: Francesca Romana Moro. All rights reserved.

Dynamics of Ambon Malay: Comparing Ambon and the Netherlands

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus,
volgens besluit van het college van decanen
in het openbaar te verdedigen op maandag 21 maart 2016
om 12.30 uur precies

door

Francesca Romana Moro

geboren op 7 september 1985

te Rome, Italië

Promotoren:

Prof. dr. P. C. Muysken
Prof. dr. M. A. F. Klammer

Copromotoren:

Dr. S. P. Aalberse
Dr. H. Hammarström (Max Planck Institute
for Psycholinguistics Nijmegen)

Manuscriptcommissie:

Prof. dr. A. van Kemenade
Prof. dr. K. A. Adelaar (University of
Melbourne, Australië)
Prof. dr. A. M. Backus (Tilburg University)

Contents

Acknowledgements	xi
List of abbreviations	xiii
Chapter 1: Introduction	1
1.1 Objectives and motivations.....	1
1.2 Heritage languages and heritage speakers	4
1.2.1 Heritage speakers and the baseline language	7
1.3 Sources of divergence and convergence in heritage languages.....	10
1.3.1 Cross-linguistic influence.....	12
1.3.1.1 Change in frequency	13
1.3.1.2 Loss or reduction	16
1.3.1.3 Grammatical reanalysis.....	18
1.3.2 Incomplete acquisition.....	20
1.3.2.1 Loss or reduction.....	21
1.3.2.2 Grammatical reanalysis.....	22
1.3.3 Attrition.....	22
1.3.3.1 Loss or reduction.....	23
1.3.3.2 Grammatical reanalysis.....	24
1.3.4 Type of input	25
1.3.5 Universal principles	26
1.3.6 Interim summary	28
1.4 Working Hypotheses.....	30
1.4.1 The Alternation Hypothesis, the Vulnerability Hypothesis, and the Functional Convergence Hypothesis	31
1.4.2 The Conceptual (Transfer) Hypothesis	33
1.5 History of Ambon Malay: the homeland community and the heritage community.....	35
1.5.1 The homeland community in the Central Moluccas	41
1.5.1.1 The homeland language.....	42
1.5.1.2 Previous studies on homeland Ambon Malay	44

1.5.2	The heritage community in the Netherlands	45
1.5.2.1	The heritage language.....	49
1.5.2.2	Previous studies on heritage Ambon Malay	52
1.6	Sketch grammar of Ambon Malay	54
1.6.1	Phonology.....	54
1.6.2	Word classes	56
1.6.2.1	Nouns.....	56
1.6.2.2	Verbs.....	56
1.6.2.3	Adverbs.....	57
1.6.2.4	Personal pronouns	57
1.6.2.5	Demonstratives	58
1.6.2.6	Quantifiers and classifiers.....	59
1.6.2.7	Prepositions	60
1.6.2.8	Conjunctions	61
1.6.3	Word formation	61
1.6.3.1	Affixation.....	61
1.6.3.2	Reduplication.....	62
1.6.4	The noun phrase	62
1.6.5	The verb phrase	64
1.6.6	The clause.....	66
1.6.6.1	Verbal clauses	66
1.6.6.2	Non-verbal clauses.....	67
1.6.7	Verb serialization.....	67
1.6.8	Negation	68
1.7	Overview of the chapters.....	69
Chapter 2: Methodology		73
2.1	Speaker sample	73
2.1.1	The baseline groups: issues and problems	76
2.1.2	Participants.....	79
2.1.2.1	Heritage speakers in the Netherlands	79
2.1.2.2	Homeland speakers in Ambon	83
2.1.2.3	First generation speakers in the Netherlands	85
2.1.2.4	Dutch speakers in the Netherlands	86
2.2	Elicitation material and tasks	87
2.3	The corpus	91

2.4	Analytic procedure.....	92
-----	-------------------------	----

Chapter 3: Nominal modification in heritage Ambon Malay..... 97

3.1	Introduction.....	97
3.2	Word order of nominal modifiers in Ambon Malay and Dutch.....	100
3.2.1	Word order of nominal modifiers in Ambon Malay.....	100
3.2.2	Word order of nominal modifiers in Dutch.....	104
3.2.3	Interim summary	107
3.3	The study	108
3.3.1	Objective.....	108
3.3.2	Research questions and hypotheses	108
3.3.3	Participants, task, and responses	110
3.4	Results and discussion	112
3.4.1	Frequency of word order patterns	112
3.4.1.1	Results	112
3.4.1.2	Discussion	116
3.4.2	Grammatical reanalysis of (in)definite markers.....	119
3.4.2.1	Results.....	120
3.4.2.2	Discussion	123
3.5	Conclusions.....	127

Chapter 4: Aspectual distinctions in heritage Ambon Malay 129

4.1	Introduction.....	129
4.2	The restructuring of aspect in heritage languages	132
4.3	Tense-aspect in Ambon Malay and Dutch.....	135
4.3.1	Tense-aspect in Ambon Malay	135
4.3.2	Tense-aspect in Dutch.....	138
4.3.3	Interim summary	140
4.4	The study	140
4.4.1	Objective.....	141
4.4.2	Research questions and hypotheses	141
4.4.3	Participants, task, and responses	142
4.5	Results and discussion	143
4.5.1	Results of the simultaneous video description task	143
4.5.2	Results of the video-clip retelling task.....	145
4.5.3	Discussion.....	150

4.6	Conclusions.....	154
-----	------------------	-----

Chapter 5: Give-constructions in heritage Ambon Malay..... 157

5.1	Introduction.....	157
5.2	<i>Give</i> -constructions in Ambon Malay and Dutch	159
5.2.1	<i>Give</i> -constructions in Ambon Malay	160
5.2.2	<i>Give</i> -constructions in Dutch	164
5.2.3	Interim summary	168
5.3	The study	169
5.3.1	Objective.....	169
5.3.2	Research questions and hypotheses	169
5.3.3	Participants, task, and responses	170
5.4	Results and discussion	172
5.4.1	Frequency of <i>give</i> -constructions	172
5.4.1.1	Results	172
5.4.1.2	Discussion	174
5.4.2	Preposition selection.....	180
5.4.2.1	Results	180
5.4.2.2	Discussion	181
5.5	Conclusions.....	182

Chapter 6: Resultative constructions in heritage Ambon Malay 185

6.1	Introduction.....	185
6.2	Resultative constructions in Ambon Malay and Dutch	187
6.2.1	Resultative constructions in Ambon Malay.....	188
6.2.2	Resultative constructions in Dutch	192
6.2.3	Interim summary	195
6.3	The study	196
6.3.1	Objective.....	196
6.3.2	Research questions and hypotheses	196
6.3.3	Participants, task, and responses	198
6.4	Results and discussion	200
6.4.1	Frequency of resultative constructions	200
6.4.1.1	Results	200
6.4.1.2	Discussion	202
6.4.2	Preposition selection.....	207

6.4.2.1	Results.....	207
6.4.2.2	Discussion	208
6.5	Conclusions.....	210
Chapter 7: Divergence and convergence in heritage Ambon Malay: structural and social factors		213
7.1	Introduction.....	213
7.2	Structural factors and language change.....	214
7.2.1	Cluster analysis.....	218
7.2.2	Discussion	222
7.3	Social factors and language change	229
7.3.1	Effect of social factors.....	232
7.3.2	Discussion	237
7.4	Conclusions.....	239
Chapter 8: Conclusions and future research.....		241
References		245
Appendix 1		273
Appendix 2		277
Appendix 3		291
Nederlandse samenvatting.....		295
Curriculum vitae		301

Acknowledgements

This dissertation is the result of the help, guidance and support of many people. I could have never done it without them.

First of all, I want to thank all the people who directly or indirectly participated in this study. My deepest thank goes to the members of the Moluccan community in the Netherlands who have participated to the interviews or have helped me to find participants. Without them this dissertation could not have been written. *Tarima kasi*, Otjep, Wim, Dina, Palau, Maroesja, Vanessa, Patrick, Jan, Loes, Lulu, Franciska, Eduardo, Alex, Albertho, Arie, Lolita, Netty, Dave, Albarthina, Ina, Max, Niko, Mas, Buce, Fera, Funreng, Marthen, Raymond, Jospheh, Daniel, Riet, Juliana, Becca, Mick, Joshua, Sarita, Greet, Henk, Helentje, Frans, Ruth, Johannes, Joyce. They have welcomed me in their community, in their families, and in their homes. I have found not only participants for my study, but also friends. Next, I want to thank the participants in Ambon, Indonesia. My deepest gratitude goes to Rose, Rovien, Meggy and their families who have hosted me and helped me during my fieldwork in Ambon, thank you very much! *Tarima kasi* also to Debby, Josefien, Neviac, Meri, Sia, Gerald, Leonora, Ferri, Joyce, Paulina, Chris, Hendrik, Andrew, Awa, Hapsa, Marce, Rina, Poli, Lucas, Dominggus, Josina, Frans, Cristal, Lovindis, Gloria, Ratna, Martha, Helena, Barnesy, George, Dolfina, Oli, Lili. I also want to thank all the Dutch participants who have been interviewed by Rowan Soolsma. Their contribution has been invaluable for the completion of this dissertation.

I want to express my sincere gratitude to my promotors, Pieter Muysken and Marian Klammer, and to my copromotors, Suzanne Aalberse and Harald Hammarström. Thank you for your unwavering support, your thoughtful advice and your patience throughout these years. Thank you for believing in me and encouraging me to complete this dissertation. I have learned a lot from each and every one of you!

I am also very grateful to all my colleagues at the Radboud University and elsewhere for the countless discussions, the valuable comments and the constant encouragement they showed me. First of all, a word of thanks goes to Pablo Irizarri van Suchtelen, my partner in crime in this heritage language research project. Writing this dissertation without you would have been much less fun. Our trips, especially the one to California, have been memorable! Thanks to you I have a picture of every squirrel at UCLA. Our 'traumatic' dinners together with Harald have

been the best anti-stress therapy! I also want to thank Betty Litamahuputty for her invaluable help in coding the data of Chapter 3, Hülya Şahin for the nice chats during our coffee and cigarette breaks, Gerrit Jan Koostra, Linda van Meel, and Francesco Bassi for their help with statistics, Margot van den Berg, Robbert van Sluis, and Lotte Hogeweg for being such wonderful and patient office mates, Pelin and Derya for their support and enthusiasm at all the conferences and workshops we attended together, Angela Kluge for the many discussions about Malay languages and fieldwork in Indonesia, Sophie Villerius, Loretta O'Connor, Kofi Yakpo, Cefas van Rossem, Robert Borges, Joshua Birchall, Olga Krasnoukhova, and Neele Müller for having created such a nice work environment in the research group.

A word of thank also goes to the internship students, student assistants and MA students who contributed to this research, namely Rose Lekawael, Rowan Soolsma, Palau Tetelepta, Jusmianti Garing and Feny Eky.

Thanks also to Hella Cranen-Jooren, Dirkje van der Aa, Lisenka Fox, and the staff members of the ICT Helpdesk for your help with bureaucracy, logistics and general life at Radboud University.

Finally, without the love and support of my family and friends, I would have never gone so far. My greatest thank goes to my parents, my brother, my friends in Rome, my friends in Nijmegen, my friends scattered around the world, who have been with me through all of the ups and downs of this research project. I owe a special thanks to Roberta, who designed the front cover. Robi, you are a great artist! Last, but not least, thank you, Alessandro, for your love and support throughout these three years, thank you for always being there for me, and thank you for the amazing work you did in editing my dissertation. I can't wait to start a new "chapter" with you.

List of abbreviations

1	First person	Fr	French
2	Second person	Ge	German
3	Third person	glm	General linear model
A	Agent argument	Gn	Generation
ACC	Accusative	HC	Times to home country
ACL	Accidental	HL	Heritage language
ADV	Adverb	INDF	Indefinite
AM	Ambon Malay	INF	Infinitive
AoA	Age of acquisition	INTENS	Intensifier
AP	Ajectival phrase	ITER	Iterative
APiCS	Atlas of pidgin and creole language structures	L	Language
Arr.	Arrival	L1	First language
ART	Article	L2	Second language
CB	Cut and break	L3	Third language
CLF	Classifier	LOC.DIST	Locative, distal
Cz	Czech	LOC.MED	Locative, medial
DEF	Definite	LOC.PROX	Locative, proximal
DIM	Diminutive	LOR	Length of residence in the Netherlands
D.DIST	Demonstrative, distal	LSEM	Landelijk steunpunt educatie Molukkers
D.PROX	Demonstrative, proximal	Ls	Other languages spoken
DL	Dominant language	M	Mean
DO	Double object	M	Masculine
DOM	Differential object marking	Mo	Mother
Du	Dutch	<i>n</i>	Number of speakers
En	English	N	Neuter
EXCL	Exclamative	NEG	Negation
EXIST	Existential verb	NL	The Netherlands
F	Feminine	NP	Noun phrase
Fa	Father	O	Object
FML	Formal	ORD	Ordinal number

Pa	Partner
PL	Plural
Po	Portuguese
PO	Prepositional object
POSS	Possessive
PP	Prepositional phrase
PRF	Perfect
PRO	Pronoun
PROH	Prohibitive
PST	Past
PTCP	Participle
R	Recipient argument
RECP	Reciprocal
REL	Relativizer
S	Subject
Sb	Sibling
SD	Standard deviation
SG	Singular
SIG	Significance value
Sp	Speaker
Spa	Spanish
SVC	Serial verb construction
Sx	Sex
T	Theme argument
V	Verb
VOT	Voice Onset Time
Vp	Verb particle
WALS	World atlas of language structure
-	Affix boundary
=	Clitic boundary
?	Missing information

CHAPTER 1

Introduction

This first chapter is an introductory chapter meant to orientate the reader for what is to follow. The next section is a statement of research objectives and motivations. The following sections outline the most relevant information regarding heritage languages and heritage speakers (Section 1.2), the sources of divergence and convergence in heritage languages (Section 1.3), the working hypotheses of this research (Section 1.4), the history of the Ambon Malay speaking community (Section 1.5), in Ambon (Section 1.5.1) and in the Netherlands (Section 1.5.2), and the grammar of Ambon Malay (Section 1.6). The last section (Section 1.7) provides a brief description of the chapters of this dissertation.

1.1 Objectives and motivations

Mobility and migration are central aspects of the modern society. People have always moved across the world for all sorts of reasons, in search of better livelihoods, to look for economic opportunities, to escape from wars and conflicts, and when they have moved, they have carried their cultures and their languages with them. As a result of voluntary or forced migration, nowadays many people live outside their homeland and an increasing number of languages is spoken far away its original region. For instance, “within the European Union there are [...] more than 60 indigenous regional and minority languages, and many non-indigenous languages spoken by migrant communities” (TNS Opinion & Social, 2012). The indigenous and non-indigenous minority languages are also called ‘heritage languages’ (Fishman, 2001; van Deusen-Scholl, 2003; Polinsky & Kagan, 2007; Benmamoun, Montrul, & Polinsky, 2010; Nagy, 2015).

Due to past and recent migration patterns, many languages have a homeland variety and a so-called heritage variety that is spoken in the countries of Europe, in the U.S.A, Canada or Australia. For instance, homeland Russian is spoken in Russia, and heritage Russian or American Russian is spoken in the U.S.A. (Polinsky, 2006, 2008a, 2008b, 2008c; Laleko, 2010), Turkish is spoken in Turkey, and heritage

Turkish or Dutch Turkish is spoken in the Netherlands (Backus, 1996; Onar Valk, 2015), Ambon Malay, the language under investigation in this dissertation, is spoken in the Central Moluccas (Indonesia) and in the Netherlands, where it is known as heritage Ambon Malay or *Melaju Sini* ‘Malay from here’ (Tahitu, 1989; Huwaë, 1992).

Through the course of time, due to the intense contact with the dominant language of the country (e.g., English or Dutch), and to the restricted domains of usage (typically the home) heritage languages have come to diverge significantly from their homeland varieties, to the point that they deserve a grammatical description on their own right (Silva-Corvalán, 1994; Backus, 1996; Polinsky & Kagan, 2007; Benmamoun et al., 2010; Laleko, 2010, among others). Quite ironically, in the case of Ambon Malay, the grammatical description of the heritage variety (Tahitu, 1989) was published almost ten years before the grammatical description of the homeland variety (van Minde, 1997). However, a systematic comparison between the two varieties has never been carried out until now. The need for such a comparison was already felt by Tahitu, who (1989, p. 159) concluded his grammar by saying:

This description of MS [*Melaju Sini*: Ambon Malay in the Netherlands] is not complete. Further research is necessary. Comparing MS with AM (the Malay language in Moluccas) can give more insight on divergence. The same variety has developed in two ways: Malay in the Moluccas alongside standard Indonesian, and the same Malay language in Holland (via Tangsi Malay) alongside a non-related language, Dutch.

The aim of this dissertation is to fill this gap by comparing heritage Ambon Malay, as spoken in the Netherlands, to its homeland variety, as spoken in Ambon, Indonesia, hence the title *Dynamics of Ambon Malay: comparing Ambon and the Netherlands*. More specifically, the present study investigates divergence from the homeland variety and convergence toward the dominant language, Dutch, by focusing on some specific areas of heritage Ambon Malay grammar and providing quantitative analysis of the observed patterns. Furthermore, this dissertation also includes information regarding the linguistic variation found among speakers and to the socio-linguistic variables that account for such variation. The following general questions are, thus, addressed in this dissertation:

- Does heritage Ambon Malay diverge from its homeland variety?

- Is heritage Ambon Malay changing under the influence of Dutch? How does this ‘on-going’ change manifest itself?
- What are the factors driving this ‘on-going’ change?

Onar Valk (2015, p. 38) correctly points out that:

[A] contact-induced change can only be established after *systematic quantitative* comparisons of a *diagnostic* linguistic feature with an earlier or pre-contact stage, with a non-contact variety, and most important, with the presumed model or source variety (emphasis mine).

The chapters of this dissertation follow these guidelines and *systematically* compare heritage Ambon Malay to homeland Ambon Malay (the pre-contact variety) and to Dutch (the source variety). Four linguistic features are selected as *diagnostic*: nominal modification (order of nominal modifiers), aspectual distinction (frequency and distribution of aspect markers), *give*-constructions and resultative constructions. These features are chosen for two reasons. First they cover a wide portion of heritage Ambon Malay grammar, from word order to constructions, and thus give an insight on the extent to which the heritage language diverges from the homeland language. Second, these grammatical areas display internal variation, and thus allow for an observation of possible cross-linguistic effects leading to convergence with Dutch. As shown in Section 1.3.1.1, alternation of structures in the heritage language is one of the loci for cross-linguistic influence. Finally, the comparison carried out in this dissertation is *quantitative* because it uses frequency rates and statistical tests as evidence for contact-induced change.

By investigating contact-induced changes in a heritage Malay variety, this dissertation aims to contribute to two fields of study: contact-linguistic and Malay linguistics.

One of the questions of contact linguistics is “what kinds of situation promote one type of outcome rather than another?” (Winford, 2003, p. 5). Studying the heritage speaker population allows us to understand the outcome of language contact in a situation where two languages of unequal status are spoken alongside each other for about 50-100 years (two to three generations). The study of heritage languages takes a bottom-up perspective on contact-induced change. Instead of starting from the resulting language and trying to reconstruct the scenario that brought it about, heritage language research knows the situation and it investigates the types of contact-induced change that are allowed and constrained by it. In other words, studying heritage languages allow us to observe the ontogeny of contact

induced change in a naturalistic setting. The bottom-up approach allows us to make clear predictions with respect to the direction and the types of change that take place in asymmetric bilingual populations. Furthermore, the results of heritage language studies can help us reconstruct the early stages of change in languages with a long history of contact (hundreds of years), where these stages are no longer visible (see Section 1.3.1.3).

Studying heritage Ambon Malay also contributes to our understanding of the development of Malay contact varieties or Pidgin-derived Malay varieties (Adelaar, 2005). Although some of these varieties have already been described in detail (see, for instance, van Minde, 1997 for Ambon Malay; Litamahuputty, 2012 for Ternate Malay; and Kluge, 2014 for Papuan Malay), quantitative studies on specific linguistic features are still lacking. More studies targeting specific constructions and supporting the investigation with quantitative data and statistical tests are needed in order to understand language internal variation. Claims such as the following one need to be backed up by statistical data, analysis and figures:

It is yet unclear what the semantic contrast is between the two constructions *ini/itu*+HEAD [demonstratives+HEAD] versus HEAD+*ini/itu* [HEAD+demonstratives]. Statistically speaking, however, the second of these is far more dominant” (van Minde, 1997, p. 147).

The quantitative analysis of grammatical features carried out in this dissertation will hopefully give more insight on the frequency and usage patterns of Ambon Malay as it is employed by its speakers, both in Ambon and in the Netherlands.

1.2 Heritage languages and heritage speakers

Generally speaking, heritage speakers are individuals who grew up in a household where the language spoken is not the dominant language of the larger national society. The language spoken in the household is referred to as the ‘heritage language’. Two definitions of heritage speakers are currently in use, a ‘broad’ definition and a ‘narrow’ definition (Polinsky & Kagan, 2007). According to the ‘broad’ definition, heritage speakers are all those individuals that “have been raised with a strong cultural connection to a particular language through family interaction” (van Deusen-Scholl, 2003, p. 222), even though they cannot actually speak the language. These individuals often enroll as students in language

classrooms aiming to re-learn their heritage languages (Carreira, 2004; Carreira & Kagan, 2011). Since they cannot speak the language, these students are comparable to L2 learners and are therefore better referred to as ‘learners with a heritage motivation’ rather than ‘heritage speakers’ (van Deusen-Scholl, 2003). According to the ‘narrow’ definition, heritage speakers are individuals raised in a home where the heritage language is spoken and who are to some degree bilingual in the heritage language and in the dominant language of the country (Polinsky & Kagan, 2007, p. 369). The crucial component of the ‘narrow’ conception is the passive and active use of the heritage language (at least during childhood). In the present study, I adopt the narrow definition (see also Section 2.1).

To give an example, a typical Ambon Malay heritage speaker in the Netherlands (under the narrow definition) is Frans. Frans’ parents arrived in the Netherlands from the Central Moluccas in 1951. His father was a soldier in the Dutch East Indies army and his parents spent a long period of time in the army camps in Java, where they spoke a divergent variety of Ambon Malay known as Tangsi Malay. Frans was born and raised in a Moluccan camp in the Netherlands, where he spoke Tangsi Malay with his parents and with the members of the other Moluccan families living in his and in the neighboring barracks (see Section 1.5.2). When Frans was five years old, his parents moved to a Moluccan ward in a Dutch town and Frans started elementary school, where he began to be consistently exposed to Dutch. When his siblings joined elementary school, he began to speak Dutch with them, and slowly, but surely, Dutch became one of the languages of communication in his household. In the course of time, by attending Dutch schools and working in Dutch-language environments, Dutch became his functionally dominant language. By the time Frans reached early adulthood, his first language (in terms of order of acquisition) had become his second language (in terms of functional dominance), and his second language had become his dominant language. Nowadays, Frans mainly speaks Dutch in his daily life, but he also speaks his heritage language when he talks to his parents, when he meets other Moluccan elders, or when he goes to Moluccan events, ceremonies or gatherings.

Figure 1.1 (based on Montrul, 2012, p. 4) shows a schematic representation of language shift in bilingual speakers like Frans. Simultaneous bilinguals acquire the two languages at the same time (from birth) while sequential or successive bilinguals acquire the heritage language (HL) from birth and the socially dominant language (DL) after the age of four or five, thus after the heritage language has already started developing. Some speakers, especially those who grew up acquiring

simultaneously the heritage language and the dominant language, experience a less severe shift during early and late childhood because they have been exposed to both languages since birth. The dotted lines in Figure 1.1 represent the amount of use of the dominant language in simultaneous bilingual heritage speakers. Regardless of the age of onset of bilingualism (birth or age of five), all heritage speakers become more fluent in the dominant language by adolescence and/or young adulthood. At that age the heritage language begins to display signs of incomplete acquisition, attrition and transfer from the dominant language (see Section 1.3).

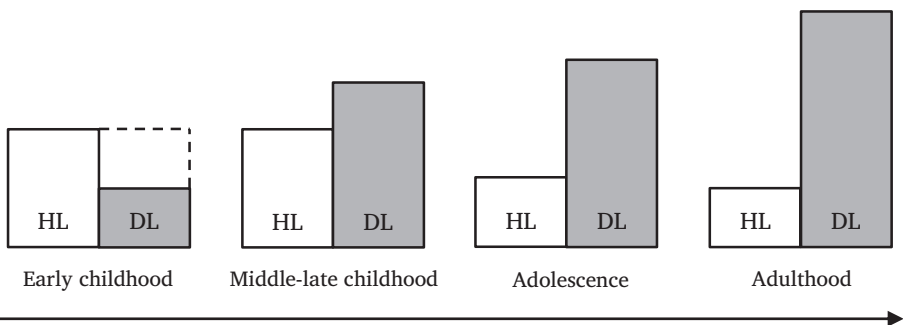


Figure 1.1: Language shift in bilingual heritage speakers.

The crucial point of heritage language speakers like Frans is that they are not isolated individuals. This characteristic differentiates them from other bilingual speakers (i.e., early bilinguals from mixed marriages or expat parents). They are part of a bilingual speech community, the heritage community, whose members share similar autobiographical characteristics and have experienced a shift in their dominant language to various degrees. Furthermore, they speak a language that has a homeland or monolingual counterpart to which their language can be compared. To summarize, the following criteria characterize a heritage language (Aalberse & Muysken, 2013, p. 3):

- a. A language acquired at an early age in a naturalistic setting (usually the home, but not necessarily)
- b. A language that is not the dominant language of the country
- c. A language which has cultural value for the speaker
- d. A language with a long history in the country of residence (about two generations)

- e. A language that is spoken by a community of people who share a common ancestry
- f. A language that is (or once was) spoken as the majority language in a country or region (homeland).

The last criterion (f) is absent from Aalberse and Muysken's (2013) definition. I added it in order to exclude urban youth languages from being considered as heritage languages. In fact, criteria (a-e) could also apply to youth languages, such as Nouchi in Cote d'Ivoire. Van Rijswijk, Muntendam and Dijkstra (in preparation) have formulated a narrower definition of heritage languages by including three additional criteria, which are (1) the L1 is an immigrant language, (2) the L1 is not fully attained, and (3) the bilingual received no or limited L1 education. Although this definition is narrow enough to rule out youth languages, it also rules out indigenous languages, such as Quechua in Peru or indigenous languages in Indonesia. Following Fishman (2001) and van Deusen-Scholl (2003), I think that any language that fits the criteria (a-f) and is spoken alongside another (dominant) language can have the status of heritage language, be it indigenous or immigrant.

In this context it is important to note that, despite the fact that all heritage speakers acquire the heritage language naturally in early age, they reach different levels of proficiency in the language. As a consequence, the heritage variety they speak diverges in various ways from the baseline language, namely the language from which it derives (Montrul, 2004, 2009, 2011; Polinsky, 2006, 2008a, 2008b, 2008c, 2011; Pires & Rothman, 2009; Montrul & Bowles, 2009; Benmamoun et al., 2010; Laleko, 2010; O'Grady, Kwak, Lee, & Lee, 2011; Shi, 2011; Onar Valk, 2015, among many others). Before discussing the sources of heritage language divergence (Section 1.3), I now turn to the notion of baseline language.

1.2.1 Heritage speakers and the baseline language

Heritage languages are usually studied in comparison with a baseline language. Different baseline languages can be chosen depending on the purpose of the research. One can choose to parallel the heritage language to its homeland variety (Onar Valk, 2015), or to the interlanguage of L2 learners (Montrul, 2011), or to the language of homeland (monolingual) children (O'Grady, Lee, & Lee, 2011), or again to the homeland variety spoken by illiterate speakers (Pires & Rothman, 2009). The

choice of one (or more) of these baseline languages is mostly dictated by the research questions one aims to answer.

The majority of heritage language studies, including the present one, compares adult (bilingual) heritage speakers to adult (monolingual) homeland speakers (Polinsky, 2008a, 2008b; Montrul, 2009; Montrul & Bowles, 2009; Irizarri van Suchtelen, 2014; Onar Valk, 2015, among others). Since heritage speakers speak a variety that is in contact with a dominant language, while homeland speakers speak a variety without such contact, the comparison between the two aims to filter out possible effects of contact (from the dominant language) on the heritage language. For instance, Onar Valk (2015) investigates the use of finite and non-finite subordination in heritage speakers of Turkish in the Netherlands and in homeland speakers of Turkish in Turkey. Onar Valk concludes that the higher incidence of finite subordination in Dutch-Turkish when compared to Turkey-Turkish is related to the influence of Dutch, the dominant language of these speakers.

Polinsky & Kagan (2007, p. 372) correctly point out that the proper “baseline language for a heritage speaker is the language that he or she was exposed to as a child”. The authors warn that comparing the language spoken by heritage speakers to the standard language of the homeland country as promoted by the school and the media would yield misleading results. For instance, it would be a mistake to compare heritage Vietnamese in the U.S.A. to Standard Vietnamese (which is based on the northern dialect), as the majority of Vietnamese heritage speakers come from southern Vietnam, where another dialect is spoken. The geographical origins of heritage speakers and the consequent dialect variation need to be taken into account when selecting the baseline language, and homeland comparators need to be chosen from the same geographical regions so as to match the origins of heritage speakers (see Section 2.1.1 and Section 2.1.2.2).

An equally possible baseline group can be that of L2 learners (Au, Knightly, Jun, & Oh 2002; Montrul, 2010, 2011; Polinsky & Laleko 2013; see Montrul, 2012 for an overview of studies comparing heritage speakers to L2 learners). Heritage speakers and L2 learners share the same set of languages (dominant language (L1)-heritage language (L2)) but they differ in terms of age of acquisition. Heritage speakers acquire the heritage language in early age, whereas L2 learners acquire the L2 after puberty. An analogy between heritage speakers to L2 learners, thus, allows for an observation of the effect of age of acquisition. Au et al.’s (2002) study shows that heritage speakers and L2 learners differ with respect to phonetics/phonology and pronunciation measures, with heritage speakers being more native-like, while they

do not differ on morpho-syntactic measures. These results have been partially confirmed by other studies and the empirical findings collected so far seem to indicate that early acquisition gives an advantage only in the areas of syntax (V2 rule) and phonology, while no advantage is attested in the areas of inflectional morphology syntax-discourse, and semantics¹ (Montrul, 2012).

We can also compare the grammatical abilities of heritage speakers to those of homeland (monolingual) children (O'Grady, Lee, & Lee, 2011; Polinsky, 2011). The two groups share a similar path of acquisition up to early childhood (age 5), after this age the exposure to the heritage language decreases drastically in the case of heritage speakers, but not in the case of monolingual children in the homeland. A comparison between the two can be used to investigate the role of (prolonged) input on the acquisition of certain linguistic features. A clear example is the acquisition of the subjunctive in Spanish. Monolingual children complete the acquisition of the subjunctive only at around the age of ten. A number of studies (Silva-Corvalán, 1994; Montrul, 2008, 2009, among others) argue that heritage speakers of Spanish, who receive less input in Spanish since about the age of five, display errors with the subjunctive because they never got the chance to complete the acquisition process (which culminates after age 10).

Finally, Pires and Rothman's (2009) study shows that heritage speakers can (and in some cases need to) be compared to illiterate homeland speakers in order to filter out possible effects of literacy and schooling on the divergence between homeland and heritage grammars. An example is the lacking knowledge of inflected infinitives in Brazilian Portuguese heritage speakers. Pires and Rothman (2009) argue that the lack of knowledge is not due to incomplete acquisition or to attrition but rather to the types of input the heritage speakers received. Inflected infinitives are acquired by homeland teenagers and adults only by means of formal education. Since heritage speakers are exposed only to colloquial input and do not typically get formal education in the heritage language, they never get the chance to acquire these properties. Thus, the lack of knowledge of inflected infinitives is not a case of incomplete acquisition because inflected infinitives are part of a register which heritage speakers have no access to.

¹ The results of Montrul's (2011) study on nominal and verbal inflection reveal a task effect, such that L2 learners were more accurate on the written tasks, while heritage speakers were more accurate on the oral task. Montrul (2011, p. 188) concludes that if one considers the oral task as more representative of implicit linguistic knowledge, then heritage speakers may be said to have more nativelike morphological skills.

We have seen that comparing heritage speakers to different baseline languages can help us to disentangle the sources of heritage grammar divergence, to which I now turn. We have seen that comparing heritage speakers to different baseline languages can help us to disentangle the sources of heritage grammar divergence, to which I now turn.

1.3 Sources of divergence and convergence in heritage languages

As seen in the previous section, the main focus of studies on heritage languages and heritage language acquisition has been to identify the sources or causes explaining the divergence between the heritage grammar and a baseline grammar, usually the grammar of the homeland variety. The main factors accounting for such divergence are cross-linguistic influence, incomplete acquisition, attrition, the different type of input heritage speakers are exposed to, and universal principles in language acquisition in contact settings (Torres Cacoullous, 2000; Polinsky, 2006, 2008a, 2008b, 2011; Polinsky & Kagan, 2007; Montrul, 2008, 2009; Pires & Rothman, 2009; Benmamoun et al., 2010; Laleko, 2010; Onar Valk 2015). Each of these factors can have different manifestations and lead to two types of outcome: divergence from the homeland language or convergence to the dominant language.

In this study I use the term ‘divergence’ to refer to any systematic difference in any area of the grammar between two languages assumed to be roughly the same at the onset of contact. I use the term ‘convergence’ to refer to “the achievement of greater structural similarity in a given aspect of grammar of two or more languages assumed to be different at the onset of contact” (Silva-Corvalán, 1994, p. 4; see also Winford, 2003; Backus, 2004; Silva-Corvalán, 2008; Matras, 2009). Winford (2003, p. 63) points out that:

Two languages can be said to have converged structurally when previous differences in grammar between them are reduced or eliminated either because one adopts structural features from the other [...], or because both adopt an identical compromise between their conflicting structures.

In situations of language contact characterized by asymmetrical bilingualism, such as heritage languages, the heritage language adopts structural features from the dominant language. Hence, convergence is mostly unidirectional. One of the

strategies of bilingual speakers is to prefer grammatical structures that are shared by both languages (Muysken, 2013). This process can also be considered unidirectional because speakers copy the frequency of the structure in the dominant language to the heritage language (Johanson, 2002), and thus bring the heritage language closer to the dominant language.

Divergence and convergence are two outcomes of language contact that can be brought about by a number of factors, both external (cross-linguistic influence) and internal (incomplete acquisition, attrition). For instance, influence from the dominant language can manifest itself as a shift in preference patterns between two equally possible options. This shift leads to a greater structural dissimilarity between the heritage language and the homeland language (divergence). At the same time the shift leads to a greater similarity between the heritage language and the dominant language (convergence). Incomplete acquisition can manifest itself as a reduction in the heritage grammar (e.g., from a three-gender system to a two-gender system), leading to greater dissimilarity between the heritage language and the homeland language (divergence), without necessarily leading to a greater similarity with the dominant language (convergence). So, while convergence toward the dominant language necessarily implies divergence from the homeland variety, divergence does not necessarily entail greater convergence.

The next sections discuss the factors yielding divergence and convergence in heritage grammars, namely cross-linguistic influence (Section 1.3.1), incomplete acquisition (Section 1.3.2), attrition (Section 1.3.3), the different types of input heritage speakers are exposed to (Section 1.3.4), and universal principles of language development in the context of language disuse (Section 1.3.5). Each of these factors can have different manifestations, such as a change in frequency between available options, loss or reduction, or grammatical reanalysis. Cross-linguistic influence, for instance, can manifest itself as a change in frequency, loss or reduction, or grammatical reanalysis. Likewise, incomplete acquisition can manifest itself as loss or reduction, or grammatical reanalysis, etc. The following sections discuss these factors separately for the sake of convenience, but the reader needs to be aware that the above mentioned factors can sometimes have the same manifestations and yield similar outcomes. Overgeneralization of overt subject pronouns, for example, can be the result of both attrition and cross-linguistic influence. These factors are likely to act in a cumulative way, and are therefore difficult to tease apart. Section 1.3.6 provides a summary of the most important notions discussed in this section.

1.3.1 Cross-linguistic influence

Cross-linguistic influence or transfer is defined as “the influence of a person’s knowledge of one language on that person’s knowledge or use of another language” (Jarvis & Pavlenko, 2007, p. 1). Following Jarvis and Pavlenko (2007), I use the terms ‘cross-linguistic influence’ and ‘transfer’ as synonyms. Transfer can go in different directions, from L1 to L2 (forward transfer), or from L2 to L1 (reverse transfer), or again from L2 to L3 (later transfer). It can also occur in different areas of language knowledge and use, such as phonology, morpho-syntax, lexicon, discourse and pragmatics. Most approaches to language contact make a general distinction between lexical and structural transfer which is based on the type of linguistic material transferred from one language into another (Thomason & Kaufman, 1988; Winford, 2003; Matras, 2009). Lexical transfer involves the transfer or ‘copy’ of the phonological form, while structural transfer involves “the organization, distribution and mapping of the grammatical or the semantic meaning while the form itself is not borrowed” (Sakel, 2007, p. 15). Winford (2003, p. 210) warns that the term transfer is often used in the literature to refer both to the “manifestations” of cross-linguistic influence and to “the psycholinguistic processes that bring them about”. Following Winford (2003), I use the term ‘transfer’ or ‘cross-linguistic influence’ to refer to the process leading to the change and the term ‘convergence’ to refer to the outcome of such change (see Section 1.3).

In the case of heritage languages, the unequal status of the two languages in terms of prestige and functional domains mostly leads to transfer from the dominant language to the heritage language, while the intense contact situation allows both lexical and structural transfer. The effects of dominant language influence on heritage languages have been particularly well-documented in the lexicon, the word order, the use of prepositions and articles, and verb subcategorization (Silva-Corvalán, 1994; Schoenmakers-Klein Gunnewiek, 1997; Benmamoun et al., 2010; Montrul & Ionin, 2010; Onar Valk, 2015). Lexical-semantic calques from the two languages, for instance, occur very frequently. Silva-Corvalán (2008, p. 217) explains that:

Transfer starts with the calquing of concrete structures in situated acts of communication, but in time the lexical units affected may change their semantic features and their possibilities of co-occurrence beyond the specific construction which was originally transferred.

For example, heritage speakers of Spanish in the U.S.A. have extended the meaning of *para atrás* ‘behind’ on the model of English *back*. Since *para atrás* has adopted the semantic subcategorization properties of its English counterpart, it is now used in expressions such as *dar [algo] para atrás* ‘give something back’ or *llamar para atrás* ‘to call back’ (Silva-Corvalán, 1994). According to Silva-Corvalán (1994, 2008) and Backus, Doğruöz & Heine (2011), lexical-semantic calques from the dominant language open the door for further structural changes. For them, contact-induced change proceeds from individual expressions and constructions to more general syntactic schemata or patterns. I come back to this point in Chapter 7, where I use Hartsuiker’s (Hartsuiker, Pickering, & Veltkamp, 2004; Schoonbaert, Hartsuiker, & Pickering, 2007; Hartsuiker & Pickering, 2008) psycholinguistic model of bilingual processing to show that the change in frequency due to Dutch influence is likely to start at the (specific) lemma level, and then to extend to a schematic (less specific) level of representation.

As briefly mentioned, I consider cross-linguistic influence as a process, and convergence and divergence as the outcomes of this process. But what does it really mean that a language influences another language? The process of cross-linguistic influence in heritage languages can have three main concrete manifestations: change in frequency between two equally possible features (Section 1.3.1.1), loss or reduction of a feature (Section 1.3.1.2), or grammatical reanalysis of a feature (Section 1.3.1.3). These three manifestations roughly correspond to the three types of cross-linguistic influence individuated by Alferink (2015), namely redistribution (change in frequency), reduction (loss), and accumulation (grammatical reanalysis). In the following section, I discuss each of these manifestations in more detail.

1.3.1.1 Change in frequency

The first form that cross-linguistic influence in heritage languages can take is a change in frequency or preference between two (or more) available options for the option that is also shared by the dominant language. This type of change, namely a change involving only the frequency distribution of already existing constructions, has been referred to as ‘indirect transfer’ by Silva-Corvalán (1994, p. 4), as ‘frequential copying’ by Johanson (2002, p. 292), and as ‘redistribution’ by Alferink (2015, p. 17). If the heritage language has two (or more) equally possible options, heritage speakers will prefer the option also present in the dominant language. Alternatively, they will turn a pragmatically marked option into a pragmatically unmarked one to match the frequency of that option in the dominant language. This

type of change is referred to as “frequential copying” by Johanson (2002, p. 292), who describes it as follows:

Frequential copying means that frequency patterns peculiar to model code [DL] units are copied onto units of the basic code [HL] so that the latter undergo an increase or a decrease in frequency of occurrence. For example, elements which already exist in the basic code [HL], though they are more “normal” in the model code [DL], may gain ground and become less marked.

Changes in frequency are driven by some structural surface similarity between the dominant language and the heritage language. In other words, there must be some isomorphism between the dominant language and the heritage language. In order to map the frequency of a structure in the dominant language to the ‘same’ structure in the heritage language, bilingual speakers need to identify the two structures, a condition that is met if the two languages are similar at the surface level. Once speakers identify some kind of equivalence relation between a heritage structure and its ‘corresponding’ dominant structure, they start using the heritage structure with a corresponding dominant equivalent more and more frequently and up to the point that it may become the only option (Johanson, 2002; see Section 7.2.2). Structural similarity between the two languages has been proposed as one of the constraints regulating transfer in heritage contact situations (Silva-Corvalán, 1994, 2008; Muysken, 2013); I come back to this in Section 1.4.1 when I discuss the Alternation Hypothesis.

Backus (2004, p. 180) classifies “changes in the use of existing constructions, such as preferential use of some structures over other options” as system-preserving, and contrasts them with system-altering changes, which alter the structure of the heritage language by adding or subtracting of a category. The distinction between system-preserving and system-altering changes echoes the debate around the definition of contact-induced change reported in Onar Valk (2015). Some scholars are critical in regarding a change in frequency as a type of contact-induced change (because no new feature is introduced in the language). An increasing number of scholars, however, acknowledge that a change in frequency can lead to deep structural changes because it affects the level of entrenchment of a particular structure in the bilingual mind (see Section 7.2.2). Furthermore, change in frequency is perhaps the most common type of change in contact settings (Backus, 2004). For instance, in his article on word order changes in contact situations, Heine (2008, p. 54, emphasis mine) states that:

Of all the factors discussed in this paper, *the most pervasive effects on grammatical replication can be seen in the extension of existing structures to new contexts and in an increased frequency of use.* [...] What frequently happens is that speakers draw on a minor use pattern – one that has a more marginal status, being used rarely and/or only in specific contexts only to build a new major use pattern by increasing the frequency of use and extending the range of contexts in which it may occur.

Changes in frequency increase the surface structural similarity between the dominant and the heritage language, thus opening the door for further influence between the two. This is labeled by Enfield (2003, p. 356) ‘the self-perpetuating process’ of language change, whereby structural transfer “naturally increases the structural compatibility of the languages, thereby increasing the likelihood of further common structural borrowing”. Thus, change in frequency is an essential step in the process of contact-induced change.

Examples of change in frequency are abundant in heritage languages (Silva-Corvalán, 1994, 2008; Boumans, 2006; Moro, 2014; Moro & Klamer, 2015; Onar Valk, 2015; Moro & Irizarri van Suchtelen, forthcoming). For instance, heritage Spanish allows an alternation between Subject-Verb (SV) or VS order with unaccusative verbs. According to Silva-Corvalán (1994), as a result of transfer from English, Spanish-English bilinguals use the SV order with a higher frequency than their monolingual peers. Similar results are presented in Onar Valk (2015), who reports data on Dutch-Turkish bilingual heritage speakers. She argues that bilinguals prefer the verb-medial order over the verb-final order in main clauses because the former is also possible in Dutch. Onar Valk (2015, p. 265) concludes that “Dutch Turkish has undergone some contact-induced change, in the form of a ‘change in preferences’ or a ‘change in frequency’, at least regarding these structures [subordination, reported speech and the matrix verb position in complex clauses]”. Another example comes from the Moroccan heritage community in the Netherlands. Heritage speakers of Moroccan Arabic show an increased preference for the analytic possessive construction (Possessed + *dya* ‘of’ + Possessor) over the synthetic construction (Possessed + Possessor), when compared to their peers in Morocco. This change in frequency, according to Boumans (2006, p. 213), “suggests a direct influence of Dutch as the socially dominant language”.

The previous examples have all been described as cases of cross-linguistic influence. However, another explanation is also possible, namely that the attested changes are due to universal principles leading to simplification, rather than due to cross-linguistic influence (see also Section 1.3.5). In fact, the language may simplify

its system by changing from a flexible-word order to a rigid word order and from synthetic constructions to analytic ones (Boumans, 2006; Polinsky & Kagan, 2007; Onar Valk, 2015). This shows that, in many cases, it is difficult to distinguish between transfer and universal principles of language development in contact settings on the basis of the data because both processes yield the same outcome (Boumans, 2006; Polinsky & Kagan, 2007; Benmamoun et al., 2010). The interaction between cross-linguistic influence and other sources of heritage grammar divergence are touched upon in the following sections (Section 1.3.1.2 and Section 1.3.3.1).

To conclude, changes in frequency in heritage languages are usually driven by structural surface similarities with the dominant language and affects structures that are already available in the heritage language. I come back to this claim in Section 1.4, where I illustrate the research hypotheses adopted in the present research, and in Chapter 3, Chapter 5 and Chapter 6, where I show that cross-linguistic influence from Dutch onto heritage Ambon Malay manifests itself as a change in the frequency of two or more structures already available in Ambon Malay and that this shift in frequency is leading heritage Ambon Malay to convergence toward Dutch.

1.3.1.2 Loss or reduction

Cross-linguistic influence from the dominant language can also manifest itself as a reduction in the frequency or loss of certain linguistic features. This leads to convergence with the dominant language and to divergence from the homeland language, with respect to which the heritage grammar seems ‘incomplete’ or indeed ‘reduced’ (Polinsky, 2006; Montrul, 2009, 2010; Benmamoun et al., 2010, among others). An example of dominant language transfer manifesting itself as loss or reduction comes from heritage speakers of Spanish in the U.S.A. Montrul (2010) studied transfer effects on features that have no equivalent in English, such as (D)ifferential (O)bject (M)arking, the overt morphological marking of animate direct objects which is expressed by the preposition *a* (e.g., *Ayer vi a María* ‘Yesterday I saw DOM Maria’). The data show that heritage speakers produce errors with DOM, incorrectly omitting *a* with animate and specific direct objects. Montrul (2010, p. 309) concludes that “this result can easily be attributed to transfer from English, since English does not mark animate direct objects overtly with morphology”. According to Montrul (2010), little degree of structural similarity or overlapping between the two languages (Spanish and English in this case) could be one of the factors triggering transfer. The rationale is that, when the dominant language does

not instantiate certain syntactic properties, the absence of these syntactic properties is transferred to the heritage language.

Despite the convincing argumentation of Montrul (2010), here again we could argue that contact phenomena, such as DOM omission, could be due to universal principles of language development in contact settings rather than to transfer. As already pointed out in the previous section, it is difficult to distinguish between transfer and universal principles, such as simplification, when these two processes yield the same outcomes. This point is investigated thoroughly by Jarvis and Odlin (2000) in their study on Finnish-speaking learners of English. With respect to spatial reference, English and Finnish are structurally quite dissimilar: English uses prepositions, while Finnish has a case system and no prepositions. The results of Jarvis and Odlin's (2000) study show that Finnish-speaking learners tend to omit prepositions when speaking in English. The authors interpret these results in light of transfer and simplification, and conclude (2000, p. 550):

Although the use of a zero preposition is a form of linguistic simplification, its use by the Finns also constitutes a form of transfer, given that the structural nature of the Finnish locative cases predisposes Finns to disregard proposed function words as relevant spatial markers. Thus, the Finns' omission of spatial prepositions in English seems to arise out of an interaction between simplification and transfer.

Understanding the interaction between cross-linguistic processes and other, universal principles, such as simplification, is but one of the important issues for future research on structural transfer in heritage languages. Benmamoun et al. (2010, p. 42) point out that:

Ideally, studies of the same heritage language with different contact languages should be undertaken to investigate the extent to which transfer from the dominant language influences the degree of divergence and simplification found in heritage language grammars.

To empirically test the role of transfer and simplification one would need to investigate the same heritage language in contact with two different dominant languages, one that overtly marks the feature under investigation, and one that does not. If heritage speakers are found to omit the feature in both scenarios, transfer can be said to play no role, and simplification can be argued to be the only cause. If *vice versa*, heritage speakers are found to make omission errors only when the dominant language does not instantiate the feature, transfer can be argued to be the most

probable cause. Another test would be to compare the heritage language against all other varieties of the homeland language (the baseline), for instance the variety spoken by L2 learners. If simplification is a universal principle it should take place in all the contact varieties of the same language.

For now, it seems that we need to limit ourselves to the observation that simplification or other universal principles act in a cumulative way with transfer (see Section 1.3.6). This observation is discussed in Chapter 5 in relation to the increase of double object constructions in heritage Ambon Malay, and I come back on the issue of universal principles in Section 1.3.5.

1.3.1.3 Grammatical reanalysis

Cross linguistic influence in heritage languages can also take the form of grammatical reanalysis, meaning that a form or a structure of the heritage language is reinterpreted on the model of the dominant language, thus leading to convergence between the two languages. Alferink (2015, p. 18) points out that, in this type of cross-linguistic process, increased similarity is achieved “by adding the specificity of one language to the other in some way, resulting in a cumulative bilingual system”. This type of cross-linguistic influence presupposes that speakers establish some kind of equivalence relation between linguistic signs and linguistic categories in their two languages (Winford, 2003; Heine & Kuteva, 2005; Matras, 2009; Gast & van der Auwera, 2012; see also Section 7.2.2).

Contact-induced grammatical reanalysis is also known as contact-induced grammaticalization, a process whereby speakers replicate grammatical structures or grammatical categories of another language by using material available in the replica language [heritage language] and grammaticalizing it into structures corresponding to those of the model language [dominant language] (Heine & Kuteva, 2008, p. 71; see also Section 1.4.2). The general mechanism behind contact-induced grammaticalization is schematized as follows:

- a. Bilingual speakers notice that in the model [dominant] language there is a semantic or a grammatical category (e.g., Tense).
- b. They create an equivalent category in the replica [heritage] language on the basis of the use patterns available in the heritage language (e.g., using an aspect marker).

Although Heine and Kuteva (2005) perhaps put too much emphasis on the meta-linguistic awareness of the speakers, it is undoubtedly true that speakers are the initiators of any type of contact-induced change, including contact induced grammaticalization. According to Sánchez (2004, 2006), bilingual speakers have the ability to map grammatical features of one language onto the morphological units of another language. This process is possible because grammatical features can be dissociated from their morphological counterparts in the mind of some bilingual individuals (see Section 1.4.1).

Contact-induced grammaticalization, that is, grammaticalization due to contact, is identified on the basis of the symptoms (a-e) listed below (Heine & Kuteva, 2005, p. 80).

- a. Extension, i.e. the rise of new grammatical meanings when linguistic expressions are extended to new contexts.
- b. Desemanticization (or “semantic bleaching”), i.e. loss (or generalization) in meaning content.
- c. Decategorialization, i.e. loss in morpho-syntactic properties characteristic of lexical or other less grammaticalized forms.
- d. Erosion (“phonetic reduction”), i.e. loss in phonetic substance.
- e. Obligatorification (or increase in the frequency of a form).

In Heine and Kuteva’s (2005) original list, obligatorification is not included as a parameter. In a later publication (2007), however, the authors acknowledge that obligatorification is a by-product of decategorialization, and, I would add, also of semantic extension. “Frequency, [in fact], is enhanced by semantic generality [...], which grants the compatibility of a marker with a large number of lexical items” (Bisang, 2011, p. 115). I included obligatorification in the list above because increase in the frequency of a given form is one of the symptoms of contact-induced grammaticalization that is discussed in Chapter 3 and Chapter 4 of the present study. Note that, grammatical reanalysis can lead to an increase in the frequency of a certain linguistic feature, but unlike change in frequency discussed in Section 1.3.1.1, it also involves a change in the meaning of the feature.

Contact-induced grammaticalization is a long and gradual process that takes place in long-term contact situations and usually extends over hundreds of years. The heritage language contact situation is a recently established contact setting, whereby two languages have been in contact for about 50-60 years. This contact

scenario is too short to allow the full grammaticalization process to reach completion. Nevertheless, in heritage languages we can find some of the symptoms listed above (a-e) and observe the incipient stages of contact-induced grammaticalization (Backus et al., 2011). Obviously we do not know whether the ‘on-going’ changes observed will lead to full grammaticalization later on, but we can at least document the early stages of it. For instance, Backus et al. (2011, p. 745) report that heritage Turkish in the Netherlands “has undergone grammaticalization processes due to Dutch influence, [but] these processes are not completed yet”. Examples are the semantic and context extension of the indefinite article *bir* ‘a’ and of the verb *yapmak* ‘do’, which are used on the model of their Dutch counterparts. That is, *bir* ‘a’ follows the distribution of the Dutch indefinite article, and thus occurs also with specific nouns (a possibility not present in Turkey). The verb *yapmak* ‘do’ is found in typical Dutch expressions, such as *ilkokul yapmak* ‘elementary school do’, rather than the homeland expression *ilkokul bitirmek* ‘elementary school finish’. Interestingly, Backus et al. (2011, p. 745) observe that the processes of contact-induced grammaticalization taking place in heritage Turkish involve extension, desemantization and increase in frequency, but there are no signs of decategorialization and erosion. As shown in Chapter 3 and Chapter 4, similar conclusions can be drawn for heritage Ambon Malay, whereby embryonic grammaticalization seems to mainly involve semantic extension and increase in frequency.

1.3.2 Incomplete acquisition

The second source of heritage language grammar divergence is incomplete acquisition. A grammar is regarded as incomplete “when it fails to reach age-appropriate linguistic levels of proficiency as compared with the grammar of monolingual or fluent bilingual speakers of the same age, cognitive development, and social group” (Montrul & Bowles, 2009, p. 363). Since heritage children start to be exposed to the dominant language by the age of five (at the latest), the amount of input and use of the heritage language drastically diminishes by that age. Consequently, the grammatical development of the heritage language begins to lag behind and the heritage grammar ends up to be incompletely acquired (Montrul, 2008).

Benmamoun et al. (2010, p. 44) correctly point out that the diagnostic of incomplete acquisition is a language feature that causes problems to both adult

heritage speakers and to children at the age of five and upwards. If a linguistic feature is equally problematic for both groups, there is a great likelihood that this feature has never reached the final stage of acquisition in the heritage grammar. Incomplete acquisition, thus, places the emphasis on the heritage language only and predicts divergence between the heritage system and the homeland system but no convergence between the heritage system and the dominant language system. In this respect, it differs from cross-linguistic influence, which focuses on the typology of the dominant language and predicts both divergence from the homeland variety and convergence to the dominant language (see Section 1.3.1). Although incomplete acquisition and cross-linguistic influence are distinct processes, they can also interact, in that cross-linguistic influence can affect the rate of incomplete acquisition. *Vice versa*, incompletely acquired structures can be more vulnerable to cross-linguistic influence (Montrul, 2010). Incomplete acquisition manifests itself in two ways: reduction (Section 1.3.2.1) and grammatical reanalysis (Section 1.3.2.2).

1.3.2.1 Loss or reduction

Incomplete acquisition mainly manifests itself as some kind of reduction or loss, such that the heritage language system appears to be incomplete or reduced with respect to the homeland system. A clear example of incomplete acquisition is provided by Polinsky (2008a), who investigates gender assignment in heritage Russian as spoken in the U.S.A. The results of the study show that heritage speakers make errors with certain feminine nouns and with neuter nouns by assigning them masculine gender. Since English has no gender, transfer is assumed to play no role in determining these errors. Polinsky (2008a) considers these errors as manifestations of incomplete acquisition because the nouns that are problematic for heritage speakers are equally problematic for L1 monolingual acquirers in Russia. This suggests that heritage speakers did not complete the acquisition process (due to the lack of input after the age of five); hence, their grammars show deficiencies in this respect. Another example comes from heritage Spanish in the U.S.A. Montrul (2009) investigates the grammatical knowledge of aspect and mood among heritage speakers and finds that heritage speakers have unstable knowledge (which is a sign of incomplete acquisition) of the subjunctive mood. Montrul (2009) concludes that, since monolingual children do not use and understand correctly the subjunctive until close to adolescence, the high error rate in heritage speakers can be accounted for by the missing development of this feature in the adults' heritage grammar due to the reduced input condition at the age of five and upwards. In addition to gender

classification (Polinsky, 2008a) and tense, aspect, and modality (Polinsky, 2008b; Montrul, 2009; Laleko, 2010), other domains have been found to be vulnerable to incomplete acquisition, namely complex syntax (e.g., long-distance dependencies such as relativization) and the syntax/discourse interface (e.g., overt vs. null subject pronouns) (Polinsky & Kagan, 2007; Montrul & Bowles, 2009; Benmamoun et al., 2010; O’Grady et al., 2011; Polinsky, 2011; Sorace, 2011, Laleko & Polinsky, 2013, among others).

1.3.2.2 Grammatical reanalysis

Incomplete acquisition can also lead to reanalysis. The only example that I am aware of, where reanalysis can be convincingly related to incomplete acquisition and not to transfer is provided by Polinsky (2008a). We have seen in the previous section that gender in American Russian has undergone a reduction from a three-gender system (masculine, feminine, neuter) to a two-gender system (masculine and feminine). The category of gender has not only been reduced, but it is also subject to reanalysis among the less proficient heritage speakers: feminine nouns ending in a palatalized consonant are consistently treated as masculine, while neuter nouns are consistently treated as feminine (Polinsky 2008a). Since English almost completely lacks the category of gender, the reanalysis of the American Russian gender system can only be ascribed to incomplete acquisition.

1.3.3 Attrition

Another source of heritage language grammar divergence is attrition. “Attrition implies that a grammatical system had a chance to develop completely and remained stable for a while before some grammatical aspects eroded later on, as a heritage speaker was using his/her language less and less” (Benmamoun et al., 2010, p. 46). The outcome of heritage grammar attrition is divergence because the heritage grammar ‘looses’ linguistic features of the baseline homeland grammar (Montrul 2004, 2005; Montrul & Bowles, 2009; Benmamoun et al. 2010; Polinsky, 2011; Montrul, Bhatt, & Bhatia 2012; Montrul & Sánchez-Walker, 2013). The diagnostic for attrition are those linguistic features that are fully acquired by children at the age of five. “If an adult heritage speaker experiences problems with such language properties there is a great likelihood that these properties underwent attrition and became weaker over the speaker’s lifespan” (Benmamoun et al., 2010, p. 47). For instance, Polinsky (2011) found that, with respect to comprehension of subject and

object relative clauses in Russian, heritage children and monolingual homeland children perform like monolingual homeland adults, whereas adult heritage speakers show a degraded performance on the interpretation of both types of relative clauses. The fact that children and adult perform alike suggests that adult heritage speakers have probably acquired the grammar of relative clauses, but have subsequently ‘lost’ it due to lack of exposure to and use of the language.

Attrition and incomplete acquisition are theoretically two distinct notions and make different predictions, as reported by Polinsky (2011, p. 306):

- a. Incomplete acquisition: if a child and an adult deviate from the baseline in the same way, it can be assumed that the feature has not been acquired.
- b. Attrition: if a child performs as his or her age-matched baseline control but the adult does not, the feature can be assumed to have been acquired but may have subsequently been lost or reanalyzed.

In practical terms, however, it can be difficult to distinguish between attrition and incomplete acquisition due to the lack of acquisition data in the child control population (Montrul & Bowles, 2009; Polinsky, 2011). Furthermore, attrition and incomplete acquisition are by no means mutually exclusive. Most of the studies on heritage languages, thus, conflate attrition and incomplete acquisition and treat them as one single category (Montrul & Bowles, 2009). Similarly to incomplete acquisition, attrition also manifests itself as reduction or loss (Section 1.3.3.1) and as grammatical reanalysis (Section 1.3.3.2). Given the scarcity of longitudinal data, it is hard to find heritage language studies that describe loss or reanalysis as unequivocally related to attrition.

1.3.3.1 Loss or reduction

Attrition mostly manifests itself as the “loss of restrictions on the application of rules [..., and it] involves the replacement of formally more complex and more narrowly distributed rules by formally less complex rules with wider distribution” (Sorace, 2005, p. 67). An example is provided by Montrul (2004), who shows that heritage Spanish speakers in the U.S.A. overgeneralize overt subject pronouns to contexts that require a null subject. These bilinguals seem to have lost the discourse-pragmatic constraints governing the overt subject distribution, such as known versus change of referent. Hence, they tend to produce redundant pronominal subjects when there is

no change of referent. Features that belong to the interface between syntax and other domains, such as the lexicon, discourse, or pragmatics have proven to be vulnerable to the effects of attrition in heritage language speakers (Montrul, 2004; 2005; Polinsky, 2011; Sorace 2011; Montrul & Sánchez-Walker, 2013).

According to Sorace (2011), attrition can interact with cross-linguistic influence and lead to the “neutralization of native distinctions toward the less restrictive L2 option”. In other words, speakers’ knowledge of rules and constraints begins to erode under the effect of the dominant language, which has less restrictive rules. A similar point is made by Montrul and Sánchez-Walker (2013, p. 127). The authors argue that DOM in heritage Spanish in the U.S.A. “is highly vulnerable to attrition, especially when Spanish is in contact with a language that does not mark objects overtly like English”. This relates to the change in frequency described in Section 1.3.1.1, namely the preference of heritage speakers for one encoding option over another equally plausible encoding option if the preferred option is also possible in the dominant language. Most of the changes in frequency in heritage languages are probably accompanied by some kind of attrition (or incomplete acquisition). The loss or neutralization of the discourse-pragmatic constraints leads heritage speakers to select the option also present in the dominant language (e.g., overt subject, SVO order) regardless of the constraints holding for the homeland language.

1.3.3.2 Grammatical reanalysis

Besides leading to the loss or the neutralization of certain rules, attrition can also lead to reanalysis. Polinsky’s (2011) article *Reanalysis in adult heritage language: A case for attrition* is the only study that I am aware of that actually teases apart the effects of incomplete acquisition and attrition in adult heritage speakers by using data from child language acquisition. The study compares the comprehension of subject and object relative clauses by four groups of Russian speakers (monolingual adults/children, heritage adults/children). The findings show that the children in both groups perform like monolingual adults, while heritage adult speakers perform at chance on object relatives and show a degraded performance on subject relatives. The native performance of heritage children undoubtedly suggests that the degraded performance in adult heritage speakers is a case of attrition rather than of incomplete acquisition. According to Polinsky (2011), the erosion of the relative clause system is accompanied by some kind of reanalysis performed by adult heritage speakers which leads them to limit relativization to the subject position.

Furthermore, since English allows both subject and object relativization, this reanalysis can only be ascribed to attrition and not to cross-linguistic influence.

The findings presented in this section and in the previous section show that attrition can interact with cross-linguistic influence but does not need to. Finally, in the heritage Russian case, reanalysis seems to rely upon universal principles of relative clause formation favoring subject relativization (Polinsky, 2011, p. 322). We have already seen in Section 1.3.1.1 and in Section 1.3.1.2 that universal principles of language development in contact situations may guide some of the contact-induced changes observed in heritage languages, an issue that I revisit in Section 1.3.5.

1.3.4 Type of input

Another possible source of divergence between the heritage grammar and the homeland grammar is the type of input heritage speakers receive. The input in the heritage language is not only quantitatively different from the input of monolinguals in the homeland, but also qualitatively different. In fact, heritage speakers are mainly exposed to colloquial varieties, and use the heritage language in a restricted number of informal domains (family, friendship). Furthermore, they rarely get formal classroom education in the heritage language. According to Chevalier (2004, p. 43), the gradual narrowing of registers among heritage speakers results from their shift to the dominant language:

As the linguistic repertoire in English [or any other dominant language] expands to include an increasing number of domains, the home-based language contracts, its functional use restricted to fewer domains, until it is ultimately limited to the home and family domain. A family, homebound language is characterized by a casual, conversational speech style, used with familiar interlocutors to a restricted set of topics focused on everyday life.

Qualitative differences in the input can obviously bring about both qualitative and quantitative differences in the output. Some linguistic features might be more frequent in the informal language than in the formal language. Since heritage speakers mainly use the language in informal settings, their language may show a higher frequency of these informal features compared to homeland speakers. Contrarily, some linguistic features might be more frequent in the formal and written language, or might be learned only via schooling. And since heritage speakers are not schooled in the heritage language, they may lack the knowledge of

these features. We have seen in Section 1.2.1 the example provided by Pires and Rothman (2009), showing that heritage speakers of Brazilian Portuguese lack knowledge of inflected infinitives because this linguistic feature is acquired only by means of formal education, which heritage speakers never get.

Another example of divergence which is due to the difference in the input comes from heritage Spanish in the U.S.A. Torres Cacoulios (2000) reports that the varieties of Spanish in the Southwest are characterized by popular oral features, such as a high frequency of the progressive *estar + -ndo*, roughly similar to the English *-ing*. She argues that the high frequency of *estar + -ndo* among bilinguals is not due to cross-linguistic influence from English, as previously argued by Klein (1980) and Silva-Corvalán (1994), but to register factors. The topics focused on in everyday situations require more often *estar + -ndo* than the topics focused on in written or formal language. Torres Cacoulios (2000) warns that the effects of language contact may only be an epiphenomenon, the real cause being the different registers monolinguals and bilinguals are exposed to. The truth probably lies in between and it is likely that contact with English might have accelerated a change already present in Spanish. I present a similar case in Chapter 3, where I argue that the high frequency of DEMONSTRATIVE-NOUN order in heritage Ambon Malay results from both qualitative differences in the input heritage speakers received and from contact with Dutch. Another example showing how differences in the input can have qualitative effects on the language of heritage speakers is presented in Chapter 5.

1.3.5 Universal principles

Divergence from the homeland grammar may also be accounted for by universal principles of language development in the context of language disuse (Laleko, 2010, p. 33-35). Some changes in heritage languages are independent of the structure of the two languages involved, and seem to be motivated by “universal regression processes or simplification under reduced input conditions (as attested in the case of creole genesis)” (Benmamoun et al., 2011, p. 53). The restructuring of the heritage language system, thus, may be partially governed by universal autonomous processes that may or may not interact with cross-linguistic influence. Seliger and Vago (1991, p. 12) suggest that in attrited grammars “some rules are transferred between the existing grammars available to the speaker, while others appear to derive from innate or universal principles of language acquisition”. Polinsky (2008c,

p. 161) comes to similar conclusions with respect to heritage language grammars, which are “presumably shaped by the interference from English [or any other dominant language] and some universal principles governing language development with limited input”.

But, what are exactly these universal principles? One principle of language development in contact settings is simplification. This term covers phenomena such as loss or reduction, regularization of paradigms, and preference for certain types of structures. For instance, the literature suggests that universal principles of human communication favor unmarked, less complex structures (such as SVO word order) over marked ones, or analytic constructions (such as finite subordination, analytic genitive) over synthetic ones. Heritage speakers of Russian, for instance, have a strong preference for subject relativization even though object relative clauses are present in their dominant language, English. Chapter 5 of this thesis reports the case of heritage speakers of Ambon Malay, who use (D)ouble (O)bject constructions (*John gave Mary a book*) more frequently than both their homeland and Dutch comparators. In the chapter, I argue that the frequency of DO constructions may result from an interaction of universal processes of language acquisition in contact situations and transfer from Dutch.

Universal principles underlie also what Enfield (2003, p. 361) calls ‘conceptual naturalness’, namely the set of “putative cognitive/conceptual universals (biologically based)”. According to Enfield (2003), some semantic extensions seem conceptually very ‘natural’ (such as the use of the word ‘fire’ for ‘light’), and are therefore more easily made than others; these extensions are usually very common cross-linguistically. For instance, in many languages, spatial prepositions are often used to indicate other, non-spatial meanings. Bowerman (2011, p. 599) reports that:

[C]hildren extend forms across semantic boundaries that must be honoured in their own language, but are collapsed in many other languages. For instance, learners of English sometimes overextend spatial morphemes to temporal meanings (e.g. BEHIND dinner to mean AFTER dinner).

This pattern is common in both polysemy and language change. Obviously, language internal changes driven by conceptual naturalness can be accelerated and/or reinforced by cross-linguistic influence. In Chapter 6 and in Chapter 8, I come back to the notion of ‘conceptual naturalness’ in relation to two ‘on-going’ changes in heritage Ambon Malay, and I show that transfer from Dutch is probably accelerating a language internal process.

Another example of heritage grammar divergence resulting from universal principles is reported in Aalberse and Moro (2014). The possessive construction ‘Possessor + *punya* ‘POSS’ + Possessed’ (e.g., *tikus punya kaki* ‘mouse POSS foot’ ‘the foot of the mouse’) has fully grammaticalized in the homeland variety of Ambon Malay where *punya* can be spelled as *pung/pu/ng* or zero (van Minde, 1997). In the heritage variety spoken in the Netherlands, the grammaticalization process has not proceeded further than the *pung/pu* stage and the zero variant is not attested. The authors argue that the halt in the grammaticalization of *punya* in the Netherlands setting relates to the low frequency and low predictability of words in the immigrant speech community and to the importance that heritage speakers attach to ease of perception over ease of articulation. The Aalberse and Moro’s (2014) findings are in line with the general observation that, since heritage speakers are unsure about their linguistic skills, they tend to select structures which convey the intended meaning explicitly (Laleko & Polinsky, 2013).

Finally, other universal strategies adopted by asymmetrical bilingual speakers include fronting well-known words to gain time while trying to retrieve a lexical item, or overusing connectives (e.g., ‘and so’, ‘and then’) to cover up for hesitations (Aalberse & Muysken, 2013).

To sum up, universal principles of language development in the context of language disuse, such as preference simplification, regularization, and naturalness, as well as universal principles of human communication, such as preference for phonologically long variants, word-fronting, etc. constitute yet another set of causes of divergence between the heritage grammar and the homeland grammar. Although universal principles are theoretically independent from cross-linguistic influence, the two may reinforce each other in shaping heritage language grammars.

1.3.6 Interim summary

The previous sections have illustrated the main sources of heritage grammar divergence, their manifestations, and the outcomes they yield. This information is summarized in Table 1.1 on the next page.

Table 1.1: Sources, manifestations, outcomes and examples of heritage grammar divergence.

SOURCE	MANIFESTATION	OUTCOME	EXAMPLES	IN THIS STUDY
Cross-linguistic influence	Change in frequency	Divergence from HL Convergence to DL	Shift in matrix verb position in Dutch-Turkish (Onar Valk, 2015)	Chapter 3, Chapter 5, Chapter 6
	Reduction or loss	Divergence from HL Convergence to DL	Omission errors with DOM in heritage Spanish (Montrul, 2010)	-
	Grammatical reanalysis	Divergence from HL Convergence to DL	Grammaticalization of the article <i>bir</i> 'a' in Dutch-Turkish (Backus et al., 2011)	Chapter 3, Chapter 4
Incomplete acquisition	Reduction or loss	Divergence from HL	Errors with subjunctive mood in heritage Spanish (Montrul, 2009)	-
	Grammatical reanalysis	Divergence from HL	Reanalysis of nouns in American Russian (Polinsky, 2008a)	-
Attrition	Reduction or loss	Divergence from HL	Loss of pragmatic constraints (Montrul, 2004)	-
	Grammatical reanalysis	Divergence from HL	Reanalysis of object relative clauses in American Russian (Polinsky, 2011)	-
Type of input	Qualitative and quantitative differences	Divergence from HL	Lacking knowledge of inflected infinitives in Brazilian Portuguese (Pires & Rothman, 2009)	Chapter 3, Chapter 5
Universal processes	Qualitative and quantitative differences	Divergence from HL	Preference for analytic structures (Boumans 2006; Onar Valk, 2015)	Chapter 3, Chapter 5, Chapter 6

The first source of divergence, cross-linguistic influence from the dominant language, is language-external because it is triggered by contact with another (dominant) language. The other three sources, namely incomplete acquisition, attrition, and type of input, are language-internal because they are solely related to the internal structure of the heritage language. The last source, universal principles, is neither internal nor external because the principles are independent of the structures of the two languages involved.

The reader needs to keep in mind that the chapters of this dissertation account for the divergence between heritage Ambon Malay and homeland Ambon only on the basis of cross-linguistic influence, type of input and universal principles. Incomplete acquisition and attrition are not examined further due to the lack of data on children L1 and adult L2 acquisition (see Section 2.1.1).

The sources of heritage grammar divergence listed in Table 1.1 are not mutually exclusive, and are likely to interact and reinforce each other in various ways. As seen in the previous sections, *cross-linguistic influence* can be reinforced by *universal tendencies* in promoting the SVO word order in heritage Turkish and heritage Spanish and analytic constructions in heritage Moroccan Arabic (Section 1.3.1.1). It can also interact with *attrition* and lead to the overgeneralization of overt subjects in heritage Spanish (Section 1.3.3.1). Finally *cross-linguistic influence* can act in a cumulative way with *type of input* and accelerate the change in favor of the *estar + ndo* constructions in heritage Spanish (Section 1.3.4). The other sources can also interact. The preference for subject relativization in American Russian, for instance, is most likely the result of *attrition* and *universal principles* (Section 1.3.3.2). To sum up, multiple causation is not only the most likely explanation for many of the contact phenomena listed above, but it is also the only explanation that we can provide until we identify criteria to disentangle all the possible sources (Torres Cacoullos, 2000; Benmamoun et al., 2010). Understanding the interaction between language-external, language-internal and universal processes is one of the important challenges for heritage language research.

1.4 Working Hypotheses

Previous research on language contact in general, and on heritage languages in particular, has led to the formulation of a number of hypotheses and predictions concerned with the direction, type and outcome of language change in various

contact situations. Illustrating all the hypotheses proposed in contact linguistics is beyond the scope of the present chapter, therefore I limit myself to the hypotheses that are relevant for the following chapters. The hypotheses that are considered here are the Alternation Hypothesis (Jansen, Lalleman, & Muysken, 1981), the Vulnerability Hypothesis (de Prada Pérez, 2015), the Functional Convergence Hypothesis (Sánchez, 2004), the Conceptual Hypothesis (Schoenmakers-Klein Gunnewiek, 1997), and the Conceptual Transfer Hypothesis (Bylund & Jarvis, 2011). Section 1.4.1 discusses the Alternation Hypothesis, the Vulnerability Hypothesis, and the Functional Convergence, while section 1.4.2 discusses the Conceptual Hypothesis and the Conceptual Transfer Hypothesis. The Alternation Hypothesis and the Vulnerability Hypothesis adopt linguistic considerations in order to predict the nature of some of the changes that may occur in heritage languages, while the Conceptual Hypothesis and the Conceptual Transfer Hypothesis rely more on cognitive considerations. The Functional Convergence Hypothesis adopts both linguistic and cognitive considerations.

1.4.1 The Alternation Hypothesis, the Vulnerability Hypothesis and the Functional Convergence Hypothesis

The Alternation Hypothesis was first proposed by Jansen, Lalleman & Muysken (1981) in the field of L2 acquisition. The hypothesis states that: “when the target language offers an alternation between two patterns [...] a second language learner will tend to overgeneralize the pattern existing in his or her first language” (1981, p. 315). For example, if the target language allows an alternation between verb-medial and verb-final, L2 learners of a verb-final language will overgeneralize the verb-final pattern when speaking the target language. In other words, the strategy adopted by these bilinguals is to select among the existing alternatives in the target language the one matching the alternative in their L1 (Heine, 2008; Muysken, 2013). The Alternation Hypothesis, thus, assumes that cross-linguistic influence occurs where the two languages present some degree of overlap, and consequently predicts that the nature of the change depends on the typology of the two languages in contact. The predictions of this hypothesis are found to hold true also for heritage speakers (Silva-Corvalán, 1994; Boumans, 2006; Onar Valk & Backus, 2013; Moro, 2014; Moro & Klamer, 2015; Alferink, 2015; Onar Valk, 2015). Silva-Corvalán (2008, p. 215) observes that, in a situation of extensive and intensive bilingualism, “it seems logical to expect that frequently used patterns in the socially dominant language will

motivate an increase in the frequency of use of parallel structures in the subordinate [heritage] language”. Hence, if applied to heritage speakers, the Alternation Hypothesis would need to be rephrased as follows: when the heritage language offers an alternation between two patterns, heritage speakers will tend to overgeneralize the pattern existing in his or her dominant language.

The Alternation Hypothesis involves two prerequisites: (i) the L2 (or the heritage language) presents some kind of variability (alternation between structures) within certain areas of its system; (ii) the L1 (or dominant language) and the L2 (or heritage language) need to share some degree of similarity in certain areas of their grammatical systems. Notably, variability is the central notion of the Vulnerability Hypothesis (de Prada Pérez, 2015), while partial similarity is the central notion of the Functional Convergence Hypothesis (Sánchez, 2004, 2006).

The Vulnerability Hypothesis establishes a categorical-variable continuum of permeability, such that variable phenomena are permeable while categorical phenomena are not (de Prada Pérez, 2015). The rationale behind this hypothesis is that variability is difficult for bilingual (heritage) speakers to acquire, and consequently it becomes the target of possible dominant language influence. De Prada Pérez (2010, p. 114) suggests that “those areas identified as more vulnerable to interlingual interference can be accounted for by the variability present in the monolingual grammar”. Adopting a usage-based perspective, one may speculate that variability becomes the target of incomplete acquisition and/or transfer because variable structures are less entrenched than categorical ones. If a structure has no rival, it is better entrenched in the speaker’s repertoire, and any rival would be ungrammatical. Variable structures, on the other hand, are less entrenched and therefore more open to change.

The Functional Convergence Hypothesis states that convergence in bilingual grammars is favored by partial similarity across the two languages (Sánchez, 2004, 2006). This hypothesis focuses on grammatical categories (e.g., TMA), rather than on syntactic structures. The hypothesis predicts that bilingual speakers will map grammatical features from one language onto the morphological units of the other language, if the two languages have partially similar features associated with the same grammatical category. For example, in Quechua evidentiality features are associated with tense, while in Spanish tense is linked to aspectual features. Sánchez (2004, 2006) shows that Quechua-Spanish bilinguals have mapped the evidentiality features on the Spanish tense system, so that the Spanish pluperfect is used by some bilinguals to convey the reportative feature. The association of a feature to a new

morphological unit takes place independently from the different morphological forms used in each language (an auxiliary in Spanish, a suffix in Quechua). This shows that the re-association process is constrained by the syntactic structures of the receiving language. Interestingly, Sánchez (2006) proposes that the obligatory nature of the mapping might explain why some features are more prone to be imported than others. Finally, the psycholinguistic process of functional convergence seems to be at the basis of long-term contact-induced changes, such as contact-induced grammaticalization (see Section 1.3.1.3).

To sum up, the Alternation Hypothesis, the Vulnerability Hypothesis, and the Functional Convergence Hypothesis predict that the areas vulnerable to contact-induced change are those that present some degree of variability and similarity across the two languages. These hypotheses, however, differ with respect to the direction of the change. The Alternation Hypothesis and the Functional Convergence Hypothesis focus on the language pair and predict change in the direction of the dominant language. The Vulnerability Hypothesis, by contrast, focuses on the permeability continuum and predicts simplification of variable phenomena regardless of the dominant language (the dominant language can, however, accelerate this internal change).

In the following chapters, I have adopted these hypotheses and tested the vulnerability of variable features to contact-induced change. The findings of Chapter 3, Chapter 5, and Chapter 6 show that the alternation between two or more syntactic structures creates the conditions for cross-linguistic influence. This influence manifests itself as an increase in the frequency of the structure that is also possible in Dutch. The findings of Chapter 4 show functional convergence in the tense-aspect system of Ambon Malay heritage speakers.

1.4.2 The Conceptual (Transfer) Hypothesis

The basic tenet of the Conceptual Transfer Hypothesis is that patterns of conceptual categorization can be transferred from one language to another (Schoenmakers-Klein Gunnewiek, 1997; Bylund & Jarvis, 2011). The Conceptual Hypothesis formulated in Schoenmakers-Klein Gunnewiek (1997) maintains that patterns of conceptualization attached to a word can be transferred from one language to the other without the transfer of the word. Heritage speakers of Dutch in Brazil, for instance, have been found to use the Dutch verb *pakken* 'to take' following the conceptualization pattern of Portuguese (Schoenmakers-Klein Gunnewiek, 1997). In homeland Dutch, the verb

pakken ‘to take’ is sensitive to the feature of intention and control, so expressions like *de trein pakken* ‘to take the train’ are allowed, while expressions like *een ziekte pakken* ‘to take a illness’ are disallowed. In heritage Dutch, *pakken* ‘to take’ has adopted the conceptual categorization of its Portuguese equivalent *pegar* ‘to take’ (which is not sensitive to the feature of intention) and therefore it can be used also with a non-agentive meaning, as in the expression *een ziekte pakken* (to take an illness).

The Conceptual Transfer Hypothesis formulated in Bylund and Jarvis (2011, p. 47) adopts a broader perspective to include also patterns of event construal:

The Conceptual Transfer Hypothesis assumes that speakers of different languages have somewhat differing patterns of conceptual categorization and construal, and that, in the case of bilinguals and second language learners, these types of conceptualization differences have the potential to transfer across languages – or, more precisely, the conceptual distinctions and patterns of conceptualization that they have acquired as speakers of one language can also affect their use of another language.

The Conceptual Transfer Hypothesis has been tested in various studies, the majority of which focus on aspect and motion events (von Stutterheim & Nüse, 2003; Flecken, 2010; Bylund & Jarvis, 2011; Schmiedtová et al., 2011). The results of these studies seem to validate the predictions of the hypothesis, showing that conceptualization patterns in bilingual speakers are indeed vulnerable to cross-linguistic influence, and that transfer can lead to an increase or decrease in the use of grammatical elements. Flecken (2010) and Bylund & Jarvis (2011) studied cases in which a particular conceptual distinction is present and productive in one language of a bilingual, but not in the other. The studies demonstrate that, if bilingual speakers are likely to attend and refer to a certain concept or pattern in their dominant language, they will do so also when speaking the other language. *Vice versa*, if they do not normally verbalize the concept in the dominant language, they will tend to disregard it also in the other language. For instance, Spanish speakers have the tendency to attend to the ongoingness of the events because Spanish has a grammaticalized progressive/imperfect contrast. Swedish speakers, by contrast, have the tendency to refer to the endpoint rather than the ongoingness of events because Swedish lacks the grammaticalized progressive aspect. Spanish-Swedish bilinguals living in Sweden have been found to use progressive aspect to a lesser extent than Spanish monolinguals and to encode endpoints more frequently. In light of these findings, Bylund and Jarvis (2011, p. 58) conclude that “bilinguals are affected by the

Swedish tendency to construe events with maximal time schemas, with the consequence that they are less attentive to features of ongoingness”.

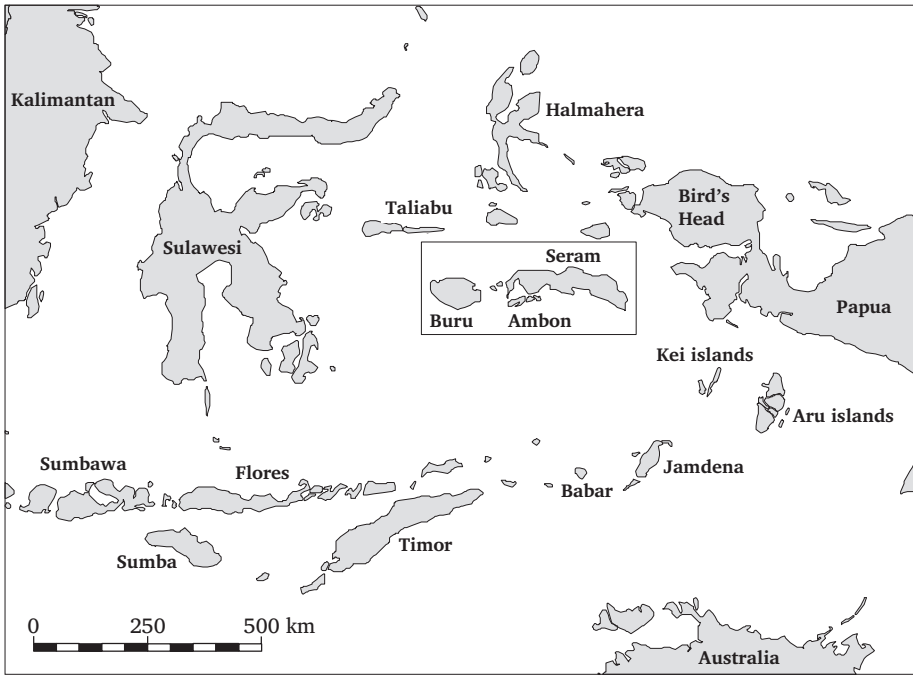
Given the above, one may ask why and how conceptual categories of one language can affect the use of another language. According to Backus et al., (2011, p. 740), in situations of intensive bilingualism, speakers tend to replicate elements that they have heard or used before, so that “everything that recurs in someone’s language experience is hypothesized to be entrenched in that speaker’s idiolect. If speakers speak two languages, patterns belonging to both languages will be entrenched in their minds”. Like words and structures, conceptual categories can also be expected to be replicated from one language to another. It is plausible to assume that the dominant language influences the way heritage speaker’s select and organize information when speaking the heritage language.

In the following chapters, I have adopted the Conceptual Transfer Hypothesis in Chapter 5 and in Chapter 6. The findings of these chapters show that heritage speakers have adopted a Dutch way of organizing information and verbalizing events.

1.5 History of Ambon Malay: the homeland community and the heritage community

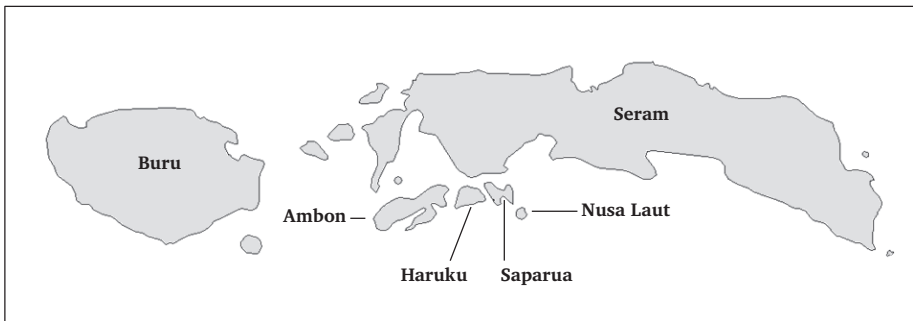
This section briefly describes the history of Ambon Malay and the most salient historical events that have influenced the development of this language in the Central Moluccas (Indonesia) and in the Netherlands. The information concerning the history and the development of Ambon Malay until the present days are mainly based on the studies of Collins (1980a), Grimes (1991), van Minde (1997; 2002), Sneddon (2003b), and Paauw (2008). The main sources for the history of the heritage Ambon Malay community in the Netherlands are Bartels (1986), Tahitu (1989), Veenman (1994) Smeets and Veenman (2000), Vermeulen and Penninx (2000), and van Wagtendonk (2008). The reader can consult these references for further information.

Ambon Malay, known as *Malayu Ambong* by its speakers, is spoken in the Central Moluccas (eastern Indonesia) by about 200.000 native speakers and by about 1.400.000 L2 speakers (Lewis, Simons, & Fennig, 2015); see Map 1.1 on the next page (the Central Moluccas are enclosed in a box).



Map 1.1: Eastern Indonesia and the Central Moluccas.

The Ambon Malay speaking area includes the islands of Ambon, Saparua, Haruku, Nusa Laut, and the coast of Seram and Buru, as shown on Map 1.2.



Map 1.2: Major Ambon-Malay speaking islands.

Since the 16th century, the island of Ambon has played a central role in the region. It was chosen by the Portuguese and subsequently by the Dutch as their economic and political center, and nowadays it hosts the capital of the province, *kota Ambon* ‘Ambon city’. Due to the pivotal role of Ambon, the Malay variety spoken there and

in the surrounding islands came to be known as ‘Ambon Malay’ or ‘Ambonese Malay’. This variety is sometime referred to as ‘Moluccan Malay’, but the term ‘Ambon Malay’ is preferred in order to avoid confusion with ‘North Moluccan Malay’, a variety spoken in the province of North Moluccas (Paauw, 2008).

Although Ambon Malay originated and developed in the Central Moluccas, this language was not an indigenous language of the Moluccas. Ambon Malay is derived from Vehicular Malay, the Malay trade language which was brought to the eastern part of Indonesia in the era preceding European contact (Paauw, 2008). The evidence shows that Vehicular Malay began to be used as a lingua franca in the Moluccas from (at least) the 16th century onwards. It was brought along to the eastern island from the west part of the archipelago by merchants who used it as a trade language (Collins, 1980a; Grimes, 1991; Paauw, 2008). First, it became the language for inter-ethnic communication among the Moluccans and the Malay-speaking traders. Later it also became a lingua franca among the Moluccans and the Europeans. When the Portuguese arrived in the Moluccas, in 1512, Vehicular Malay was already so widespread across the islands that the Portuguese missionary Antonio Galvao (1566) compared the use of Malay in the Moluccas to the use of Latin in Europe (quoted in Collins, 1980a, p. 7). During the Dutch colonial period (1606-1948), a standardized form of Malay, known as Literary Malay or High Malay, came to be used in schools and in the church. Standard Malay became established as the language for political and cultural communication, and Standard Malay, as well as Ambon Malay, was used by Islamic and Christian missionaries to spread their religions.

Despite its widespread use, Ambon Malay remained a second language used as a lingua franca up to the beginning of the 19th century. We do not know with certainty when Ambon Malay began to change from a second language to a native language used by a specific community. There is evidence, however, that by the 19th century some Christian villages on Ambon Island had shifted from indigenous languages to Ambon Malay (Grimes, 1991). While almost all the Christian villages shifted to Ambon Malay as their first language by the middle of the 19th century, some Muslim villages have preserved regional languages as a sign of religious identification. Due to this historical division, regional languages in Christian villages are being lost at a more rapid rate than languages spoken in Muslim villages (Florey & van Engelenhoven, 2001).

Over the centuries, Ambon Malay has borrowed many words from Portuguese and Dutch. Throughout the Dutch colonial period (1606-1948) Ambon Malay had

been exposed to Dutch influence more than any of the other contact Malay varieties. According to van Minde (2002, p. 213), 68.9% of loanwords in Ambon Malay are from Dutch, while 24.3% are from Portuguese. The majority of loanwords are nouns referring to concrete objects, such as *dos* ‘carton, box’ (<Du *doos*) or *martelu* ‘hammer’ (<Po *martelo*), although a number of verbs and function words have been borrowed as well (see Section 1.6.2.4 and Section 1.6.2.8). Beside Portuguese and Dutch, Ambon Malay has also been influenced by the indigenous languages of the region² (Collins, 1981, 1983; Grimes, 1991; van Engelenhoven, 2008). One example is the order of quantifier and noun. In five indigenous languages spoken on Ambon Island and on the neighboring islands, quantifiers follow the noun, as in Buru (*humar telo* ‘house tree’). This order is found also in Ambon Malay, which differs in this respect from western inherited Malay varieties, where quantifiers usually precede the noun (Grimes, 1991, p. 109).

The complex socio-historical scenario that gave rise to Ambon Malay has led some scholars to classify this language as a creole. For instance, Grimes (1991, p. 118) concludes that Ambon Malay is a creole that developed from a long standing pidgin because the multi-ethnic and multilingual setting where Ambon Malay originated is compatible with pidgin formation. The conclusion of Grimes (1991) is shared by Adelaar and Prentice (1996), who list Ambon Malay among the ‘Pidgin-Malay derived’ varieties, and by McWhorter (2001b), who also considered all vernacular Malay varieties as creoles (but cf. Gil, 2001). In contrast, Collins (1980a) holds a different view. By comparing Ambon Malay (a contact variety) to Trengganu Malay (a non-contact dialect of the Malay peninsula), Collins (1980a) shows that many linguistic features used as evidence for creolization, such as lack of inflection, few productive affixes, absence of copula, are also present in a non-contact Malay variety, and thus there is no direct link between these typological features and their ontogeny. According to Collins (1980a, p. 56), we need to be careful in using “the term “creole”” to refer to languages which developed in a creole-like cultural situation but perhaps the linguistic development occurred after a long period of contact”, especially if the linguistic criteria do not support creolization.

What is important for us now is that Ambon Malay was spoken in Ambon and in the surrounding islands since about 500 years, at a point when a political event

² Indigenous languages of the Central Moluccas belong to the Austronesian language family. Most of them are now nearly extinct or threatened by Ambon Malay and Standard Indonesian (Florey & van Engelenhoven, 2001). The most important (in terms of number of speakers) is Alune (spoken on Seram).

changed the history of this speech community. The Republic of Indonesia was proclaimed in 1945, but several years passed before the Dutch parliament recognized its independence and the Dutch began the decolonization process. After the transfer of sovereignty in 1949, the Dutch government was forced to dismantle its *Koninklijk Nederlandsch-Indisch Leger* (KNIL) (the Royal Dutch-Indies Army), which was formed mainly by Moluccan soldiers. The Moluccans, in general, and the Ambonese in particular, had always rendered their services to the Dutch; they were treated by the Dutch “as the ‘favored sons’ and were often given preferred positions in the military and colonial government” (Grimes, 1991, p. 101). These soldiers often lived or spent long periods of times in the barracks of the KNIL army or in army camps in Java where a divergent form of Ambon Malay was spoken. This language variety, known as Tangsi Malay or Barracks Malay was heavily influenced by Ambon Malay, Javanese and Dutch.

When the KNIL army had to be dismantled, a political problem arose. In fact, on April 24th, 1950, an independent movement proclaimed the independent Republic of South Moluccas, *Republik Maluku Selatan*, (RMS) on Ambon Island. At that point, the Indonesian government disallowed any demobilization of the KNIL soldiers on the Moluccas, fearing that the soldiers would fight for the RMS rebels.³ Demobilizing the soldiers in other Indonesian territories was also dangerous, as the Moluccans were afraid of reprisals from the nationalists who saw them as Dutch allies.⁴ The Dutch government, then, saw no other solution than to ‘temporarily’ bring the KNIL soldiers and their family (about 12.500 people) to the Netherlands. A conspicuous number of Moluccans (ex KNIL soldiers) who arrived in the Netherlands spoke Tangsi Malay, and this language is considered one of the sources, if not the major source, from which heritage Ambon Malay developed (Tahitu, 1989, Adelaar & Prentice, 1996; see also Section 1.5.2.2).

³ During the period from April to November 1950, the ‘new-born’ Indonesian army, the TNI (Tentara Nasional Indonesia), attacked and defeated the ‘rebels’ of the RMS on Ambon. A guerilla movement, however, continued on Seram Island for about a decade, until the leader of the movement, Dr. Christiaan Soumokil was arrested 1963, and executed in 1966 by order of Suharto. After Soumokil’s death, J. Manusama became president of the RMS exile government in the Netherlands.

⁴ Some Moluccan ex-KNIL soldiers decided to be demobilized on New Guinea, a region that was still under Dutch control at that time. When, in 1963, the Dutch transferred the sovereignty in New Guinea to the Indonesian government, several thousand of Moluccans choose to move to the Netherlands.

The arrival of ca. 12.500 Moluccans in the Netherlands had important consequences not only for the lives of these individuals, but also for their language. In fact, since 1951, two Moluccan communities are recognized: (i) the community in the Central Moluccas speaking the homeland language (see Section 1.5.1) and (ii) the community in the Netherlands speaking the heritage language (see Section 1.5.2) The homeland language spoken nowadays in Ambon is the direct descendent of the Malay variety that was spoken on Ambon Island since the 16th century. The heritage language, by contrast, is the descendent of Tangsi Malay, and to a lesser extent of Ambon Malay. In the past 50-60 years, the homeland language and the heritage language have begun to diverge, the former under the influence of (Standard) Indonesian, the latter under the influence of Dutch and to a lesser extent of (Standard) Indonesian. Figure 1.2 schematically presents this development.

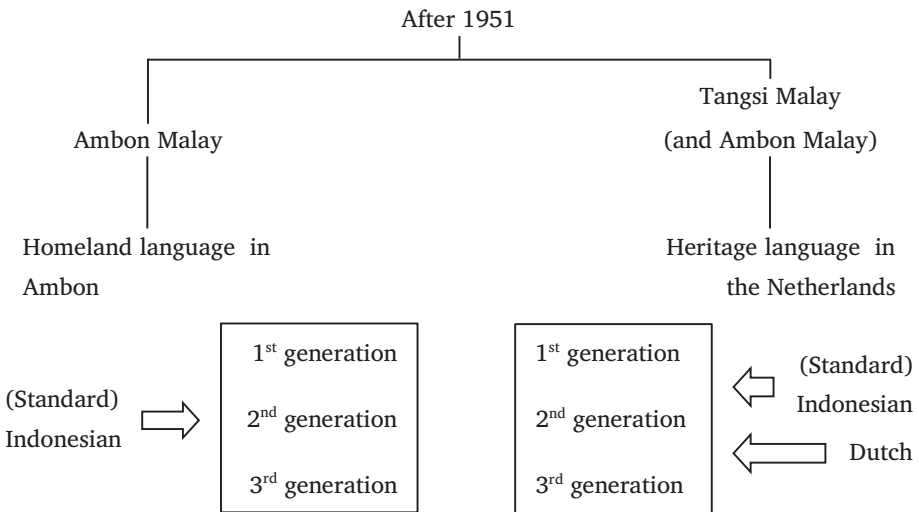


Figure 1.2: The parallel development of Ambon Malay in Ambon, Indonesia, and in the Netherlands

The reader needs to bear in mind that it is not entirely correct to say that the removal of the present-day Moluccan community to the Netherlands in the early 1950s constitutes the start of Dutch influence on the language that would become heritage Ambon Malay. First, there had already been considerable influence from Dutch on Ambon Malay throughout the Dutch colonial period (1606-1948). Second, extensive exposure to Dutch by the Moluccan immigrants arguably did not start in

1951 but was delayed some twenty years because of the relative isolation of Moluccas from Dutch mainstream society in the 1950s and 1960s.

1.5.1 The homeland community in the Central Moluccas

After independence, the Indonesian government established 26 provinces. All the Moluccan islands formed a single province until 1999, when a decentralization reform split off the northern islands into a separate province (*Provinsi Maluku Utara*, Province of North Moluccas). Nowadays, the Province of Moluccas (*Provinsi Maluku*) has a population of 1.664.631 (BPS⁵ 2013), and the capital of the province, Ambon city, counts 223.173 inhabitants (in 2002; see Leirissa, Pattykaihatu, Luhukay, Talib, & Maelissa, 2004, p. 183).

In 1999, the Central Moluccas experienced the explosion of extreme violence due to an inter-ethnic and inter-religious conflict between Christian Ambonese and Muslim migrants that was exacerbated by the Indonesian transmigration program and by the economic crisis of 1997. Amirrachman (2012, p. 48) reports that in the Moluccas the non-native Muslim population increased from 5% of the entire population in 1971 to more than 14% in 1995. By 1999, almost 100.000 people have transmigrated to the islands of Ambon, Seram and Buru. The violent conflict, which caused about 2.000 victims, was resolved in 2003, but it left a profound wound in the Moluccan community. "Before the conflict, we could still find some Muslim families living in a Christian area and also the other way around, but after the conflict this was almost no longer the case" (Amirrachman, 2012, p. 56); now villages are either entirely Christian or Muslim.

The economy of the Central Moluccas is mainly based on trade, commerce and agriculture (BPS). In 2002, the economic profile of Ambon city, for instance, was subdivided as follows: 28.36% trade, 23.47% commerce, hotels, and restaurants; 21.68% agriculture (Leirissa et al., 2004, p. 185). For the trade and commerce sector, BPS (2012) reports that there are 9.306 companies in the province, which employ 27.280 workers. Out of these, 4.006 are trading companies. Ambon city alone hosts 21% of all companies in general and 37.9% of all trading companies. In the agricultural sector, cassava and unmilled rice are the main crops planted in the region, with an annual production of 97.813 ton and 96.807 ton, respectively (BPS, 2013). The main horticultural products are bananas (7.742 ton), mangos (1.543 ton)

⁵ BPS stands for *Badan Pusat Statistik Provinsi Maluku* (Center for statistics of the Moluccas Province), online URL <http://maluku.bps.go.id/> [Last accessed 30 June 2015].

and oranges (1.215 ton) (BPS, 2013). Finally, many families living in the small coastal communities maintain themselves with the exploitation of marine resources (Harkes & Novaczek, 2002).

1.5.1.1 The homeland language

The homeland language spoken nowadays in Ambon is the direct descendent of Ambon Malay, the contact Malay variety spoken on Ambon Island since the 16th century. Nowadays, Ambon Malay stands in a diglossic situation with regional languages on the one hand, and with Standard Indonesian, the national language, on the other (Grimes, 1991; Paauw, 2008). It is considered a “High” variety when compared to regional languages, but a “Low” variety when compared to Indonesian (for some speakers Ambon Malay is an inferior variety of Indonesian).

The diglossic situation with Standard Indonesian is better described as a continuum because no clear division exists between the two languages. Depending on the situation and on their knowledge of the standard language, speakers may use more Ambon Malay-like features or more Indonesian-like features (Sneddon, 2003b; Paauw, 2008). For instance, speakers may make their speech more formal by replacing the Ambon Malay pronouns, with Indonesian pronouns, or they may use the Indonesian negator *tidak* ‘NEG’, instead of Ambon Malay *seng*. Standard Indonesian is the language of government, administration, law, formal education, and mass media (Sneddon, 2003b). In contrast, Ambon Malay is used when talking with family members and friends, and also with outsiders. One homeland informant (female, 19 years old) on Ambon Island reports that:

Homeland speaker:

“If I speak with friends, I use Ambon Malay, but in the school environment [when we follow classes], Indonesian is used [...], if I speak with my mother, I use Ambon Malay, the everyday language [...], with my relatives I also use Ambon Malay but not the rough version [...], with friends also I used Ambon Malay, yes my whole family speaks Ambon Malay [...], yes Ambon Malay is important, but Indonesian is also important, it depends on the situation in which we are.”⁶

⁶ In the original: “*kalo deng tamang-tamang, bahasa Ambong tapi kalo dalam lingkungan kulia di tanpa studi, musti harus bahasa Indonesia [...], kalo bicara deng mama, bahasa Ambong, bahasa sehari-hari [...], kalo deng sodara lai bahasa Ambong tapi seng kasar [...], kalo deng tamang lai bahasa Ambong, ya secara keseluruhan keluarga samua katong bahasa Ambong [...], iya bahasa, bahasa Ambong penting, bahasa Indonesia jua penting, tergantung situasi dimana katong berada.*”

Ambon Malay speakers seem to be aware of the fact that Standard Indonesian and Jakarta Indonesian (the variety of Indonesian spoken in the capital)⁷ have a strong impact on their language, and that this is due to education and to the exposure of young generations to the mass media, songs, soap operas, and TV shows. This is expressed very clearly in the words of one homeland informant (male, 32 years old):

Homeland speaker:

“Ambon Malay is certainly already influenced...it is mixed with Indonesian, because now children and young people...many of them already started speaking a mix [of Ambon Malay] with Indonesian [...], yes as the electronic media and the books also, everything is in Indonesian, rarely something is in Ambon Malay. So if every day they use Indonesian, Ambon Malay begins to diminish [...] There is influence from songs and music [...] because if...even more young children now, they love songs from abroad [...], many of them don't like to listen to Ambon songs; but take the Indonesian bands and also the international bands, oh they love them.”⁸

Although the language of mass media is mainly (Standard or Jakarta) Indonesian, it is important to point out that in recent years a number of publications have appeared which use Ambon Malay. One of them is the monthly magazine *Kacupeng*, issued since 2007, which provides a platform for exchanging ideas and discussing local news. Although most of the articles are actually written in Standard Indonesian, the magazine contains also some texts and comic strips written in Ambon Malay. Other publications in Ambon Malay are the translations of Bible portions made available by SIL International.

Despite the constant influence of Indonesian, Ambon Malay is not a threatened language at the moment. On the contrary, it continues to gain speakers, to the

⁷ Colloquial Jakartan Indonesian is gaining increasing prestige throughout the archipelago due to its use in films and TV series. This variety is becoming extremely popular also because it is associated to the ‘wealthy and attractive’ people living in the capital (Sneddon, 2003b).

⁸ In the original: “*Memang bahasa Ambong su mulai dipengaruhi deng e...tacampor deng bahasa Indonesia, karna kebanyakan skarang dari ana-ana muda sampe ana-ana kacil jua banya su mulai bicara tu tacampor deng bahasa Indonesia [...], iya untuk media-media elektronik deng kaya buku-buku samua, itu kan su deng bahasa Indonesia, untuk deng bahasa khususnya deng bahasa Ambong tu jarang. Jadi kalo setiap hari orang tinggal kong deng bahasa Indonesia tetap, oh iya bahasa Ambong mulai berkurang [...]. Tetap ada pengaruh sa dari lagu-lagu deng musik itu [...] karna kalo...apalagi kalo ana-ana muda sekarang tu kan, dengar paleng suka par dengar lagu-lagu dari luar [...], dengar lagu daera Ambong ni jua, banya yang seng suka, tapi kalo bilang kata bend-bend dari maksudnya Indonesia punya trus yang dari luar negri punya, ah itu dong lebe suka.*”

detriment of indigenous languages. For instance, out of 27 informants recorded in Ambon, only two reported speaking indigenous languages. One informant (female, 58 years old) reported speaking Alune with her husband and sometimes with her children. The other informant (female, 73 years old.) reported knowing several indigenous languages, including Hulaliu,⁹ which was the language spoken by her parents, and Galela and Tobelo,¹⁰ which she learned later on in life.

One possible reason for the limited presence of indigenous language speakers in the dataset is that the interviewed informants were all from Christian villages. As described in Section 1.5, indigenous languages are better preserved in Muslim communities than in Christian villages. This situation is also acknowledged by Ambon Malay speakers themselves. The following is a fragment from an interview conducted with a homeland speaker living in Amahusu, on Ambon Island (male, 57 years old):

Homeland speaker

“Specifically about those languages [the regional languages], in Muslim areas, Muslims still use them because they still maintain their language, those languages, they have never lost them, but [in] the majority of the Christian villages, it is over.”¹¹

1.5.1.2 Previous studies on homeland Ambon Malay

Van Minde’s (1997) dissertation, which includes a grammatical description of the language as well as texts, represents the most comprehensive source on Ambon Malay to date. Besides, scholars, such as Collins (1983), Tjia (1992, 1997, 2004), and Litamahuputty (1994) have investigated a variety of different features of Ambon Malay grammar in more detail. In his (1992) article, Tjia describes a number of particles (aspectual, modal, sentence-final, etc.) used in contemporary Ambon Malay, while in his MA thesis (1997), he illustrates the structural criteria and the functions of Ambon Malay serial verb constructions. Tjia (2004) gives a synchronic and a diachronic explanation of two typical Ambon Malay constructions, the possessive and the passive. The meaning and function of the Ambon Malay aspect

⁹ Hulaliu is an Austronesian language spoken on Haruku.

¹⁰ Galela and Tobelo are Papuan languages spoken in the Province of North Moluccas.

¹¹ In the original: “*Khusus untuk bahasa-bahasa bagitu, kalo daera-daera muslim, muslim dong kan mase pake karna dong tetap pake dong Bahasa, bahasa itu, dong seng perna ilang, tapi sebagian basar daera-daera kresten, abis.*”

marker *suda* 'PRF', with its short form *su*, are discussed in van Minde (1997) and Tjia (2002). Collins (1983) explores the syntactic changes that have occurred in the Ambon Malay possessive construction. Finally, Litamahuputty (1994) examines causative constructions with *biking* 'make' and *kasi* 'give' in Ambon Malay, Manado Malay and Ternate Malay.

Other publications on Ambon Malay include a study on Portuguese loan words by Abdurachman (1972), and a study on loan words from European languages by van Minde (2002). A comparison of Ambon Malay to Indonesian is presented in Collins (1974). Collins (1980a) is dedicated to the question whether Ambon Malay can be correctly classified as a creole using typological and the comparative approaches. Collins (1980b) describes the effects of Ambon Malay on the indigenous language Laha, spoken on Ambon Island, while Collins (1981) describes the effect of indigenous languages on Ambon Malay. Code switching between standard Malay and non-standard varieties, including Ambon Malay, is examined in Nivens (1994).

A socio-historical view of the development of Ambon Malay is given in Grimes (1991). After describing some features of Ambon Malay, with a focus on the differences with Standard Indonesian and the similarities with the indigenous languages, the author examines linguistic data in light of the in socio-historical and socio-linguistic facts.

Besides, three dictionaries are available for Ambon Malay: the *Kamus bahasa Melayu Ambon-Indonesia* (Ambon Malay-Indonesian dictionary) (Takaria, Pieter & Muis 1998), the *Daftar kata-kata Moluks Maleis-Nederlands* (Words list Moluccan Malay-Dutch) (Tahitu 2000), and the *Kamus bahasa-harian: dialek orang Ambon* (Dictionary of the everyday language: the 'dialect' of the Ambonese) (Mailoa 2006).

1.5.2 The heritage community in the Netherlands

Between March and June 1951, 12,578 Moluccans arrived in the Netherlands on board of eleven ships. They were subdivided as follows: 821 corporals, 372 sergeants, 35 sergeant-majors, 6 clergymen, 3 adjutants, 2341 privates, and 9000 family members (van Wagtendonk, 2008, p. 45). No ethnic breakdown for the total of 12,578 people is available, but Bartels (1986, p. 27) reports that 76.1% of the troops came from the Central Moluccas, while the remaining 23.9 % came from the southeastern Moluccan islands. Besides, a number of soldiers had wives of non-Moluccan origin (mostly Javanese). This picture probably also reflects the ethnic composition of the immigrant population. Unfortunately, the percentages reported in

Bartels (1986) only tell us about the origins of the soldiers. They say little as to where the soldiers had lived prior to their arrival in the Netherlands and what language(s) they spoke.

During the first period of residence, little effort was made by the Dutch and by the Moluccans to facilitate the integration process because both parties believed that the stay of the Moluccans in the Netherlands would be temporary. The idea of a short-term stay led the Dutch government to exclude the Moluccans from the housing and the labor markets. After their arrival, the Moluccans were housed in 34 camps (*woonoorden*)¹² spread all over the Netherlands (see Map 1.3 on the next page), such as in the ex-WWII concentration camp of Vught. The camps were situated in rural and isolated areas of the country, where the Moluccans lived segregated until the 1960s. In these camps, the Moluccans lived without work and without income, as gas, electricity and water were free, while food and clothes were subsidized. This may seem positive. However, as van Wagendonk (2008, p. 13) points out, these subsidies fostered a sense of ‘economic impotence’ amongst the people living in the camps.

In the beginning of the 1960s, it became clear that the stay of the Moluccans would become permanent.¹³ Therefore, the Dutch government decided to close the camps and to move the Moluccans to newly built wards (*woonwijken*) on the outskirts of small towns. This process, however, took a long time, and the last camp, Lunetten, was closed in 1989. Vermeulen and Penninx (2000, p. 9) report that “though they [the Moluccans] increasingly disperse from these areas, they are still the least likely of all six groups [Turks, Moroccans, Southern Europeans, Surinamese, Antilleans, and Moluccans] to be found in big cities”.

The 1960s and 1970s saw little economic progress for the Moluccans. They were mostly employed as unskilled or semiskilled laborers in factories due to their limited levels of education and inadequate mastery of Dutch. What dominated collective thought in those years was the idea to return to an independent state in the Moluccas (Bartels 1986). In the 1970s, the sense of anger and frustration that grew within the community escalated in four train hijackings and hostage taking actions. The aim of these attacks was to force the Dutch and the Indonesian governments to establish a free Republic of South Moluccas, independent of Indonesia, to where they and their parents could finally return. Although these attacks did not achieve their

¹² The number of camps increased to 71 in 1958 (Veenman, 1994, p. 7).

¹³ This conclusion, however, was made public only in 1978 (Smeets & Veenman, 2000).

intended goals, they attracted the attention of the Dutch government, which began to spend money on projects in favor of the Moluccans.



Bron: LSEM/MHM 2003

Map 1.3: Moluccan camps in the Netherlands (1951-1989).

Groningen

1. Carel Coenraad Polder (Finsterwolde)
2. Marum (Nuis)

Friesland

3. Oranje (Fochteloo)
4. Ybenheer (Fochteloo)
5. Wite Pael (Sint Johannesga)
6. Wyldemerck (Balk)

Drenthe

7. De Fledders (Norg)
8. Schattenberg (Westerbork)
9. Pieterberg (Westerbork)
10. Mantinge (Westerbork)
11. Ruinen (Ruinen)
12. Stuijfzand (Ruinen)
13. Geesbrug (Oosterhesselen)

Overijssel

14. Beenderribben (Steenwijkerwold)
15. Pikbroek (Steenwijk)
16. Eind van 't Diep (Steenwijk)
17. Beugelen (Staphorst)
18. Conrad (Rouveen)
19. Laarbrug (Ommen)
20. Eerde (Ommen)
21. Vossenbosch (Wierden)
22. St. Joseph (Glanerbrug)

Gelderland

23. De Zwaluwenburg (Elburg)
24. Vaassen
25. Teuge (Voorst)
26. De Biezen (Barneveld)
27. De Schaffelaar (Barneveld)
28. Elzenpasch (Tiel)
29. Schutsluizen (Tiel)
30. Overbroek (Echteld)
31. Lingebrug (Opheusden)
32. De Haar (Randwijk)
33. De Bruynhorst (Edeveen)
34. Golfliks (Arnhem)
35. Onderlangs (Arnhem)
36. Snodenhoek (Elst)
37. Klein Baal (Haalderen)
38. De Hogehorst (Groesbeek)
39. De Kemp (Wehl)
40. Vosseveld (Winterswijk)

Noord-Holland

41. Oude Zeug (Wieringermeer)
42. Medemblik
43. Coehoorn (Muidenberg)
44. Almere (Huizen)

Zuid-Holland

45. Kazerne (Woerden)
46. Utrechtse Straatweg (Woerden)
47. Singel (Woerden)
48. IJsseloord (Capelle a/d IJssel)
49. Kamp Q (Slikkerveer)

Zeeland

50. Noordwelle (Westerschouwen)
51. Burghsluis (Westerschouwen)
52. Brydorpe (Middenschouwen)
53. Kerkwerve (Middenschouwen)
54. Grijskerke
55. Westkapelle
56. Serooskerke
57. Koudekerke
58. Vlissingen I
59. Vlissingen II
60. Vliegvelde Souburg
61. Middelburg I
62. Middelburg II
63. Kruiningen I
64. Kruiningen II
65. De Haven (Breskens)
66. Duinoord(Groede)
67. Wilgenhof (Oostburg)
68. Rodanborgh (Aardenburg)

Noord-Brabant

69. Wouw
70. Lunetten (Vught)
71. Donzel (Nistelrode)
72. Villheide (Mill)
73. Vierlingsbeek
74. Lage Mierde
75. Baarschot (Middelbeers)

Limburg

76. Plasmolen (Mook)
77. Genapium (Gennepe)
78. Oude Molen (Welt)
79. Vlakwater (Venray)
80. Tienray (Meerlo/Wanssum)
81. Blerick (Venlo)
82. Tungelroy (Weert)
83. Heythuizen
84. Maashaven (Roermond)
85. Montfort
86. Op de Loop (Echt)
87. Lillbosch (Echt)
88. Graetheide (Geleen)
89. Rijckholt (Eijsden)
90. Capucijnse Klooster (Eijsden)

At the beginning of the 1980s, the Moluccan population had increased to 35.000 family units. Despite the efforts for the social integration, there was still a high level of unemployment and 80% of the population was still living in camps or wards. Things began to change in 1990s, though, when the unemployment rate fell from almost 40% (in 1983) to 15% (in 1990) (Veenman, 1994, p. 31), and Moluccans began to move out of camps and wards. In 1994, 60% of the Moluccans lived outside a camp or ward (in 1978 the figure was only 20%; Veenman, 1994, p. 15). Veenman (1994) reports that Moluccans living in a municipality with a Moluccan ward were less likely to have contact with the Dutch than those living in a municipality without a Moluccan ward. Furthermore, generally speaking, young women enjoyed a higher level of informal participation in Dutch society than young men, who tended to have contact predominantly with their own group. The increase in the frequency and intensity of contact with the Dutch has also led to an increase in the numbers of mixed partnerships (Dutch and Moluccan), so that the proportion of mixed partnerships for males rose from 45% (in 1983) to 51% (in 1990), and for females from 25% (in 1983) to 40% (in 1990) (Veenman, 1994, p. 24).

Nowadays, the Moluccan population is estimated at 40.000. The rate of unemployment is the same as for the indigenous Dutch population, and “young Moluccans began to find places in the professional and academic realms, and established themselves in new communities throughout the country (van Wagtenonk, 2008, p. 39).

1.5.2.1 The heritage language

The variety of Ambon Malay spoken in the Netherlands by the heritage community is known as *Melayu Sini* ‘Malay from here’ (Tahitu, 1989). In the present work, however, this variety is referred to as ‘the heritage variety’ or ‘heritage Ambon Malay’ in order to highlight the fact that this language (i) is spoken by bilinguals with another dominant language (Dutch), (ii) has a cultural value for the speakers, and (ii) is a language with a rather long history in the Netherlands (see Section 1.2).

Heritage Ambon Malay began to develop as an independent language variety after 1951, when about 12.5000 Moluccans arrived in the Netherlands. As we have seen above, Tangsi Malay, a divergent form of Ambon Malay spoken by ex-KNIL soldiers, is probably the progenitor of heritage Ambon Malay (Tahitu, 1989; van Engelenhoven, 2008, p. 112). Tangsi Malay was a vernacular contact language spoken in the military barracks of the KNIL army on Java, Sumatra, and Sulawesi. Structurally, Tangsi Malay was similar to other Malay varieties, but its lexicon was

heavy influenced by Ambon Malay (the language of most of the soldiers), Javanese and Dutch (Adelaar & Prentice, 1996). Many of the ex-KNIL soldiers that were brought to the Netherlands spoke Tangsi Malay, or a variety of Ambon Malay with a strong Tangsi flavor, and for this reason we find typical Javanese elements, such as *kuping* ‘ears’ (instead of *talingang* ‘ears’), or the pronouns *aku* ‘1SG’ and *kue* ‘2SG’ in the lexicon of heritage Ambon Malay (see Section 1.4.2.4). Chapter 3 and Chapter 5 discuss two other Tangsi Malay elements inherited by heritage Ambon Malay, namely the definite marker =*nya* ‘DEF’ and the use of the preposition *buat* ‘to, for’ give-constructions.

After their arrival in 1951, the Moluccans were housed in camps and wards in rather remote and isolated areas of the Netherlands. Although the isolated housing situation caused a delay in the integration process, it strengthened mutual links and it facilitated language maintenance (Veenman, 1994). Tangsi Malay was the language used in the barracks by first generation immigrants among themselves and with their children (see Section 1.2). Despite the fact that some Moluccans had some knowledge of Dutch before their arrival and some married a Dutch partner (see Section 2.1.2.3), Tangsi Malay (and Ambon Malay) remained the main means of interaction among the Moluccans, especially in the first two decades of their stay.

Nowadays, after about 60 years, heritage Ambon Malay is still spoken in the community, although its domains of use are becoming limited. Most of the participants in the study report speaking heritage Ambon Malay when interacting with elder people, when communicating with family and friends in the homeland, at weddings, funerals, parties or other Moluccan events and in church. Some still speak it with their parents, friends and neighbors; this behavior, however, is more common among Moluccans who still live in Moluccan municipalities (see Chapter 7). Switching between Dutch and heritage Ambon Malay is a common practice for the Moluccans, who report speaking a mix (*campor*) of the two languages (see Section 1.5.2.2 for an overview of studies on code-switching).

Through code-switching, in particular, and through bilingualism, in general, many Dutch features (words and structures) have entered the language. A number of these innovative features are discussed throughout the following chapters. Quite predictably, heritage speakers are more aware of the influence of Dutch on their lexicon than on their grammar. When asked about their opinions on the Ambon Malay variety spoken in the Netherlands, most of the participants pinpoint the use of Dutch words, as illustrated in the following fragment provided by a heritage speaker (male, 34 years old):

Heritage speaker

“Actually they speak Ambon Malay, or actually not really Ambon Malay, but the Malay from here, mixed with Dutch words.”¹⁴

In addition to Dutch words, the heritage variety diverges from the homeland variety with respect to the use of archaisms and the lack of newly coined words. This discrepancy is recognized by heritage speakers, but also by homeland speakers. The following two fragments - the first by a heritage speaker (male, 45 years old) and the second by a homeland speaker (male, 17 years old) identify the absence of new terms and the use of phonologically long forms that are perceived as archaic as two possible sources for divergence:

Heritage speaker

“Yes there they use new... new words, we do not have them here.”¹⁵

Homeland speaker

“For me, their language is different from ours, what I mean, actually it is almost the same, but sometimes, like we say *katong* [‘1PL’ shortened form], they say *kitorang* [‘1PL’ long form] or...like that, there is a small difference.”¹⁶

Another interesting point, made by a homeland speaker (male, 46 years old) whom I interviewed on Ambon Island, is that male heritage speakers seem to preserve the language more firmly than women. This may relate to the fact that (young) Moluccan women are more prone to have contact with the Dutch than Moluccan men (see Section 1.5.2).

Homeland speaker

“Men, their Ambon Malay is good, if we speak about the men, but women, overall (their Ambon Malay) is not so good, [...] they have already forgotten a lot.”¹⁷

¹⁴ In the original: “*Dong sebenarnya omong bahasa Melayu Ambong, atau sebenarnya bukannya Melayu Ambong, tapi Melayu dari sini sebenarnya, campuran deng kata-kata Belanda.*”

¹⁵ In the original: “*Iya di sana to, dong pake baru, itu kata-kata baru di belong ada.*”

¹⁶ In the original: “*Kalo par beta sandiri, dong pung bahasa Melayu tu beda deng katong, maksudnya memang amper sama, cuma kadang-kadang, yang kata katong bilang “katong”, dong bilang “kitorang” ato kaya bagitu-bagitu, ada sadiki beda.*”

Finally, it is important to point out that Indonesian is also exerting influence on heritage Ambon Malay (although to a lesser extent than Dutch). Many heritage speakers, in fact, are exposed to (Standard) Indonesian in the Moluccan churches in the Netherlands, and also when interacting with people from other parts of the archipelago or when visiting Jakarta, Bali or the homeland (see Chapter 5).

1.5.2.2 Previous studies on heritage Ambon Malay

The most comprehensive publication on heritage Ambon Malay is Tahitu's (1989) dissertation, which includes a description of the phonology, morphology, and some basic syntactic features, as well as some texts. Tahitu (1988) presents a comparison of the sound system of the heritage variety to the sound systems of the homeland variety and of Standard Indonesian. Another study that looks at a specific aspect of the heritage grammar is van Engelenhoven (2008), which describes the use of deictic elements, such *ini* 'D.PROX' and *itu* 'D.DIST' in the Malay speech of Southwest Moluccan migrants in the Netherlands. Tahitu and Lasomer's (2001) article covers various topics related to heritage Ambon Malay. It presents a historical sketch and a grammatical sketch, and discusses language use and language attitudes between different generations and in different socio-cultural contexts.

Other publications on heritage Ambon Malay mainly focus on patterns of language use. Huwaë (1992), for instance, examines language change and code-switching patterns in the heritage Ambon Malay variety spoken by six Moluccan women (ages ranging from 24 to 37) in the community of Wierden. She investigates language change by looking at lexical and grammatical knowledge of Ambon Malay items (words and structures) elicited by means of translations tasks and interviews. Her findings show that older women score higher in Dutch-Ambon Malay translation tests than younger women. The results also show that words can be divided into five groups according to their translation likelihood. That means that words such as 'sleep' and 'chair' are likely to be known by every informant, while the equivalents for words such as 'bitter' and 'suddenly' are known only by the older ones. In order to tests grammatical knowledge, Huwaë (1992) also examined the use of a number of different structures, such as possessive marking, tense and aspect marking, causatives, and passives. Her findings reveal a shift in preference in the domain of possessive marking, such that speakers often use the *dari* 'from' possessive

¹⁷ In the original: "Ehm kalo yang laki-laki, dong ng bahasa Ambong bagus, Melayu Ambong bagus, kalo ana laki-laki, tapi kalo yang parampuang, rata-rata tidak terlalu bagus mereka, [...] dong su lupa banya paskali."

construction (Possessed + *dari* 'from' + Possessor) as an alternative to the more Ambonese *punya* 'have, possess' possessive construction (Possessor + *punya* 'poss' + Possessed). Her findings also show that the use tense-aspect marker is very limited, especially in younger speakers. Finally, Huwaë (1992) reports that the most common type of code switching is the intrasentential type, which occurs within sentence boundaries. This is followed by the extrasentential type, which occurs at the sentence margin (with the insertion of a tag), and by the intersentential type, which occurs between sentences.

Voigt's (1994) study similarly deals with code-switching and language shift in a group of four Moluccan speakers of the second generation living in Breda. The analysis of the data (obtained by means of personal interviews and by recording the rehearsals of a Moluccan rock band) reveals that switching from heritage Ambon Malay to Dutch is far more frequent (90.5%) than from Dutch to heritage Ambon Malay (9.5%). Another observation is that intrasentential code switching is the most common type of switching (84.5%), followed by intersentential code switching (9.3%) and extrasentential code switching (6.2%). These findings are partially in line with those reported by Huwaë (1992), whereby intrasentential code-switching is also the most common type of switching. The data further shows that code-switching occurs more frequently with nouns (31.6%) and with conjunctions (13.5%) than with other word classes. According to Voigt (1994), the high percentage of nouns is expected, as nouns are easily borrowed due to their referential function. The high percentage of conjunctions, by contrast, is somehow unexpected. Voigt explains that, in his corpus, the conjunctions *maar* 'but' and *dus* 'so, thus' occur always in Dutch and almost never in heritage Ambon Malay. Hence, he concludes that Ambon Malay shows a tendency to converge toward Dutch with respect to conjunctions. Voigt, however, does not seem to be aware that *mar* 'but' (<Du *maar*) 'but' and *des* 'so' (<Du *dus*) are also Ambon Malay words (see Section 1.6.2.8).

Vos (2008) investigates code-switching patterns of fifteen second and third generation Moluccan speakers in order to find out how they express their ethnic identity through language use. She collected data in a Moluccan church in Hoogeveen where she recorded parts of the Mass and the group discussion after the service. The results of her analysis show that there are five ways in which second generation Moluccans switch from Dutch to heritage Ambon Malay to build their identity. First, they make use of words or expressions which refer to items, concepts or values typical of the Moluccan culture. Second, they avoid translating words or expressions that have a specific connotation in heritage Ambon Malay, but lose this

connotation if rendered in Dutch. Third, they switch to heritage Ambon Malay when the subject of the discussion represents a cultural value for the Moluccan community. Fourth, communications for and concerning the Moluccans and the Moluccan community are usually uttered in heritage Ambon Malay and introduced by the first person plural pronoun *ketong* ‘1PL’ meaning ‘we, the Moluccans’. Fifth, they switch to heritage Ambon Malay when the person that is talked about is a Moluccan; in this way the language is used to identify the ethnicity of the subject.

Finally, Lekawael’s (2011) thesis is the first attempt to describe heritage Ambon Malay or *Melayu Sini* in the heritage language framework. By comparing the grammatical structures found in the speech of eight heritage speakers to those provided by four first generation speakers, the author is able to identify a number of features that differentiate Ambon Malay in the Netherlands from Ambon Malay in Ambon and from Standard Malay.

1.6 Sketch grammar of Ambon Malay

This section presents a sketch of Ambon Malay grammar, based on the data collected for the present study and on the work of Collins (1980a, 1983), Tahitu (1989), Tjia (1992, 1997, 2004), van Minde (1997), and Paauw (2008). It serves to enable the reader to appreciate the following chapters, which describe innovative features of the Ambon Malay variety spoken in the Netherlands (the heritage variety). Due to the limitation of space, the description here is kept brief and general. The following sections describe in short the phonology (Section 1.6.1), word classes (Section 1.6.2), word formation (Section 1.6.3), the noun phrase (Section 1.6.4), the verb phrase (Section 1.6.5), the clause (Section 1.6.6), verb serialization (Section 1.6.7), and negation (Section 1.6.8).

1.6.1 Phonology

Ambon Malay has five phonemic vowels¹⁸ and 19 phonemic consonants, as shown in Table 1.2 and Table 1.3 (on the next page), respectively. The orthographic representation is in parenthesis.

¹⁸ According to Tahitu (1988, 1989), the heritage variety has three additional vowel phonemes, namely /i, ə, o/. In the homeland variety, [ɪ] and [o] are the allophones of /i/ and /u/, respectively. Since, in my corpus, I did not find any of the (near) minimal pairs listed in Tahitu (1988), I considered them allophones also in the heritage variety. There is no schwa phoneme

Table 1.2: Ambon Malay vowel phonemes.

	FRONT	CENTRAL	BACK
High/Close	i		u
Mid	ɛ (e)		ɔ (o)
Low/Open		a	

Table 1.3: Ambon Malay consonant phonemes.

	LABIAL		ALVEOLAR		PALATAL		VELAR		GLOTTAL
Stops	p	b	t	d	c	ɟ (j)	k	g	
Nasals		m		n		ɲ (ny)		ŋ (ng)	
Fricatives	f		s						h
Lateral				l					
Trill				r					
Approximants	w					j (y)			

In word-final position, stops do not occur and nasals are velarized, except in loanwords. Note, however, that words that have a cognate in Standard Indonesian are sometimes pronounced following the Indonesian convention with the final stop. The labio-dental fricative /f/ and the glottal fricative /h/ occur in loanwords. The most common types of syllables are V, CV and CVC. Most words are disyllabic (70.5%), while the others are trisyllabic (18.0%), or monosyllabic (6.3%) (van Minde, 1997, p. 53). The stress falls on the penultimate syllable, unless that syllable

in the homeland variety, as Proto-Malayic *ə became either /a/ or /e/ (Adelaar, 2005, p. 206). It is important to note, however, that there is considerable inter- and intra-speaker variation, and that the increasing influence of Standard Indonesian may affect the pronunciation of some words, so that *sakarang/sekarang/skarang* ‘now’ are all to be found. The schwa in the heritage variety is represented orthographically as <e>, following the conventions for Standard Indonesian.

contains an original *ə, in which case the stress falls on the ultimate syllable, as in *ba'li* (Proto-Malayic **bəli*) ‘to buy’ (but see Maskikit-Essed & Gussenhoven, forthcoming). The orthography of Ambon Malay follows the conventions for Standard Indonesian and for the other Malay varieties, which have phonemic orthographies (unless specified otherwise).

1.6.2 Word classes

Ambon Malay has two open word classes and a number of closed classes. The open classes are nouns (1.6.2.1) and verbs (1.6.2.2). The major closed classes are adverbs (1.6.2.3), personal pronouns (1.6.2.4), demonstratives (1.6.2.5), quantifiers (1.6.2.6), prepositions (1.6.2.7) and conjunctions (1.6.2.8). A characteristic of all Malay varieties, including Ambon Malay, is that there is considerable overlap between categories, so that “word-forms which semantically appear to be verbs easily and without further morphological modification occur in nominal functions and vice versa” (Himmelmann, 2005, p. 127).

1.6.2.1 Nouns

The category of noun includes common nouns (e.g., *ruma* ‘house’), proper nouns (e.g., *Haruku* ‘Haruku Island’) and classifying nouns (e.g., *orang* ‘person’). Nouns can function as clausal subjects or objects. In addition, they can also function as clausal predicates (without any copula), and they can be modified by a number of attributes (see Section 1.6.4). Nouns are not marked morphologically for number, case, definiteness or any other grammatical category, so that “an unaffixed common noun such as *ana* could, depending on contextual and situational indicators, be interpreted as ‘child; a child; the child; children; the children’” (van Minde, 1997, p. 60). The correct interpretation is usually provided by the context.

1.6.2.2 Verbs

Verbs function as predicates in clauses. Besides, they can also function as modifiers in noun phrases (see below), and they can be modified by mood-aspect markers (see Section 1.6.5). Verbs can be monovalent (intransitive), bivalent (transitive), or trivalent. Monovalent verbs are either dynamic (e.g., *masu* ‘enter’) or stative (e.g., *basar* ‘be big’). Stative verbs express states or qualities, and they often occur as

modifiers in the noun phrase.¹⁹ Trivalent verbs are limited in number, and they rarely occur with three full NPs (see Chapter 5 for a more detailed discussion). Verbs are not marked morphologically for person, number, mood, voice, tense and aspect. Similar to nouns, the correct interpretation is inferred from the context.

1.6.2.3 Adverbs

Adverbs modify the predicate. Their main function is to indicate time, degree, aspect, modality and negation (for negation see Section 1.6.8). They usually precede the constituent they modify, but there are some exceptions. Time adverbs, such as *beso* ‘tomorrow’, *nanti* ‘in a while’, *tadi* ‘a while ago’ usually occur at the beginning or at the end of a clause. Degree adverbs, such as *lebe* ‘more’ and *paleng* ‘very’ precede the verb, while *batul* ‘really’ and *sakali* ‘very’ follow the verb. Aspect adverbs, such as *suda* ‘already’ and *masi* ‘still’ precede the verbs. Modality adverbs, such as *mangkali* ‘maybe’ and *pasti* ‘surely’, can either precede the verb or they can occur at the beginning of the clause.

1.6.2.4 Personal pronouns

Personal pronouns refer to animate or inanimate entities and fill the position of a noun phrase in a clause. Personal pronouns are differentiated for person, number and politeness. They have a full form and one or more abbreviated forms. The full forms can occur in the subject, direct object, and indirect object slots. The abbreviated forms can take the subject slot, but there are some restrictions for the object and the indirect object slots. The homeland variety and the heritage variety present some dissimilarity in the pronoun system, which reflect the different historical trajectories of their speakers (see Section 1.5). The personal pronoun systems of both varieties of are presented in Table 1.4 on the next page.

The homeland variety has *beta* for the first person singular and *ose* (<Po voce) or *ale* for the second person singular; *ose* and *ale* express familiarity and are usually used to address peers. In the heritage variety, *beta* (1sg) is used to indicate respect, while *aku* (and the abbreviated form *a*) has more colloquial functions; *kue* (2sg) is used by youngsters to address peers. The heritage forms *aku* (1sg) and *kue* (2sg) are borrowings from Malay varieties spoken on Java into Tangsi Malay (see Section

¹⁹ Since stative verbs express the properties that are typically expressed by adjectives in European languages, I refer to stative verbs as adjectives in the following chapters. This is done to facilitate the comparison with Dutch (see Chapter 3 and Chapter 6).

1.5.2). The homeland and the heritage varieties use proper nouns or kinship terms, such as *bapa* ‘father’, *usi* ‘older sister’ and *om* ‘uncle’ (<Du *oom*), to address older people or people with a higher social status.

The remaining pronouns are rather similar in both varieties. Regarding the forms for the third person singular formal, my impression is that the form *antua* is preferred in the homeland variety, while *ontua* is preferred in the heritage variety. Probably, *ontua* was the original form that has been preserved in the heritage language, while the form *antua* is a later development that has become very widespread in the Central Moluccas. The full forms of the plural pronouns are based on the singular pronouns + the noun *orang* ‘people’ (e.g., *dia* ‘3SG’ + *orang* ‘people’ = *dorang* ‘3PL’). Nowadays, however, the shorter forms are more commonly used.

Table 1.4: The pronoun system of homeland Ambon Malay and heritage Ambon Malay.

	HOMELAND AMBON MALAY	HERITAGE AMBON MALAY
1SG	<i>beta</i>	<i>aku, a, beta</i>
2SG	<i>ose, se, ale</i>	<i>kue</i>
3SG	<i>dia</i>	<i>dia</i>
3SG.FML	<i>antua, ontua</i>	<i>antua, ontua</i>
3SG.N	<i>akang</i>	<i>akang</i>
1PL	<i>katong, tong</i>	<i>ketorang, ketong, torang, tong</i>
2PL	<i>dorang, dong</i>	<i>kemorang, kemong</i>
3PL	<i>dorang, dong</i>	<i>dorang, dong</i>

1.6.2.5 Demonstratives

The Ambon Malay demonstrative system is based on the relative distance from the speaker. The system has a two-way split for demonstrative pronouns and for similitive demonstratives, and a three way split for locative demonstratives (see Table 1.5 on the next page).²⁰ The demonstrative pronouns are not only used for spatial orientation, but also for anaphoric and temporal reference (see Chapter 3).

²⁰ This table is a simplified version of Table 3 presented in Cleary-Kemp (2007, p. 330).

Table 1.5: The demonstrative system of homeland Ambon Malay and heritage Ambon Malay.

	PRONOUN	SIMILATIVE	LOCATIVE
Near speaker	<i>ini/in/ni</i> ‘D.PROX’	<i>bagini</i> ‘like.this’	<i>sini</i> ‘LOC.PROX’
Away from speaker	<i>itu/it/tu</i> ‘D.DIST’	<i>bagitu</i> ‘like.that’	<i>situ</i> ‘LOC.MED’
Far away from speaker			<i>sana</i> ‘LOC.DIST’

The full forms *ini* ‘D.PROX’ and *itu* ‘D.DIST’ can be used substantively as clausal subjects or objects. The full forms and the short forms can be used also adnominally, in which case they can either precede or follow the noun. Statistically speaking, the post-nominal position is the preferred one (see Section 3.2.1). Finally, demonstratives can co-occur in combinations, such as *ini*+NOUN+*ni*, or *itu*+NOUN+*tu*. According to van Engelenhoven (2008, p. 95), when two demonstratives co-occur, the first signals spatial distance, while the second signals speech distance, as in *itu Opel tu* ‘that Opel I talked about’.

1.6.2.6 Quantifiers and classifiers

Ambon Malay quantifiers are the numerals and the word *samua* ‘all’. Quantifiers can either precede or follow the noun they modify, and they can be accompanied by a classifier. Van Minde (1997, p. 153) states that the position of the numeral with respect to the noun depends on what “is made more prominent”. For the quantifier *samua*, van Minde (1997, p. 156) specifies that “if *samua* precedes the noun the individuality of composite entities is stressed, whereas if it follows the head noun collectivity is stressed”. Although van Minde does not express this explicitly, it is probable that what he observes for *samua* ‘all’ holds also for the other numerals. In another contact Malay variety, Papuan Malay, the order of the numerals conveys a meaning similar to that of *samua* ‘all’ in Ambon Malay: “noun phrases with preposed numerals express a sense of individuality by signaling the composite nature of their referents, [...] noun phrases with post-head numerals signal exhaustivity of definite referents” (Kluge, 2014, p. 387). In Ambon Malay, the NOUN+NUMERAL (+CLASSIFIER) order is more common, but the NUMERAL(+CLASSIFIER)+NOUN order is becoming increasingly frequent due to the influence of Standard Indonesian (see Chapter 3).

Finally, Ambon Malay has three main classifiers: *orang* ‘person’ for humans (e.g., *dua orang ana* ‘two person child’), *ekor* ‘tail’ for animals (e.g., *satu ekor gaja* ‘one tail

elephant) and *bua* ‘fruit’ for fruits and other inanimate objects (e.g., *tiga bua bola* ‘three fruit ball’). The use of classifiers is optional and rather infrequent. In the simultaneous video descriptions provided by homeland speakers (see Section 2.3), only 2.3% of all nouns occur with a classifier (the video descriptions contain a total of 4.842 noun tokens and 113 classifier tokens).

1.6.2.7 Prepositions

Ambon Malay employs prepositions to express a number of semantic relations (i.e., location, direction, recipient). The prepositional phrase usually occurs after the verb, or after the object noun phrase. The most common prepositions are listed in Table 1.6:

Table 1.6: Ambon Malay prepositions.

PREPOSITION	GLOSS	SEMANTIC RELATION
<i>di</i>	‘at, in, on, to’	LOCATIVE
<i>dalang</i>	‘in(side)’	LOCATIVE
<i>dekat</i>	‘close to’	LOCATIVE
<i>tangada</i>	‘across, opposite’	LOCATIVE
<i>dari</i>	‘from’	SOURCE, ORIGIN
<i>ka</i>	‘to’	DIRECTION
<i>for</i>	‘for, to’	RECIPIENT, BENEFACTIVE, PURPOSE
<i>par</i>	‘for, to’	RECIPIENT, BENEFACTIVE, PURPOSE
<i>buat</i>	‘for, to’	RECIPIENT, BENEFACTIVE, PURPOSE
<i>dengang (deng)</i>	‘with’	INSTRUMENTAL, COMITATIVE
<i>sampe</i>	‘until’	DIRECTION (TOWARD A TEMPORAL ENDPOINT)

The preposition *for* is a loan from Dutch *voor* ‘for, for the benefit of’, or probably a combination of Dutch *voor* and Portuguese *por* ‘to, for, through, because of’. *Par* is a loan from Portuguese *para* ‘for, to, in order to’. Chapter 5 and Chapter 6 discuss the differences in preposition selection between homeland speakers in Ambon and heritage speakers in the Netherlands.

1.6.2.8 Conjunctions

Ambon Malay has coordinating and subordinating conjunctions. The main coordinating conjunctions are: *tapi*, *mar* (<Du *maar*) ‘but’; *jadi*, *des* (<Du *dus*) ‘so’; *la/lalu*, *lantas* ‘and then’; *tarus/trus* ‘and then, next’. The form *mar* ‘but’ is more colloquial than *tapi* ‘but’, while the forms *des* ‘so’ and *lantas* ‘and then’ are more common in the heritage variety than in the homeland variety. Finally the forms *jadi* ‘so’ and *tarus* ‘and then, next’ are also verbs meaning ‘to become’ and ‘to continue’, respectively.

The main subordinating conjunctions are: *kata* ‘that’ and *yang* ‘REL, that’ (for complement clauses); *yang* ‘REL, that’ (for relative clauses); *waktu* ‘when’ (for temporal clauses); *par* ‘to, for’ and *supaya* ‘so that’ (for purpose clauses); *biar* ‘although’ (for concessive clauses); *barang* ‘because’ and *tagal* ‘because’ (for reason clauses); *kalo* ‘if’ (for conditional clauses); and *padahal* ‘whereas, but actually’ (for contrastive clauses).

1.6.3 Word formation

Ambon Malay mainly has two word formation processes: affixation (1.6.3.1) and reduplication (1.6.3.2).

1.6.3.1 Affixation

Ambon Malay has a number of derivational affixes. Some are fully productive, some are only marginally productive, while others are unproductive. The main affixes are discussed here below.

- The prefix *ta-* ‘ACL’ is fully productive. It attaches to verbs and adds the meaning of ‘sudden or unintentional’ (e.g., *rabe* ‘to tear’, *ta-rabe* ‘torn unintentionally’).
- The prefix *baku-* ‘RECP’ is also fully productive. It attaches to verbs and indicates reciprocity (e.g., *pukol* ‘hit’, *baku-pukol* ‘to fight with each other’).
- The prefix *ka-* ‘ORD’ attaches to cardinal numbers to form ordinal numbers (e.g., *dua* ‘two’, *ka-dua* ‘second’).
- The prefix *ba-* is marginally productive. It derives intransitive verbs from nouns and (transitive) verbs (e.g. *dara* ‘blood’, *ba-dara* ‘to bleed’; *kumpul* ‘to collect’, *ba-kumpul* ‘to come together’). Some verbs prefixed with *ba-* have the same meaning as their bases (e.g., *pikir* ‘to think’, *ba-pikir* ‘to think’).

Demonstratives specify the distance of a referent in both spatial and psychological terms. Additionally, they can also be used to indicate definiteness. The short forms, *ni* ‘D.PROX’ and *tu* ‘D.DIST’, in particular, have article-like functions. In (2), *tu* ‘D.DIST’ is used to keep track of the new referent (the cake) that has been introduced. It signals that the referent of *kue* ‘cake’ is identifiable by both speaker and hearer.

- (2) *Ada se-ekor tikus yang badiri di muka kue*
 EXIST one-CLF mouse REL stand at face cake

trus dia ciom-ciom kue tu
 next 3SG ITER-smell cake D.DIST
 ‘There is a mouse standing in front of a cake, then it (repeatedly) smells the cake.’

The quantifiers can precede or follow the noun (see Section 1.6.2.6 and see Chapter 3), as shown in (3).

- (3) *satu dos – dos satu*
 one box box one
 ‘one/a box’

The possessive noun phrase is formed with the possessive marker *punya* (or its phonological variants *pung/pu/ng/zero*) and has the structure ‘Possessor + *punya* + Possessed’, as shown in (4).

- (4) *dia pung dos*
 3SG POSS box
 ‘his/her box’

The genitive noun phrase follows the noun it modifies. Generally speaking, such a sequence of two nouns indicates that there is some kind of relation between the first noun (head noun) and the second one (genitive noun phrase). This is illustrated in (5), where the genitive noun phrase *kore api* ‘matches’ is an attribute of the head noun *dos* ‘box’.

- (5) *dos kore api*
 box match fire
 ‘box of matches’

Finally, stative verbs, such as *basar* ‘be big’, and relative clauses follow the noun they modify, as shown in (6) and (7), respectively.

- (6) *Dia menyimpan bola-bola di dalam dos basar*
 3SG store PL-ball at inside box big
 ‘He puts the balls into a big box.’

- (7) *Dia badiri nganga baju yang ada di atas pohong*
 3SG stand look.at cloth REL EXIST at above tree
 ‘He stands looking at the T-shirt that is (stuck) on the tree.’

All the modifiers can co-occur, as illustrated in (8), where the noun *talingang* ‘ear’ is modified by a possessive noun phrase, a quantifier, and a demonstrative.

- (8) *tikus pung talingang dua tu*
 mouse POSS ear two D.DIST
 ‘the two ears of the mouse’

1.6.5 The verb phrase

Verbs can be modified by aspect and mood particles, and by adverbs (see Section 1.6.2.3). Aspect and mood particles precede the verb. The main Ambon Malay aspect markers are *ada* ‘EXIST’, *su* ‘PRF’ and *mau/mo* ‘want’ (see Chapter 4). The main modal auxiliaries are *musti* ‘must’, *bisa* ‘can’, and *mau* ‘want’. Like other preposed possessor languages, Ambon Malay does not have voice alternation; verbs occurring with the actor voice prefix *me-* or *ma-* and the undergoer prefix *di-* are borrowed from Indonesian.²²

²² Symmetrical voice alternation is a typical feature of some Austronesian languages. In the symmetrical voice alternation, there are two voices marked morphologically on the verb (actor voice and undergoer voice) neither of which is clearly the basic form (Himmelmann, 2005).

The existential verb *ada* ‘to be (somewhere)’ can function as a marker of progressive aspect²³, as shown in (9), but in some contexts it can also have a perfect meaning, as illustrated in (10) (see Section 4.3.1).

- (9) *Parampuang* *ada* *tidor*
 girl EXIST sleep
 ‘A girl is sleeping.’
- (10) *Bungkusang* *ada* *datang*
 parcel EXIST come
 ‘The parcel has come.’ (van Minde, 1997, p. 191)

The marker *su* ‘PRF’ is an aspect-temporal marker whose functions overlap with those of the English perfect (*have –ed*) and of the adverb ‘already’, as illustrated in example (11) (see also Section 4.3.1)²⁴.

- (11) *Padahal* *gaja* *su* *makang* *kue* *tu*
 but.actually elephant PRF eat cake D.DIST
 ‘(The mouse wants to eat the cake), but actually the elephant has already eaten the cake.’

The marker *mau* ‘want’ has a long form *mau* and a short form *mo*. The long form *mau* mainly functions as a modal indicating volition, as illustrated in (12).

- (12) *Dia* *mau* *makang* *pisang*
 3SG want eat banana
 ‘She/he wants to eat a banana.’

²³ A typical feature of contact Malay varieties of eastern Indonesia is that the existential verb *ada* ‘to be (somewhere)’ also functions as a progressive aspect marker (Adelaar, 2005).

²⁴ Dahl (quoted in Olsson, 2013, p. 4) has coined the term ‘iamitive’ to refer to more or less grammaticalized markers in Southeast Asian languages that have functions shared by ‘already’ and the perfect. Unfortunately this term is still uncommon and it remains obscure to many scholars. For this reason here *su* is glossed as perfect ‘PRF’.

The short form *mo* functions mainly as a marker of prospective aspect, as illustrated in (13) (see also Section 4.3.1).

- (13) *Supaya baju mo jato*
 so.that cloth want fall
 ‘(The boy throws a ball against the tree,) so that the cloth will fall (from the tree).’

The modal auxiliary *musti* ‘must’ indicates necessity and obligation, as illustrated in (14) (see Moro, 2015). In (14), the speaker reports that, at the workplace, he and his colleagues are required to speak Indonesian, the standard language.

- (14) *Katong musti pake bahasa Indonesia*
 1PL must use language Indonesia
 ‘We have to speak Indonesian.’

The modal auxiliary *bisa* ‘can’ expresses possibility or ability, as illustrated in (15).

- (15) *Katong bisa pake bahasa Indonesia*
 1PL can use language Indonesia
 ‘We can speak Indonesian.’

1.6.6 The clause

Ambon Malay has verbal clauses (1.6.6.1) and non-verbal clauses (1.6.6.2).

1.6.6.1 Verbal clauses

In verbal clauses, the verb functions as the predicate. The basic order in verbal clauses is SV(O), with prepositional phrases following the verb in intransitive clauses (16), and the object noun phrase following it in transitive clauses (17).

- (16) *Kuda deng anjing ada lari di lapang*
 horse with dog EXIST run at field
 ‘A horse and a dog are running in a field.’

- (17) *Parampuang rabe kaeng deng dia pung tangang*
 girl tear cloth with 3SG POSS hand
 ‘A girl tears a piece of cloth with her hands.’

A specific type of intransitive clause is the existential clause, which is introduced by the existential *ada* ‘to be (somewhere)’. In existential clauses, the prepositional phrase can precede or follow the verb and the argument. An example is given in (18).

- (18) *Ada botol dalang bakol*
 EXIST bottle inside basket
 ‘There is a bottle in a basket.’

1.6.6.2 Non-verbal clauses

The most common types of non-verbal clauses are clauses where a noun or a preposition functions as the predicate. These clauses usually have an equative or a locative function, as shown in (19) and (20) respectively.

- (19) *Bapa dolo guru*
 father earlier teacher
 ‘(My) father (was) a teacher.’
- (20) *Dia su di pinggir*
 3SG PRF at side
 ‘She/he (is) already on a side.’

1.6.7 Verb serialization

In Ambon Malay, verb sequences or serial verb constructions (SVCs) are used to convey various meanings, such as directional (21), passive (22), or resultative (23). Verbs in a SVC share the same aspect/mood/negation marker and they share at least one argument.

- (21) *Dia bajalang kaluar*
 3SG walk exit
 ‘He walks out.’

- (22) *Ana satu dapa pukol*
 child one get hit
 ‘One/a child got beaten.’

- (23) *Dia su pata akang jadi dua*
 3SG PRF break 3SG.N become two
 ‘She/he has broken it into two (lit. She/he has broken it became two).’

1.6.8 Negation

Negation is expressed by independent lexical morphemes, which usually precede the predicate. The negators *seng* ‘NEG’ and *balong* ‘not yet’ negate verbal, existential, and nonverbal prepositional clauses, as shown in (24) and (25).

- (24) *Dia seng maeng bola*
 3SG NEG play ball
 ‘She/he doesn’t play with the ball.’

- (25) *Dolo-dolo balong ada televisi*
 INTENS-earlier not.yet EXIST television
 ‘In the old days there was no television yet.’

The negator *bukang* ‘NEG’ negates non-verbal nominal clauses, as shown in (26), while *jangang* (*jang*) ‘PROH’ negates imperatives, as shown in (27). The negator *jang* can also negate purpose clause introduced by *supaya* ‘so that’ (see Section 1.6.2.8), as illustrated in (28).

- (26) *Katong bu kang orang Seram*
 1PL NEG person Seram
 ‘We are not people from Seram Island.’

- (27) *Jang makang ikang*
 PROH eat fish
 ‘Don’t eat fish.’

(28) *Tikus ika talinga supaya jang dapa dengar*
 mouse tie ear so.that PROH get hear

gaja pukol poro
 elephant hit belly

‘The mouse ties his ears so that it doesn’t get to hear (the noise of) the elephant hitting its belly.’

1.7 Overview of the chapters

This dissertation consists of an integrated collection of articles reporting on a number of case studies of heritage Ambon Malay grammar. The case studies are linked together by a common theoretical and methodological framework (see Section 1.3 and Section 1.4) and make use of the same dataset (see Chapter 2). Even though each chapter focuses on a particular aspect of heritage Ambon Malay grammar, the reader may find some overlap and duplication among the chapters, particularly in the introduction and methodology sections. This arrangement provides the opportunity to read the chapters separately, or in a different order, depending on where one’s interests lie.

Chapter 2 illustrates the methodology used in the present study to sample the speakers, collect data and analyze the results. The chapter first describes which types of speakers were included in the sample, how they were selected, and when and where the fieldwork was conducted. Subsequently, the chapter discusses the problems of identifying the proper Ambon Malay baseline group and issues related to data collection in the first generation group. The second part of the chapter illustrates the elicitation material used for the present research and the procedures employed to transcribe, code and analyze the data.

Chapter 3 examines word order changes in the domain of nominal modification. In Ambon Malay, demonstratives, numerals, adjectives, and definite markers occur after the noun (i.e., post-nominal position), but demonstratives and numerals can also occur before the noun (i.e., pre-nominal position). In Dutch, nominal modifiers always occur in pre-nominal position. The chapter shows that partial overlap in surface structure between Ambon Malay and Dutch has a concrete effect on word order in heritage Ambon Malay. This effect manifests itself as an increase in the frequency of word order patterns shared by Dutch (i.e., pre-nominal demonstratives,

and pre-nominal numeral ‘one’). The chapter also tests whether the change in the linear order of these morphemes correlates with their grammatical re-interpretation on the model of Dutch articles. The results show that no grammatical re-interpretation process has taken place yet. What may be undergoing an initial process of contact-induced grammaticalization is the definite marker =*nya*, which is found to occur statistically more frequently in heritage speakers than in homeland speakers. The chapter argues that the high frequency of =*nya* ‘DEF’ in heritage speakers may be partly due to grammaticalization of the category ‘definiteness’ as a result of Dutch influence, and partly due to the different language variety heritage speakers were exposed to (Tangsi Malay).

Chapter 4 examines the aspectual system of heritage Ambon Malay. The chapter first describes the tense/aspectual distinction of homeland Ambon Malay and Dutch, and then investigates signs of divergence and convergence by comparing the frequency of usage of aspect markers in the heritage group and in the three control groups. The findings reveal that the heritage variety diverges from the homeland variety in two ways: the marker *ada* ‘EXIST’ is overextended to new contexts, whereas the marker *su* ‘PRF’ and verbal reduplication are used significantly less often. The frequency of the marker *mau* ‘want’ seems rather stable. Following grammaticalization theories and the findings of other studies, the chapter argues that, in addition to functioning as a progressive marker, *ada* has acquired the function of a present tense/finiteness marker, an innovation that is arguably due to cross-linguistic influence from Dutch. The chapter also argues that language-internal factors, such as form-meaning mapping, frequency and acoustic salience, play an important role in determining the maintenance versus loss of aspectual forms.

Chapter 5 focuses on the expression of *give*-events in heritage Ambon Malay. The chapter describes the different strategies in which *give*-events are expressed in Ambon Malay and Dutch. It shows that both languages allow the ‘Double Object (DO) construction’ (*John gave Mary a book*) as well as the ‘Prepositional Object’ (PO) construction (*John gave a book to Mary*), although with different preferential tendencies. Ambon Malay always prefers PO, while Dutch has a bias for DO in corpus data and for PO in de-contextualized elicited data. Furthermore Ambon Malay can express *give*-events using two predicates in a single sentence (i.e., the ‘two predicate construction’). A quantitative analysis of the data reveals that heritage speakers use the DO construction significantly more often than homeland speakers, while they use the ‘two predicate construction’ significantly less often. The chapter argues that cross-linguistic influence from Dutch and universal principles in

language development under reduced input are responsible for these changes in frequency. The chapter illustrates that there are also qualitative differences between heritage and homeland speakers of Ambon Malay: the different prepositions both groups choose for the PO construction reflect their different social histories and the different type of input heritage speakers were exposed to (Tangsi Malay).

Chapter 6 takes resultative constructions as a case study. The chapter describes the various strategies in which resultative events are expressed in Ambon Malay and Dutch. It shows that, even though Ambon Malay prefers serial verb constructions (SVC) (*She breaks a stick becomes two*) and Dutch prefers verb particles (*She cuts off a branch*), both languages also allow preposition phrases (PP) (*She breaks a stick in two*) and adjectival phrases (AP) (*She hits a vase broken*). A quantitative analysis of the data reveals that heritage speakers use SVCs significantly less often than homeland speakers, whereas they use PPs and APs significantly more often. The shift of heritage speakers away from the encoding preferences of homeland speakers, and towards those of Dutch speakers clearly identifies cross-linguistic influence from Dutch as the main source for divergence. The chapter also discusses changes in the choice of prepositions used in PP constructions. Heritage speakers of Ambon Malay predominantly use the prepositions *ka* ‘to’ and *dalang* ‘in(side)’ (*ka dua* ‘into two’) The chapter argues that the semantic extension of these prepositions is an internal process driven by universal principles, which is accelerated by contact with Dutch.

Chapter 7 brings together the findings of all preceding chapters to examine the structural and social factors responsible for the patterns of divergence and convergence observed in various areas of heritage Ambon Malay grammar. The first part of the chapter focuses on structural factors. By using hierarchical cluster analysis, the chapter shows that the linguistic features of heritage Ambon Malay fall into two main clusters: the innovative ‘Dutch-like’ features and the more conservative ‘Malay-like’ features. The Dutch-like features are all those features adopted by heritage speakers to maximize the compatibility between Ambon Malay and Dutch. The chapter uses Hartsuiker’s (Hartsuiker et al. 2004; Schoonbaert et al. 2007; Hartsuiker & Pickering 2008) psycholinguistic model of bilingual processing to illustrate how speakers can copy frequency patterns from one language to the other with the aim of increasing the similarity of the two systems. The second part of the chapter, which focuses on social factors, individuates the *place where the speaker lives* as the best predictor of linguistic innovations. Speakers living outside a Moluccan ward have a higher rate of Dutch-like features than speakers living in a Moluccan ward. Furthermore, the chapter shows that the interaction between *age of*

onset of bilingualism and *place where the speaker lives* can have important effects on the linguistic performance of heritage speakers.

Chapter 8 concludes the dissertation by answering the research questions posed in Chapter 1, and by giving suggestions for further research.

CHAPTER 2

Methodology

The following sections illustrate the methodology used in the present study to sample the speakers, to collect data and to analyze the results. Section 2.1 describes which types of speakers were included in the sample, how they were selected, and when and where the fieldwork was conducted. Section 2.1.1 discusses the problem of identifying the right Ambon Malay baseline group and issues related to data collection in the first generation group. Section 2.1.2 gives a detailed description of the heritage Ambon Malay participants (section 2.1.2.1), of the homeland Ambon Malay participants (section 2.1.2.2), of the first generation Ambon Malay participants (section 2.1.2.3) and of the Dutch participants (section 2.1.2.4). Section 2.2 illustrates the elicitation material used for the present study, while section 2.3 presents the corpus used to conduct the present research. Finally section 2.4 outlines the procedures employed to transcribe, code and analyze the data.

2.1 Speaker sample

Speakers were sampled from four populations: (i) heritage speakers of Ambon Malay in the Netherlands (early bilinguals), (ii) homeland speakers of Ambon Malay in Ambon, Indonesia, (the homeland); (iii) first generation speakers of Ambon Malay in the Netherlands (late bilinguals); and (iv) speakers of Dutch in the Netherlands. A detailed list of the participants is presented in the following sections. Heritage speakers are the so-called ‘experimental group’: they represent the individuals under investigation. The other three groups, the homeland speakers, the first generation speakers and the Dutch speakers are so-called ‘control groups’, they are the individuals with whom heritage speakers are compared (see Section 2.1.1 and Section 1.2.1).

Speakers in the four groups were selected on the basis of their linguistic autobiography, following the sociolinguistic approach outlined in Nagy (2015). Heritage speakers are individuals with Moluccan origins, who grew up and currently live in the Netherlands, and who are (early) Dutch-Ambon Malay bilinguals with Dutch as their dominant language. No proficiency test was used as a selecting

criterion in order to avoid the risk of excluding very innovative speakers.²⁵ As it is pointed out by Nagy (2015, p. 314), if we wish to cast a wide net, in order to be able to describe the range of performance of all types of heritage speakers, we should refrain from administering proficiency or fluency tests as inclusion/exclusion criteria.

Speakers in the three control groups were also selected with reference to their autobiographical characteristics. Homeland speakers are individuals who grew up and currently live in the Central Moluccas and who are (mostly) Ambon Malay monolinguals. The same holds for the Dutch group, which is made by individuals who grew up and currently live in the Netherlands and are Dutch monolinguals. First generation speakers are defined on the basis of autobiographical and linguistic criteria. First and foremost, they are individuals with central Moluccan origins. Furthermore, many of them are typically ex-KNIL soldiers or family members of the soldiers, and they speak Tangsi Malay or Ambon Malay with a strong Tangsi Malay influence. They arrived in the Netherlands after age 14 (the majority around their 20s), and therefore classify as late Dutch-Ambon Malay bilinguals.²⁶ Note that in this dissertation, first generation speakers are defined by the age of onset of bilingualism, and not by “year of birth” (as was done in other studies, such as Huwaë, 1992, and Veenman, 1994). The age of onset of bilingualism is a valid criterion for distinguishing adult bilingualism from child bilingualism, whereas the “year of birth” is not informative in this respect.

Having illustrated which types of speakers were selected, I now turn to how they were selected. The participants were selected in a quasi-random fashion in order to meet the requirement of representativeness, which ensures that the sample selected is representative of the whole population (Tagliamonte, 2006, p. 23). This requirement is accomplished by selecting individuals on the bases of age, sex, and place of birth, so that the linguistic diversity in the targeted community is represented in the sample (as much as possible).²⁷ The selection criteria are labeled quasi-random because the researcher first identifies the categories of speakers to be

²⁵ Since speakers had to perform a production tasks, only heritage speakers who were able and willing to speak were included (‘overhearers’ and ‘weak’ speakers were excluded).

²⁶ Most linguistic studies consider age 12 as the threshold for late bilingualism (Montrul, 2008, among others).

²⁷ Socio-economic class and educational level were not included as variables because the Moluccan community has rather homogenous socio-economic characteristics (see Section 1.5.2).

studied, and then selects randomly within the categories. In the present study, the selection within the categories was done by means of the social network approach, a type of approach based on the concept of social ties or 'friend of a friend' network (Tagliamonte, 2006, p. 21).

In the social network approach, the researcher is introduced into the speech community by an intermediary who presents the research as a 'friend' to his or her group. The main benefit of using this approach is that the researcher selects individuals from a pre-existing social group and not from an abstract category. Furthermore, the fact that the researcher is not perceived as an outsider by the community ensures spontaneous and reliable data. According to Tagliamonte (2006, p. 28), combining (quasi) random sampling and the social network approach is the most fruitful fieldwork technique because "whereas random survey methods ensure representativeness of the sample, a social network approach goes a long way towards mitigating the observer's paradox and reaching the right people".

The data collection for the present research was conducted following these guidelines. The fieldwork in the Moluccan community in the Netherlands was carried out by Rose Lekawael,²⁸ a Moluccan MA student in Nijmegen who had personal ties with members of the community, and by myself. Rose Lekawael collected data in the area of Middelburg, Vlissingen and Nijmegen, where many Moluccans, including her own family members, live. I was introduced into the community by Otjep Rahantoknam, an employee of the *Landelijk Steunpunt Educatie Molukkers* (National Support Centre for the Education of the Moluccans), who provided me with names, telephone numbers and email addresses of his relatives and friends, who kindly accepted to be interviewed, and who, in turn, introduced me to other relatives, friends and neighbors. Furthermore, I recruited some participants through an advertisement that I was able to place in the magazine *Marinjo*, thanks to Otjep Rahantoknam.²⁹

The original plan was to limit data collection to the areas of Middelburg-Vlissingen and Nijmegen, but this turned out to be very hard as I did not find an 'intermediary' in those areas. In addition, the 'friend of a friend' approach and the

²⁸ Some of the interviews were also conducted with the help of two students doing an internship, Jasmianti Garing and Feny Eky.

²⁹ *Marinjo* is an independent Moluccan magazine published bimonthly since 1995, where Moluccans can read articles (mainly written in Dutch) concerning cultural activities, sports, music, history and politics.

advertisement put me in contact with people living in other areas. As a consequence, I decided to broaden the area of investigation, and I collected data in the area of Amsterdam, Breukelen, Eindhoven, Leiden, Lunteren, Nijmegen, Uden, Venlo, and Zwolle. Fieldwork in the Netherlands was conducted over the period March-May 2011 by Rose Lekawael, and in the period March 2012- April 2013 by myself.

Fieldwork in Ambon was also conducted partly by Rose Lekawael and partly by me. Rose Lekawael, who returned to live in Ambon after graduation, collected data in the village of Amahusu (Ambon), where she interviewed her relatives, her neighbors, and her friends. I collected most of the data in Ambon city. The participants I interviewed were members of my host family, their relatives, their neighbors and their friends. I also went to the village of Amahusu to record three participants. Furthermore, two heritage speakers in the Netherlands put me in contact with their relatives in Ambon, who I also managed to interview. Fieldwork in Ambon was conducted over the period November 2011-February 2012 by Rose Lekawael, and in the period September-October 2012 by myself.

Dutch participants in the Netherlands were interviewed by Rowan Soolsma, a Dutch MA student in Nijmegen, who collected data in the Middelburg-Vlissingen area over the period May-October 2012. She selected the participants among her own family members and friends.

2.1.1 The baseline groups: issues and problems

The baseline group is the control group against which the heritage group is to be compared. We have seen in Section 1.2.1 of the previous chapter that heritage speakers can be compared to different control groups (i.e., monolingual adults, monolingual children, late bilinguals, L2 learners, illiterate monolingual adults) in order to answer different questions regarding heritage grammars, such as does bilingualism matter? Does the age of acquisition matter? Does formal schooling matter?. In the present research, the heritage group is compared systematically to (i) homeland speakers, (Ambon Malay “monolingual” adults³⁰), (ii) to first generation speakers (Dutch-Ambon Malay late bilinguals), and (iii) to Dutch speakers (Dutch monolingual adults).

³⁰ Strictly speaking, speakers of Ambon Malay in the homeland are not monolingual. Many of them also speak Standard Indonesian, while some of them also speak an indigenous language (see Section 2.1.2.2). I refer to them as monolingual here for the sake of comparison with the other two groups.

Homeland adults in Ambon form the first baseline group, because they are the monolingual control. In other words, their language represents the language variety without any possible significant Dutch influence (posterior to 1951). Note that, the heritage speakers' baseline language is Ambon Malay, and not Standard Indonesian, or Jakarta Indonesian, which are the languages used in the school system, literature, and the media. Heritage speakers speak this particular geographical variety (to various degrees), and thus are to be compared with speakers of this variety (the issue of dialect variation in heritage language acquisition is addressed in Polinsky & Kagan, 2007, p. 372; see also Section 2.1.2.1 and Section 2.1.2.2). As shown in the next sections, homeland speakers were selected from the same geographical regions in order to match the origins of heritage speakers. Obviously, since 1951, Ambon Malay in the Central Moluccas has not existed in a vacuum; it has been and still is under the constant and increasing influence of Indonesian. This factor, however, does not undermine the importance of the homeland baseline; as long as a possible Indonesian influence is controlled for in the data analysis (see Chapter 3). Another factor, which is discussed in Chapter 5, is the influence of Standard Indonesian on the heritage variety in the Netherlands.

First generation speakers in the Netherlands form the second baseline group. They represent the 'true' baseline, as they are the ones who provided the actual input for the heritage speakers under investigation. The language they spoke is the language heritage speakers acquired. Many first generation speakers spoke Tangsi Malay, a divergent form of Ambon Malay with some Javanese and Dutch elements which developed in the barracks of the KNIL army (see Section 1.5.2.1). Since Tangsi Malay is now extinct, first generation speakers are the only source we have to individuate traces of Tangsi Malay in the heritage speakers' grammar (see Chapter 3 and Chapter 5). There are, however, two main issues related to first generation speakers. The first is that, being late bilinguals, first generation speakers present a considerable degree of attrition in their language; this is especially true for the speakers with a long period of residence in the Netherlands (about 60 years)³¹. As a consequence of attrition, the language that they speak now and that we record is not

³¹ Attrition begins already in the first decade of residence in a foreign country. Schmidt (2011), however, shows that length of residence is not the only predictor of language attrition, but it interacts with the amount of L1 use. For instance, the impact of length of residence on attrition is stronger for migrants who use the L1 very infrequently and those who use it extremely frequently. For individuals with an intermediate L1 use, length of residence does not play a significant role.

likely to be the same language that they spoke to their children 40-50 years ago. Furthermore some speakers show considerable aging effects in their language (e.g., a very slow speech rate). The second problem is that, with few exceptions, first generation speakers are quite old (about 80 years). It turned out to be rather difficult to recruit participants in this group, because many of them have unfortunately already passed away. In addition, some were too ill or weak to properly describe stimuli presented on a laptop, while others were extremely shy or reluctant to take part in an interview. Due to these factors, we were able to collect data only from six speakers. As a result, the comparison between heritage speakers and first generation speakers carried out in the following chapters is qualitative in nature.

Dutch monolingual speakers and Dutch grammar are also used as a control. Although Dutch cannot be considered the baseline language of heritage speakers, a comparison of heritage Ambon Malay to Dutch is fundamental because Dutch is, after all, the dominant language of heritage speakers. However, in order to understand the possible influence of Dutch on heritage Ambon Malay, careful attention must be paid to Dutch spoken input, as opposed to prescriptive Dutch grammar as described in textbooks. In the following chapters, Dutch grammar is used as a control, when the grammatical domain under investigation does not present any variation (e.g., the order of nominal modifiers with respect to the noun is completely fixed in Dutch; see Chapter 3), while data from Dutch speakers are used, when the domain under investigation does present some variation (e.g., the use of tense-aspect markers, see Chapter 4; the use of *give*-constructions, see Chapter 5; and the use of resultative constructions, see Chapter 6)

Finally, we have seen in Section 1.2.1 that heritage speakers can also be compared to monolingual children, to L2 learners and to illiterate monolingual adults. A comparison with monolingual children and L2 learners was not carried out in this dissertation due to the lack of data. As far as I know, no data on child acquisition of Ambon Malay are available, and collecting such data for the present study would have been too complex and time consuming considering the purpose of this research. Data from L2 learners are also not available as Ambon Malay is not a language taught in schools, be it in Ambon or abroad. The Moluccan Malay course made available by the LSEM, which I attended, targets a rather high variety of Ambon Malay which combines Standard Indonesian with Moluccan expressions and cultural values. Lastly, a comparison with illiterate speakers was not strictly necessary because Ambon Malay does not have specific constructions that are

acquired through formal education (as it is the case for inflected infinitives in Portuguese, see Pires & Rothman, 2009). Generally speaking, the properties of the spoken language are the same as those of the written language.

2.1.2 Participants

A total of 80 people participated in the study.³² Of these, 32 are heritage speakers of Ambon Malay in the Netherlands, 27 are speakers of Ambon Malay in Ambon, six are first generation speakers of Ambon Malay in the Netherlands, and ten are speakers of Dutch in the Netherlands. The participants in the four groups were matched as much as possible for age, sex and geographical area.

2.1.2.1 Heritage speakers in the Netherlands

The heritage speaker sample includes 16 females and 16 males (tot=32), with an age ranging from 14 years to 62 years old ($M = 43.8$, $SD = 12.6$) (see Table 2.1 on the next pages). The sample contains six pairs of siblings, four pairs of spouses, and four parent-child pairs.

The majority of the heritage speakers in the sample (19 speakers) belong to the second generation, meaning that they were either born in the Netherlands with first generation parents, or they arrived in the Netherlands at a very early age (three second generation speakers were born in Indonesia but arrived before the age of two). Other seven participants belong to the generation 2.5, they have one parent from the first generation and the other parent from the second generation. Finally, six speakers belong to the third generation as they have both parents from the second generation. Twenty-eight speakers have two Moluccan parents, while only four speakers (two from the second generation and one from the third generation) have a Dutch mother and a Moluccan father.

In order to control for dialect variation as much as possible, we tried to select only heritage speakers with parents originating from the Central Moluccas. This, however, was not always possible as we had to consider the trade-off between time to complete the data collection and optimal sampling. As a consequence, the sample contains twenty-three heritage speakers with Central Moluccan origin (Ambon,

³² Five speakers, four from Ambon and one from the Netherlands, who participated in the experiment, were excluded because the Ambon Malay variety that they used showed too much influence from Standard Indonesian. Despite the various requests for speaking 'the everyday language', they performed the task in Indonesian.

Haruku, Nusa Laut, Saparua, and Seram), six speakers with probable ‘mixed’ origins (two speakers of Central Moluccan and Javanese origin; four speakers of Central Moluccan and South Moluccan origin), two speakers with South Moluccan origin, and one speaker with North Moluccan origin.

All speakers in the sample acquired Ambon Malay (with some Tangsi influence) from birth in a naturalistic setting; typically at home. The speakers with non-Central Moluccan parents also acquired Ambon Malay as they grew up in camps or wards where Tangsi Malay was commonly spoken. The age of onset of acquisition (AoA) of Dutch varies: ten speakers are sequential bilinguals (AoA of Dutch: from age four or five), while 22 speakers are simultaneous bilinguals (AoA of Dutch: from birth). For all speakers, Dutch is the functionally dominant language. Twelve speakers report mainly speaking Ambon Malay with both parents during childhood,³³ eight report speaking Ambon Malay with one parent, and Dutch with the other, while 12 speakers report mainly speaking Dutch with both parents. Nevertheless, this group of 12 was still exposed to their parents speaking Ambon Malay with each other and with others. Furthermore, they also acquired the language through interactions with other community members such as grandparents, friends, and neighbors. Some of them report being ‘overhearers’ until puberty and becoming more active toward the age of 15 and later on when they started going out with friends, joining Moluccan organizations, or marrying a Moluccan partner.

All the speakers report having at least some basic knowledge of English, with some even reporting intermediate or advanced knowledge. About half of the speakers report knowing some German, while a few report knowing French (English, German and French are the languages usually taught in Dutch schools). One speaker is very fluent in Czech because he worked in Czech Republic for few years.

In terms of place of upbringing, ten speakers grew up in Moluccan camps, 13 grew up in Moluccan wards or municipalities, and nine grew up in cities. Of the ten speakers growing up in camps, eight moved to a Moluccan municipality, while two moved to a city. Of the 13 speakers growing up in Moluccan municipalities, 11 remained, and one moved to a city. Of the nine speakers growing up in cities, seven remained in the city, while three moved to a Moluccan municipality. To sum up, 22 speakers live in a Moluccan municipality and ten live in a city. A breakdown of this demographic information is listed in Table 2.1 on the next pages:

³³ Two speakers who report mainly speaking Ambon Malay with the parents are classified as simultaneous bilinguals because their parents explicitly taught them some Dutch.

Table 2.1: Autobiographical characteristics of heritage speakers. In this Table, F = female, M = male, AoA = age of acquisition of Dutch.

Sp	Age	Sx	Gn	LIVE	GREW	AoA	L Mo	L Fa	ORIGIN Mo-Fa	Ls
H1	55	M	2	ward	camp	4-5	AM	AM	Nusa Laut- Ambon	En, Ge
H2	61	F	2	ward	camp	4-5	AM	AM	?- Ilmarang	En, Ge, Fr
H3	62	F	2	ward	camp	0	AM	AM-Du	Saparua- Saparua	En
H4	56	F	2	ward	camp	4-5	AM	AM	Haruku- Haruku	En, Ge
H5	52	F	2	ward	camp	4-5	AM	AM	?- Ilamarang	En, Ge, Fr
H6	53	M	2	ward	city	0	Du	AM	NL- Ambon	En, Ge
H9	33	M	2	ward	ward	4-5	AM	AM	Saparua- Saparua	En
H12	59	M	2	ward	camp	4-5	AM	AM	Kei Besar Kei Besar	En
H14	49	F	2	city	city	birth	Du	Du	Seram- Saparua	En
H16	57	M	2	ward	camp	6	AM	AM	Seram- Seram	En
H21	49	M	2	ward	city	birth	Du	Du	Saparua- Ambon	En
H22	50	M	2	ward	camp	birth	Du	Du	Saparua- Haruku	En, Ge
H24	62	M	2	city	camp	5-6	AM	AM	Java- Saparua	En, Spa
H25	47	F	2	ward	ward	birth	D	AM	NL- Saparua	En, Ge

The table continues on the next page.

Sp	Age	SX	G	LIVE	GREW	AoA	L Mo	L Fa	ORIGIN Mo-Fa	Ls
H26	31	F	2	ward	ward	birth	AM	AM-D	Kei Besar- Kei Besar	En
H28	59	M	2	city	camp	5-6	AM	AM	Java- Saparua	En, Ge
H30	48	F	2	city	city	birth		Du	Seram- Saparua	En
H31	31	F	2	city	city	birth	AM	Du	Ambon- Ambon	En
H32	31	F	2	city	city	birth	Du	Du	NL- Halmahera	En
H8	45	M	2.5	ward	ward	birth	Du	Du	Seram- Seram	En, Ge
H11	34	M	2.5	city	city	birth	Du	Du	?-Ambon	En
H17	45	M	2.5	city	city	birth	Du	AM	Ambon- Saparua	En, Cz, Ge
H20	14	M	2.5	ward	ward	birth	AM	Du	Seram- Ambon	En
H23	43	F	2.5	ward	ward	birth	Du	Du	Seram- Seram	En
H29	51	F	2.5	city	city	birth	Du	AM	Ambon- Saparua	En
H33	30	F	2.5	city	ward	birth	AM	Du	Tanimbar- ?	En, Ge, Fr
H7	34	F	3	ward	ward	birth	AM-Du	AM-Du	Ambon-Kei Besar	En
H13	25	F	3	ward	ward	birth	Du	Du	Haruku- Ambon	En, Ge
H15	27	M	3	ward	ward	birth	Du	Du	Ambon- Ambon	En
H18	30	M	3	ward	ward	5-6	AM	AM	Saparua- Saparua	En

The table continues on the next page.

Sp	Age	SX	G	LIVE	GREW	AoA	L Mo	L Fa	ORIGIN Mo-Fa	Ls
H19	42	F	3	ward	ward	birth	Du	Du	NL- Ambon	En, Ge, Fr
H27	37	M	3	ward	ward	birth	Du	Du	Saparua- Haruku	En, Ge

2.1.2.2 Homeland speakers in Ambon

The homeland speaker sample includes 15 females and 12 males (tot=27), with an age ranging from 17 years to 90 years old ($M=43.3$, $SD=18.3$) (see Table 2.2 on the next page). The sample contains one pair of siblings, one pair of parent-child, two pairs of spouses, and two aunt-niece pairs. Furthermore, two homeland speakers were relatives of heritage speakers in the Netherlands.

We tried to match the geographical origin of homeland speakers to the origin of the heritage speakers in order to control for dialect variation (see Section 2.1). The parents of heritage speakers in the Netherlands came predominantly from the Central Moluccas, some came from the South Moluccas or from the North Moluccas, while some were from Java or from Sulawesi. We tried to replicate this distribution in the homeland sample. Twenty-two homeland speakers are from the Central Moluccas; four have ‘mixed’ origin. Finally, one speaker is from the South Moluccas.

All homeland speakers in the sample have Ambon Malay as their mother tongue. This also includes speakers who originate from the South and North Moluccas but who grew up and currently live in Ambon. Furthermore, Ambon Malay is also spoken on the coastal areas of some South Moluccan islands. Two speakers also report speaking one of the indigenous languages, Alune, Hulaliu, Galela and Tobelo. However, they do not speak it on a daily basis. Three other speakers report having a passive knowledge of Halmahera,³⁴ Ilmarang³⁵ and Kei. Five speakers have a good

³⁴ The speaker says that when he was young he could speak *bahasa Halmahera* ‘Halmahera language’ because he was born and spent few years in Ternate. Ethnologue (Lewis et al., 2015) does not list any language under the label ‘Halmahera’, it is thus likely that the speaker is referring to one of the Papuan languages spoken on Ternate (possibly Sahu).

³⁵ The speaker reports passive knowledge of *bahasa Ilmarang* ‘Ilmarang language’, but Ethnologue (Lewis et al., 2015) does not include any language with this name. Most likely, the speaker is referring to Davelor, an Austronesian language spoken on Dawera, which is the island hosting the village of Ilmarang.

knowledge of English, while six speakers remember a few words of Dutch. None of them, however, was raised bilingually in English or Dutch. Table 2.2 summarizes these demographic details.

Table 2.2: Autobiographical characteristics of homeland speakers. In this Table, F = female, M = male.

Sp	Age	Sx	LIVE	ORIGIN Mo-Fa	LS
B1	51	F	Ambon	Sulawesi-Kei Besar	Kei
B2	32	M	Ambon	Seram-Seram	-
B3	32	F	Ambon	Haruku-Ambon	-
B4	17	F	Ambon	Ambon-Ambon	-
B5	24	F	Ambon	Ambon-Ambon	D
B6	33	F	Ambon	Ilmarang-Ilmarang	En, Du, Ilmarang
B7	38	M	Seram	Seram-Seram	Du
B8	19	F	Ambon	Ambon-Ambon	-
B9	46	M	Ambon	Ambon-Ambon	En, Du
B10	90	M	Ambon	Ambon-Ambon	Du
B11	52	M	Ambon	Halmahera-Rote	Halmahera
B12	73	F	Ambon	Haruku-Haruku	En, Du, Hulaliu; Galela; Tobelo
B13	58	F	Ambon	Seram-Seram	Alune
B14	34	M	Ambon	Seram-Java	-
B15	66	F	Ambon	Haruku-Ambon	-
B16	57	M	Ambon	Ambon-Ambon	-
B17	57	F	Ambon	Saparua-Saparua	-
B18	17	M	Ambon	Java-Nusa Laut	En
B19	40	F	Ambon	Ambon-Ambon	-
B20	22	M	Ambon	Ambon-Ambon	-
B21	32	F	Ambon	Seram-Seram	-
B22	33	F	Ambon	Haruku-Ambon	En
B23	40	M	Ambon	Saparua-Saparua	-
B24	54	M	Ambon	Seram-Seram	-
B25	52	F	Ambon	Ambon-Ambon	-
B26	68	F	Ambon	Ambon-Ambon	-
B27	32	M	Ambon	Ambon-Ambon	-

2.1.2.3 First generation speakers in the Netherlands

The first generation speaker sample includes four females and two males (tot=6), with an age ranging from 38 to 85 years old ($M=69.1$, $SD=19.5$) (see Table 2.3 on the next page). Three first generation speakers have personal ties with four heritage speakers described in the previous section: one first generation speaker is the mother of one heritage speaker, one first generation speaker is the mother of one heritage speaker and the wife of another heritage speaker, and one first generation speaker is the uncle of one heritage speaker.

First generation speakers are classified as such, because they all arrived in the Netherlands after puberty, five speakers arrived between the ages 21-24, while one arrived at the age of 14. Their length of stay in the Netherlands goes from a minimum of 16 years to a maximum of 62 years ($M=48.6$, $SD=20.0$). Four first generation speakers arrived in 1951, one is an ex-KNIL soldier and three are family members of the soldiers. They all have parents originating from the Central Moluccas. Some of them were born or grew up in the barracks of the KNIL army in Jakarta and Cimahi (Java), Makasar (Sulawesi) or Banjarmasin (Kalimantan). The other two speakers came later, one in 1981 and one in 1995, because they both married a man of the Moluccan community in the Netherlands. One of them came from Ambon, while the other came from the South Moluccas (Kei Besar).

All first generation speakers speak Ambon Malay or Tangsi Malay as their mother tongue. One speaker also reports speaking Aboru,³⁶an Austronesian language of Haruku. They are considered late bilinguals, because they learned Dutch or started using Dutch consistently after their arrival in the Netherlands. Three first generation speakers report that they had learned Dutch in schools during the Dutch colonial period, but that they spoke predominantly Ambon Malay or Tangsi Malay with their family members and friends. The other three did not know Dutch prior to their arrival and learned Dutch once they arrived in the Netherlands. Nowadays, they all live in Moluccan municipalities. Table 2.3 on the next page summarizes these demographic details.

³⁶ Ethnologue (Lewis et al., 2015) does not list any language under the name 'Aboru'. The speaker is probably referring to the variety of Haruku spoken in the village of Aboru.

Table 2.3: Autobiographical characteristics of first generation speakers. In this Table, F=female, M=male.

Sp	Age	Sx	Age Arr.	Year Arr.	LOR	ORIGIN Mo-Fa	BACKGROUND INFORMATION	Ls
B28	38	F	22	1995	16	Ambon-Ambon	married a 2 nd generation speaker	Du
B29	52	F	21	1981	31	Kei Besar-Kei Besar	married a 2 nd generation speaker	Du
B30	83	F	21	1951	62	Jawa-Ambon	son of a KNIL soldier, grew up in Kalimantan	Du
B31	85	M	24	1951	61	Ambon-Ambon	KNIL soldier, spent two years in Cimahi	Du
B32	75	M	14	1951	61	Ambon-Ambon	son of a KNIL soldier, grew up in Makasar	Du
B33	82	F	21	1951	61	Seram-Seram	married a KNIL soldier, spent time in Jakarta and Kalimantan	Du, Aboru

2.1.2.4 Dutch speakers in the Netherlands

The Dutch speaker sample includes six females and four males (tot=10), with an age ranging from 19 to 84 years old ($M=39.9$, $SD=21.6$) (see Table 2.4 on the next page). The sample contains seven parent-child pairs, two pairs of spouses, one pair of siblings, and one grandparent-grandchild pair. They all live in the area of Middelburg and Vlissingen.

All the speakers in the sample are native speakers of Dutch, and none of them was raised bilingually. The majority of them have learned some English, German and French at school. Only the eldest speaker (age 84 y. o.) has no knowledge of a foreign language. All speakers report speaking some form of Zeeuws or Zeelandic, a South Western Dutch dialect. This knowledge, however, was restricted to using a few Zeelandic words and having a Zeelandic accent. None actually spoke the local

dialect fluently. Table 2.4 summarizes the demographic characteristics of the Dutch speakers.

Table 2.4: Autobiographical characteristics of Dutch speakers.

Sp	Age	Sx	LIVE	Ls
D1	19	F	Middelburg-Vlissingen	En, Ge, Fr
D2	21	M	Middelburg-Vlissingen	En
D3	49	F	Middelburg-Vlissingen	En, Fr
D4	24	M	Middelburg-Vlissingen	En
D5	51	M	Middelburg-Vlissingen	En
D6	84	F	Middelburg-Vlissingen	-
D7	23	F	Middelburg-Vlissingen	En, Ge, Fr
D8	52	M	Middelburg-Vlissingen	En
D9	55	F	Middelburg-Vlissingen	En, Ge, Fr
D10	21	F	Middelburg-Vlissingen	En, Fr

2.2 Elicitation material and tasks

All the studies presented in the following chapters made use of the same material, namely (i) a set of videos, (ii) a set of short video-clips, and (iii) a sociolinguistic interview. This elicitation material was assembled as a standard elicitation kit for researchers affiliated to the Traces of Contact project (ERC Project #230310)³⁷. The data were elicited orally and were digitally recorded using Audacity software version 1.3 beta. The video stimuli were played on a laptop in front of the participant, with the instruction to describe ‘what is going on’ in the target language (Ambon Malay for homeland, heritage and first generation speakers; Dutch for Dutch speakers). The participants were always and consistently addressed in the target language and they were not encouraged to code-switch. However, they were told that they could use Dutch words or phrases in case of lexical retrieval problems (this holds for heritage and first generation speakers only). The elicitation sessions took place in familiar environments, such as the home, the office, or the house of

³⁷ The ERC-project Traces of Contact (2009-2013) aimed to establish criteria by which results from language contact studies can be used to strengthen the field of historical linguistics, online URL <http://www.ru.nl/linc/projects/erc-traces-contact/> [Last accessed 25 June 2015].

friends, relatives or neighbors. Occasionally participants were recorded in other places, as in the university where they studied or worked (only two young participants were recorded in my office at Radboud University). Participants performed three tasks, each associated with a specific set of stimuli.

The simultaneous video description task: The set of videos was used for the simultaneous video description task. In this task, the participants were asked to describe 14 videos while watching them (in a fashion similar to the running commentary of a football match). The videos depicted every-day activities, eight with cartoon characters (a mouse and an elephant), and six with human characters (a man and two boys) (for a complete list of the videos see Appendix 1). The eight videos with cartoon characters were extracted from the episodes of a famous German children's series, whose protagonist is a mouse. The six videos with human characters come from two sources: three videos were created by Sotaro Kita and colleagues at the Max Planck Institute for Psycholinguistics, Nijmegen, and three other videos were created by Geoffrey Haig and Stefan Schnell at the University of Kiel. The videos had almost no sound (only a music background). The average length of the videos was about half a minute (the shortest lasted nine seconds; the longest lasted one minute 56 seconds). The videos were presented in three different randomized orders.

The advantage of using the simultaneous video description task is that, as a timed task, it leaves the participants little time to rely on metalinguistic awareness, and therefore it guarantees spontaneous data. A disadvantage is that it forces the participants to speak in a non-natural way. Furthermore, the presence of animation segments created some trouble, especially for elder speakers. Some of them had problems recognizing the animal and spent time trying to guess which animal it was instead of going on with the video description. This problem was enhanced by the fact that the stimuli were presented on a laptop screen, an object elder speakers (especially in Ambon) are not familiar with (a problem discussed in Bower 2008, p. 83). In addition, one first generation speaker found it hard to refer to activities performed by animals, as if they were humans. When she watched the videos where the mouse is cooking or sleeping, she could not help but remark that the mouse is a dirty animal and it is not allowed to be in a kitchen or in a bed (see example 1 below). She could not abstract from the images and talk about the mouse as if it was a human character, as other first generation speakers did (see example 2 below). The video descriptions provided by this informant were, thus, not included in the corpus.

- (1) *Oh tikus mar di tana Belanda tikus nda bole,*
 EXCL mouse but at land Dutch mouse NEG be.allowed

nona lia, tikus nda bole deka makanang
 girl see mouse NEG be.allowed close food

‘Oh a mouse, but in the Netherlands a mouse is not allowed..., look, it is not allowed to let a mouse (stay) close to the food.’

- (2) *Gambar ini ada tikus ada masa di oven,*
 picture D.PROX EXIST mouse EXIST cook at oven

dia mangkali dia biking telur goreng of pannenkoek
 3SG maybe 3SG make egg fry or pancake

‘In this video there is a mouse cooking (something) on the stove, maybe he is frying an egg or a pancake.’

The video-clip retelling task: The set of video-clips was used for the video-clip retelling task (see Appendix 2 for a complete list of the video-clips). Participants were asked to describe a total of 68 short video-clips (the shortest lasted two seconds, the longest 34 seconds). The video-clips were selected from a range of different sources (see Appendix 2). Participants watched two video-clips at a time, and then described what they had just seen. Like the videos, the video-clips also depicted every-day activities, but only with human characters and had no dialogue. The aim of the video-clip retelling task was to elicit a rich corpus for data mining research. Some video-clips targeted a number of grammatical domains and grammatical constructions, such as aspect marking, argument structure, double-object construction, spatial descriptions, etc. The subsets of video-clips used in each chapter are listed in the sections of Appendix 2. The video-clips were arranged in three different randomized orders. An example of a pair of video-clips described by a homeland speaker in Ambon is given in (3) below:

The advantage of using the video-clip retelling task is that it yields naturalistic data: the participants can take as much time as they want to describe what they have just seen. The disadvantage of this task is that it can become rather long and repetitive. Some participants employed 40-60 minutes to complete the task and some of them manifested boredom and frustration (cf. Bower 2008, p. 89). An additional problem was that speakers sometimes forgot the first video-clip in the

pair, so they had to repeat the task. This made the elicitation session even longer and more tiring. Again, elder speaker were among those who had more troubles with this task due to working memory limitations.

(3) *Clip pertama ada nyong satu bajalang turung dari trap-trap*
clip first EXIST boy one walk descend from PL-stair

clip yang ka-dua ada ana muda satu pegang
clip REL ORD-TWO EXIST child young one hold

bunga matahari bagitu lalu dia ciong akang
flower sun like.that then 3SG smell 3SG.N

‘(In) the first clip there is a boy who walks down the stairs, (in) the second clip there is a young boy holding (something) like a sunflower and then he smells it.’

The sociolinguistic interview: The third task was a socio-linguistic interview aimed at collecting biographical information. The participants were asked various questions concerning their life (from birth until adulthood), their origins, their language habits, their ethnic identity, and a self-evaluation of their linguistic skills (see Appendix 3 for the complete list of questions). The length of the interviews varies dramatically (from 3 minutes to 1 hour and 15 minutes), depending on the attitude of the speaker, and on the circumstances. Since the interview was the last task to be performed, in some case (especially in Ambon) it had to be kept short due to the needs and the obligations of the participants. The sociolinguistic interviews of the homeland participants are usually shorter than those of heritage or first generation speakers also because these speakers grew up in a somehow less complex language situation. Most of the homeland participants were raised in Ambon Malay and are now monolingual in this language.

The advantage of collecting data by means of sociolinguistic interviews is that this task provides naturalistic conversational speech (Bower, 2008; Nagy, 2015). Nagy (2015, p. 324) lists a number of features that differentiate experimental tasks (i.e., grammaticality judgments, picture manipulation, and controlled elicitation) from the sociolinguistic interview including “unfamiliar tasks vs. typical every day conversational behavior”, “lab or classroom setting vs. familiar setting”, and “requirement to choose a single answer, structure or form vs. options to avoid a

particular structure by circumlocution or changing the conversational topic”. Although sociolinguistic interviews undoubtedly yield naturalistic speech, the data obtained are not always comparable across speakers because, as Nagy (2015) points out, speakers can choose to avoid structures or forms. The lack of comparability between responses is the reason why the linguistic data elicited via the interviews were not used in the present research. I come back to this issue in the next section.

2.3 The corpus

The data obtained by means of the simultaneous video description task and the video-clip retelling task constitute the language corpus used for the present study (total: 44 hours, 17 minutes, 59 seconds). The language corpus, thus, includes 74 video descriptions (Ambon Malay: 13 hours, 40 minutes, 4 seconds; Dutch: 1 hour, 43 minutes, 6 seconds) and 75 video-clip retelling descriptions (Ambon Malay: 26 hours, 15 minutes, 50 seconds; Dutch: 2 hours, 38 minutes, 18 seconds). The descriptions provided by the simultaneous video description task (Section 2.2) were all considered as valid responses (except the descriptions provided by one first generation speaker, see Section 2.2). The descriptions provided by the video-clip retelling task had to fulfil inclusion criteria. The inclusion criteria are discussed in more detail in the methodology section of each chapter. Overall, the general requirement was that the response contained an adequate description of the action described in the video-clip.

The corpus used for the present research is, thus, formed solely by semi-spontaneous speech. Semi-spontaneous elicited data obtained by means of video stimuli lie somehow in between the two types of data individuated by Nagy (2015), namely highly controlled data obtained through experimental tasks (i.e., grammaticality judgments, controlled elicitation) and spontaneous naturalistic data obtained through a sociolinguistic interview (see Section 2.2). On the one hand, it may be argued that the semi-spontaneous speech is not as natural as conversational speech because the speaker is forced to describe a stimulus, but on the other hand, it is not so highly controlled because the speaker can freely choose how to describe the stimulus (cf. Bowern, 2008, p. 82). The choice of semi-spontaneous data was motivated by the need to find a compromise between ecological validity (the degree to which the data represent real-life language) and outcome effectiveness (getting from the data what I need). Since the focus of this dissertation is to investigate

possible changes in heritage Ambon Malay and to test specific hypotheses related to the grammatical features that are more prone to change, I need to elicit data that allow me to test these hypotheses. As correctly pointed out by Bower (2008, p. 73), “some aspects of a language are only discoverable through elicitation – they will appear in texts so seldom that it will be almost impossible to get enough information about them”.

The metalinguistic data extrapolated from the sociolinguistic interviews were coded and used for the interpretation of the linguistic data obtained from the other two tasks (see Chapter 7). As mentioned in the previous section, the linguistic data elicited via the interviews were not used due to the lack of comparability between responses. Furthermore, I got the impression that the speakers tended to use a more formal style during the interviews than during the other two tasks (e.g., using the Indonesian negation *tidak* instead of the typical Ambon Malay *seng*). This is in line with the observation made by Sneddon (2003b) in Indonesia, where he investigated the use of formal and informal varieties of Indonesian. Sneddon (2003b, p. 535) reports that:

The interviews, where one person is asked questions and encouraged to talk about himself or herself, tend to exhibit more characteristics of formal language, being situations in which a certain amount of self-conscious linguistic behavior might be expected.

Further investigation is needed to establish with certainty whether the same is true in other settings, such as in Ambon, or in the Central Moluccas.

2.4 Analytic procedure

The data collected through the three tasks described in Section 2.2 were transcribed by means of the software ELAN version 4.01. All the transcriptions of the Ambon Malay data were done by myself; 19 transcriptions were double checked with the transcriptions done by Rose Lekawael (see Section 2.1). The transcriptions of the Dutch data were done by Rowan Soolsma (see Soolsma, 2013). Generally speaking, the data were annotated following broad transcriptions guidelines (phonemic rather than phonetic transcription) in order to facilitate word search in the files (see Bower, 2008).

The data were coded in ELAN version 4.01 and in Excel for various type of features including total number of (animate and inanimate) nouns (in Chapter 3), total number of predicates (in Chapter 4), rate of Double Object and Prepositional Object constructions (in Chapter 5), rate of Serial Verb constructions and other types of resultative constructions (in Chapter 6), etc. When necessary new categories or labels were created in order to facilitate the comparison between the language varieties. This is the case of the ‘two predicate construction’, a term coined to refer a particular type of construction in the Ambon Malay varieties (see Chapter 5), and the ‘adjectival phrase construction’, another term coined for the sake of cross-linguistic comparability between Ambon Malay and Dutch (see Chapter 6).

The analysis of the various linguistic features carried out in this dissertation is mostly quantitative in nature. The individual rates are systematically compared across the group speaking the contact variety (heritage Ambon Malay) and the group speaking the non-contact variety (homeland Ambon Malay), following the guidelines outlined in section 1.1. In addition, the rates of the first generation group (the first generation language is the parental language for heritage speakers) and of the Dutch group (Dutch is the possible source language) are also reported for qualitative comparison. Data are analyzed statistically using the statistical software SPSS version 22 and the software R version 3.2.0. The statistical tests employed in this dissertation are the independent samples *t*-test, regression analysis (general linear model and generalized linear mixed effects model) and hierarchical cluster analysis³⁸.

I performed the independent samples *t*-test using SPSS in Chapter 3, Chapter 4, Chapter 5 and Chapter 6. The independent samples *t*-test is used to compare the means of two populations in different experimental conditions with different participants in each group (Field, 2005, p. 296). In the present study, the two populations are represented by the homeland group and the heritage group, which differ with respect to one variable, namely the language variety they speak: homeland speakers speak a variety without Dutch influence (posterior to 1951), whereas heritage speakers speak a variety influenced by Dutch. If the means of the two groups are found to diverge significantly, this means that the language they speak has a real effect on the linguistic feature under investigation. In this dissertation, I used the two-tailed independent samples *t*-test because I wanted to be neutral with respect to the direction of change and because the two-tailed *t*-test is

³⁸ The statistical tests were performed with the help of my supervisor Dr. Harald Hammarström and my colleagues, Linda van Meel and Dr. Gerrit Jan Kootstra, to whom I am deeply grateful.

more conservative than the one-tailed *t*-test (Field, 2005, p. 29). The *t*-test presupposes that the variances in the two groups are roughly equal. The variance is a measure of how much the observations vary around the mean. A naturalistic experiment may or may not have the property that the variances are roughly the same. To test this, Levene's test is used. If Levene's test is significant ($p < .05$), the variances are not equal (*equal variances not assumed*); if Levene's test is not significant ($p > .05$), the variances are equal (*equal variances assumed*). According to the results of Levene's test, I report the *t*-statistics of the corresponding row in the table (in SPSS). When reporting the results of the *t*-test, I also report the effect size (*r*). Effect size (*r*) is a useful value because it indicates the magnitude of the observed effect. The following conventions are generally used to interpret the results (Field, 2005, p. 32):

- $r = .10$ (small effect) the effect explains 1% of the total variance
- $r = .30$ (medium effect) the effect explains 9 % of the total variance
- $r = .50$ (large effect) the effect explains 25 % of the total variance

I performed the generalized linear mixed effects model using R (glmer in the lme4 package) in Chapter 3 and Chapter 4. The generalized linear mixed effects model is used to assess the effect of a number of factors on a given linguistic outcome or response (Baayen, 2008, p. 263). The model incorporates both fixed effects factors and random effects factors. Fixed effect factors have a fixed set of possible levels, while random effect factors have an infinite number of levels as they are randomly selected from a population. For instance, in the generalized linear mixed effects model used in this dissertation, the fixed effect factor is the group to which a speaker belongs (i.e., homeland vs. heritage). If we repeat an experiment and select new participants, they will always belong to either of these two groups. The random effect factor is the speaker. If we repeat an experiment and select new participants, we would introduce new levels that were not present in the previous experiment, and therefore we include new 'random' variation.

I performed the univariate general linear model (glm) using SPSS in Chapter 7. The glm aims to predict a dependent variable (response variable) from one or more independent variables (predictor variables), the latter can be categorical or continuous (Field, 2005, p. 144). In the glm used in Chapter 7, the response variable is a continuous linguistic variable (i.e., rate of double object constructions, rate of pre-nominal *itu* 'D.DIST', etc.), while the predictor variables are always categorical:

age of onset of bilingualism (sequential vs. simultaneous), the place where the speaker lives (Moluccan ward vs. outside a Moluccan ward), and generation (2nd, 2.5 or 3rd generation). Additionally, the glm produces interactions between predictors, because it assumes that the different levels of a predictor (e.g., sequential vs. simultaneous bilingualism) can combine with the levels of another predictor (e.g., living in a Moluccan ward vs. outside a Moluccan ward) and have different effects on the response variable. For instance, we see in Chapter 7 that the interaction between two independent variables (*age of onset of bilingualism* and *place where the speaker lives*) can explain the use of double object constructions (the response variable). When reporting the results of the glm, I report the partial eta squared (PES) and the *p*-value of each predictors, and the adjusted R square (R) of the overall model. PES measures the proportion of variance explained by the individual variable independent of all other variables; thus a large PES means a large effect size. The *p*-value tells us whether the individual variable better predicts the outcome than the null hypothesis. If the *p*-value is $< .05$, the variable has a significant effect on the outcome. R is the coefficient of determination and quantifies the proportion of variance in the data that is explained by the model (Baayen, 2008, p. 96).

Finally, I performed hierarchical cluster analysis using SPSS in Chapter 7. Hierarchical cluster analysis groups objects, in this case linguistic variables, into clusters (Ward's method) based on pairwise distances (Euclidean distance) (see Ward, 1963). The proximity (similarity) or distance (dissimilarity) between the object is measured as distance matrix. The graphical representation of hierarchical clustering is the dendrogram, or clustering tree, where the linguistic features are grouped together in a hierarchical fashion from the closest (most similar) to the furthest apart (most different). We see in Section 7.2.1 that two macro-groups of features can be identified: the 'Dutch-like' or 'innovative' features and the 'Malay-like' or 'conservative' features.

CHAPTER 3

Nominal modification in heritage Ambon Malay

3.1 Introduction

Word order change is one of the most prominent issues in language contact research. A number of studies have shown that the linear order of words is extremely vulnerable to change in short-term contact situations, namely when two languages are in contact for less than a hundred years (Hartsuiker & Westenberg, 2000; Silva-Corvalán, 1994, 2008; Albirini, Benmamoun, & Saadah, 2011; Onar Valk, 2015), as well as in long-term contact situations, when languages are in contact for centuries (Thomason & Kaufman, 1988; Ross, 2007; Heine, 2008; Matras, 2009). Although there is a difference in degree and scale – in short-term contact, the change is ongoing and it is actuated by a group of speakers, while in long-term situations the change is complete and is actuated by the whole speech community – the mechanisms by which word order changes emerge and develop are the same in both situations (Backus et al., 2011).

Researchers studying bilingual populations (here taken to mean L2 learners and bilingual heritage speakers) have formulated various hypotheses to explain why some structural features, including word order patterns, are more prone to cross-linguistic influence than others. The Alternation Hypothesis and the Vulnerability Hypothesis have pointed to the role of variability (alternation of structures) as the locus for cross-linguistic influence (see Section 1.4.1). Although many researchers do not explicitly adhere to these hypotheses in their interpretation of results, there is a general consensus that cross-linguistic influence from one language to the other is likely to occur when there is perceivable overlap of structures between the two languages, i.e. the L1 has option A, while the L2 has options A and B, and reversely it is less likely when there is not much perceivable overlap, that is when the L1 has option A, and the L2 only option B (Backus, 2004; Silva-Corvalán, 2008; Montrul & Ionin, 2010; Muysken, 2013). An example of word order alternation as the locus of convergence comes from Spanish-English bilingual heritage speakers (Silva-Corvalán, 1994). Spanish allows an alternation between (S)ubject-(V)erb or VS order with unaccusative verbs. The SV order is shared with English, the dominant

language of these bilinguals, while the VS order is not. Silva-Corvalán (1994) observes that Spanish-English bilinguals use the SV order with a higher frequency than their monolingual peers and concludes that this shift in preference is the result of transfer from English. Similar results are presented in Albirini et al. (2011), who report data on Egyptian-English bilingual heritage speakers. The authors argue that, due to transfer from English, bilingual speakers in the U.S.A. prefer the SVO order, although Egyptian Arabic allows also for VSO as an alternative option. Another example comes from Turkish in the Netherlands. Onar Valk (2015) shows that Dutch-Turkish bilingual heritage speakers more often select the verb-medial order, which is present in Dutch and is also possible in Turkish (although pragmatically marked).³⁹

There is another factor, however, that can have a potential effect on word order, namely the meaning of the word(s). Since words are not merely forms, but they convey also a meaning, a change in word order may entail also a change in meaning (Backus et al., 2011). For instance, in Latin the demonstrative *ille* could precede or follow the noun, but when the demonstrative became grammaticalized to the definite article in Romance languages, its order became strictly pre-nominal (Lehmann, 1992, p. 403). Another example is that of nineteenth-century Baba Malay, a Malay variety spoken in Singapore, in which *ini* ‘D.PROX’ and *itu* ‘D.DIST’ could be used as demonstratives and definite articles: when they preceded the noun, they functioned as demonstratives; when they followed the noun, they functioned as definite articles. Thurgood (2001, p. 743) explains that the Baba Malay system of demonstratives and definite articles originated when Hokkien speakers shifted to Malay and were forced to learn a new word order. In Hokkien, the position of the demonstrative is pre-nominal. In Malay, the position is generally post-nominal. However, the pre-nominal position is also allowed in contact Malay varieties (Paauw, 2008). Bilingual speakers found “a compromise between the distributional patterns found in Malay and in Hokkien” (Thurgood, 2001, p. 486), and as a result

³⁹ The higher incidence of SV order in these three heritage languages could also be due to simplification, rather than to transfer from English or Dutch (see also Section 1.3.1 and Section 1.3.5). Two simplification processes may be responsible for the higher incidence of SV: (i) SV is claimed to be the least complex or the most unmarked order (McWhorther, 2001a; Kusters, 2003); (ii) the language simplifies its system by changing from flexible-word order to rigid word order (Polinsky & Kagan, 2007, p. 382; Albirini et al., 2011, p. 298; Onar Valk, 2015, p. 245). The most likely scenario is that simplification and English/Dutch influence reinforce each other in promoting heritage language change.

of this compromise, the prenominal position came to be associated with the demonstrative use to match the Hokkien order, and the post-nominal position came to be associated with the article use.

To summarize, partial overlap of structures between two languages and grammatical re-interpretation can have an effect on the linear arrangement of words. These two factors can either act separately, or they can have an incremental effect.

In this chapter, I follow this line of investigation and focus on word order changes in heritage Ambon Malay. In order to test the role of variability and grammatical re-interpretation on word order, I selected a domain where both variability and lack of variability are found, and where items are prone to grammatical re-interpretation. This testing ground is the domain of nominal modification. Indeed, the order of nominal modifiers in Ambon Malay partially overlaps with the order of nominal modifiers in Dutch. Furthermore, while Dutch has a system of fully-fledged articles, Ambon Malay has only weakly grammaticalized definite and indefinite articles. Since (in)definiteness is obligatorily marked in Dutch, we may expect heritage speakers to replicate this grammatical category by using Ambon Malay morphemes (see the Functional Convergence Hypothesis in Section 1.4.1). Heine and Kuteva (2008, p. 79) point out that “in order to replicate an indefinite article of the model language, most likely they [bilingual speakers] will select their numeral ‘one’, or a demonstrative attribute to replicate a definite article”. I therefore hypothesize that heritage speakers of Ambon Malay will select the demonstratives and the numeral ‘one’ to replicate Dutch articles.

The central questions addressed in this chapter are: Does the partial (surface) overlap between Ambon Malay and Dutch trigger word order changes in the heritage variety? Does grammatical re-interpretation play a role in the word order change of the demonstratives and the numeral ‘one’? The findings of this chapter show that heritage speakers prefer the word order patterns that have a parallel in the dominant language, Dutch. A comparison of the overall frequency and distribution of demonstratives and the numeral ‘one’ in homeland and heritage speakers suggests that these morphemes are not reinterpreted as articles in the heritage variety. Hence, heritage speakers and homeland speakers use them with a different linear order, but apparently with the same meaning.

This chapter is structured as follows. Section 3.2 describes the linear order and the basic functions of nominal modifiers in the languages of bilingual heritage speakers: Ambon Malay (section 3.2.1) and Dutch (section 3.2.2). Section 3.3

illustrates the design of the present study. The results are presented and discussed in section 3.4. Section 3.5 summarizes the conclusions.

3.2 Word order of nominal modifiers in Ambon Malay and Dutch

This section presents a brief descriptive overview of nominal modifiers in Ambon Malay (Section 3.2.1) and Dutch (Section 3.2.2), the languages that are combined in the same, bilingual heritage speaker of Ambon Malay. Here I adopt a general definition of modifier and consider as modifiers all constituents which can modify a noun, including demonstratives and determiners (Payne, 1997; van Minde, 1997; Loos et al., 2004). The nominal modifiers under investigation here are demonstratives, numerals, adjectives and (in)definite markers. For possessive noun phrases in heritage Ambon Malay the reader is referred to Huwaë (1992, pp. 31-33), Muysken (2005, pp. 14-15), and Aalberse and Moro (2014, pp. 154-157). Ambon Malay and Dutch display some similarities, as well as interesting differences in the linear arrangement of these nominal modifiers; these are summarized in Section 3.2.3. The examples presented in this section and elsewhere in this chapter are all from the dataset collected for this dissertation (see Section 2.2 and Section 2.3).

3.2.1 Word order of nominal modifiers in Ambon Malay

In Ambon Malay the order of the demonstratives *ini/in/ni* ‘D.PROX’ and *itu/it/tu* ‘D.DIST’ is variable: it can be either DEMONSTRATIVE-NOUN, as illustrated in (1), or NOUN-DEMONSTRATIVE, as illustrated in (2). The sentences in (1) and (2) are responses provided by two speakers in Ambon to the same video-clip showing a boy holding a baseball bat who hits a ball that someone has thrown to him. Van Minde (1997) states that it is still unclear what the semantic contrast is between the pre-nominal and the post-nominal order; statistically speaking, however, the post-nominal order is far more frequent.⁴⁰ Originally, however, Ambon Malay had a higher incidence of DEMONSTRATIVE-NOUN (1), as this was a typical feature of eastern Malay varieties (Paauw, 2008, p. 299). It is likely that Standard Indonesian,

⁴⁰ Unfortunately, van Minde’s (1997) grammar does not contain information about the type of data on which quantitative statements like these are based.

which has NOUN-DEMONSTRATIVE order, has further entrenched the post-nominal order over the years (2).

- (1) *Ada orang buang bola kasti itu par dia*
 EXIST person throw ball k.o.game D.DIST to 3SG
la dia pukol itu bola
 then 3SG hit D.DIST ball
 ‘There is a person who throws a tennis ball to him (and) then he hits the/that ball.’

- (2) *Ada laki-laki satu dia pegang kayo lalu bola*
 EXIST male one 3SG hold stick then ball
datang par dia lalu dia pukol bola itu
 come to 3SG then 3SG hit ball D.DIST
 ‘There is a man, he holds a bat, then a ball comes to him (is thrown to him), (and) then he hits the/that ball.’

Ambon Malay demonstratives do not indicate only spatial distance, but also speech distance (van Minde, 1997, p. 147; Cleary-Kemp, 2007; van Engelenhoven, 2008). Van Minde (1997, p. 147) states that “both *ini/in/ni* and *itu/it/tu* share the feature definiteness: they indicate that the speaker assumes that the referent of the head noun is identifiable by the hearer.”

Definiteness and indefiniteness refer to the knowledge/mind state of the speaker and/or the hearer in the discourse. A noun is marked as definite if both the speaker and the hearer can identify the referent of the noun. On the other hand, a noun is marked as indefinite if the speaker and the hearer cannot identify a unique referent for the noun (Payne, 1997, p. 233). In some languages, such as Dutch, articles are used to express the feature of (in)definiteness. Ambon Malay does not have articles, but it can use demonstratives to perform this function. It is *itu* ‘D.DIST’, in particular, that has article-like features, while *ini* ‘D.PROX’ is used with an article-like function only when the speaker wants to emphasize the role or importance of a referent. Since referents of primary importance are usually humans or animals, *ini* ‘D.PROX’ occurs mostly with animate nouns; cf. Thurgood, 2001, p. 484. According to Cleary-Kemp (2007), Ambon Malay exhibits a pattern similar to that of Baba Malay, with pre-nominal *ini/itu* functioning as demonstratives, and post-nominal *ini/itu*

functioning as articles. The author, however, does not support her statement with quantitative data, as done by Thurgood (2001) for Baba Malay. In the dataset used for the present study both pre-nominal and post-nominal demonstratives seem to have article-like features, as shown by example (1) and (2) above and by example (3) and (4) here below. In (3) and (4), when a new participant, *kue* ‘cake’ or *apel* ‘apple’, is first introduced, it is left unmarked. When the participant is mentioned for the second time, it is referred to as *kue itu* ‘the cake’ or *itu apel* ‘the apple’, with *itu* functioning as definite marker.

- (3) *Ada se-ekor tikus yang badiri di muka kue*
 EXIST one-CLF mouse REL stand at face cake

trus dia ciom-ciom kue tu
 next 3SG ITER-smell cake D.DIST
 ‘There is a mouse standing in front of a cake, then he repeatedly smells the cake.’

- (4) *Ada satu nyong yang... dia cuci apel,*
 EXIST one young.man REL 3SG wash apple

trus dia makang itu apel
 next 3SG eat D.DIST apple
 ‘There is a young man who..., he washes an apple and then he eats the apple.’

The order of the *numerals* is also variable in Ambon Malay: it can be NUMERAL-NOUN, as shown in (5) or NOUN-NUMERAL, as shown in (6). The sentences in (5) and (6) are responses provided by two speakers in Ambon, Central Moluccas, to the same video-clip showing a grass field with three balls. In Ambon Malay, preposed quantifiers seem to be used to stress the individuality of the referent(s), while postposed quantifiers stress collectivity (see Section 1.6.2.6). Although the postposed numerals seem to be preferred in Ambon Malay (6), the preposed numerals (5) are becoming increasingly frequent due to the influence of Standard Indonesian, in which the order is NUMERAL-NOUN (Paauw, 2008, p. 411).

- (5) *Lapangang kosong, ada tiga bola*
 field empty EXIST three ball
 ‘(I see) an empty field, there are three balls.’

- (6) *Ada lapangang kosong, di sini ada bola tiga*
 EXIST field empty at LOC.PROX EXIST ball three
 ‘There is an empty field, here there are three balls.’

The numeral *satu* ‘one’ may occasionally function as the indefinite article when it occurs with first-mention nouns. This is especially true with first-mention nouns referring to animate referents (cf. Thurgood, 2001, p. 481), but its use is optional and largely dependent on the speaker. We have seen in examples (3) and (4) that bare nouns (*kue* ‘cake’ or *apel* ‘apple’) can also be interpreted as indefinite. An example of pre-nominal and post-nominal *satu* used as indefinite article is given in (7) and (8). In (7) and (8), two new participants, *nyong* ‘young man’ and *nona* ‘girl’ are first introduced as *satu nyong/satu nona* or *nyong satu/nona satu* ‘a young man/a young girl’ with the numeral *satu* functioning like an indefinite article.

- (7) *Satu nyong deng satu nona, dong dua badiri*
 one young.man with one girl 3PL two stand
 ‘(In this video I see) a young man and a young woman, they are standing.’

- (8) *Nyong satu deng nona satu dong dua ada badiri*
 young.man one with girl one 3PL two EXIST stand
 ‘(In this video I see) a young man and a young woman, they are standing.’

Ambon Malay lacks a morpho-syntactically definable class of *adjectives*. Words referring to qualities are stative intransitive verbs which can occur in an NP to modify a noun (van Minde, 1997, p. 66). For the sake of comparability with Dutch, stative intransitive verbs are referred to as ‘adjectives’. I have adopted this choice also in Chapter 6. Adjectives in Ambon Malay always follow the noun, as illustrated in (9).

- (9) *Dia ambel kayo panjang/ *panjang kayo*
 3SG take stick long long stick
 ‘He takes a long wooden stick.’

Finally, Ambon Malay can encode definiteness by means of the definite marker = *nya* ‘DEF’, an enclitic that follows the noun or the nominal modifiers. In the homeland variety, = *nya* seems to be a very marginal form, and van Minde (1997, p. 363) only mentions it in a footnote. In Indonesian (and probably also in Ambon Malay) = *nya* has an associative reference meaning which can be inferred from the context (Englebretson, 2003; Ewing, 2005; Arka, 2011). It encodes possession when the type of association is one of ownership or part-whole relationship. It encodes definiteness when the referent of the NP to which it is cliticized is identifiable through association with a frame or scheme shared by the speaker and the hearer. In (10), the referent of *nasi* ‘rice’ has not been previously mentioned but is marked as definite because in Indonesia the act of eating is implicitly associated with consuming rice.

- (10) *Kalau mau makan, nasi = nya di lemari*
 if want eat rice = DEF at pantry
 ‘If you want to eat, the rice is in the pantry.’ (Sneddon, 1996, p. 151)

The use of = *nya* as a definite marker is optional among speakers of Ambon Malay. This form is very marginal in the homeland variety, while it is rather common in the heritage variety (Tahitu, 1989; Huwaë, 1992; Aalberse & Moro, 2014). In the heritage Ambon Malay dataset used for the present study, = *nya* can also be used anaphorically. In (11), = *nya* indicates that the referent of *kore api* ‘match’ is identifiable to the speaker and the hearer because it has already been introduced.

- (11) *Ini kore api, kore api = nya sekarang ta-bakar*
 D.PROX match fire match fire = DEF now ACL-burn
 ‘This is a match/Here there is a match, now the match is burned.’

3.2.2 Word order of nominal modifiers in Dutch

In Dutch, all nominal modifiers precede the noun. With respect to demonstratives, Dutch has a fixed DEMONSTRATIVE-NOUN order (Haeseryn, Romijn, Geerts, de Rooij, & van den Toorn, 1997, 5.6), as illustrated in (12) and in (13).⁴¹

⁴¹ *Deze* ‘D.PROX’ and *die* ‘D.DIST’ are used before common gender nouns, *dit* ‘D.PROX’ and *dat* ‘D.DIST’ are used before neuter gender nouns (Haeseryn et al., 1997, 4.2; Donaldson, 2008, p. 32).

Demonstratives usually indicate spatial distance relative to the speaker.⁴² However, they can also signal a relative degree of focus placed on a referent by the speaker, regardless of spatial distance (Kirsner, 1979). For instance, in (12), the proximal demonstrative *deze* ‘D.PROX’ draws attention on the noun *lap* ‘cloth’ which undergoes a change of state as a result of the action. In (13), *die* ‘D.DIST’ is ambiguous between indicating spatial distance, and signaling a low degree of attention placed on the noun *stoel* ‘chair’.

- (12) *Een jonge vrouw heef-t een groene lap*
 ART.INDF young woman have-3SG ART.INDF green cloth

in haar hand-en en scheur-t deze lap doormidden
 in 3SG.POSS.F hand-PL and tear-3SG D.PROX cloth in.half
 ‘A young woman has a green cloth in her hand and tears this (piece of) cloth in half.’

- (13) *Een man staa-t in een ruimte,*
 ART.INDF man stand-3SG in ART.INDF room

achter hem staa-t een stoel
 behind 3SG.M.ACC stand-3SG ART.INDF chair

en hij gaa-t op die stoel zitt-en
 and 3SG go-3SG on D.DIST chair sit-INF
 ‘A man stands in a room, there is a chair behind him, and he goes to sit on that chair.’

Numerals always precede the noun (Haeseryn et al., 1997, 7.1), as illustrated in (14).

- (14) *Ik zie drie ball-en op een grasveld ligg-en*
 1SG see.1SG three ball-PL on ART.INDF grass.field lie-INF
 ‘I see three balls lying in a grass field.’

⁴² In some Dutch ethnolects, demonstratives, especially *die* ‘D.DIST’, are sometimes used instead of the definite articles (Muysken, 2010a, p. 21).

Adjectives also always precede the noun (Haeseryn et al., 1997, 6.1), as shown in (15).

- (15) *Een grote taart staa-t op een serveerschaal*
 ART.INDF big cake stand-3SG on ART.INDF serving.dish
 ‘There is a big cake on a serving dish.’

The definite article *de/het*⁴³ always precedes the noun and marks the definiteness of a noun phrase (Haeseryn et al., 1997, 4.2). In Dutch, the use of the definite article is obligatory with already mentioned nouns (anaphoric use). In (16), the second instance of the noun *taart* ‘cake’ is preceded by the definite article *de* because the referent has already been introduced, thus it is identifiable to both the speaker and the hearer.

- (16) *Er kom-t een muis aangelop-en naar een*
 there come-3SG ART.INDF mouse walk-INF to ART.INDF

hele grote taart en ruik-t met zijn neus
 very big cake and smell-3SG with 3SG.POSS.M nose

aan de taart
 on ART.DEF cake
 ‘A mouse comes walking toward a very big cake and smells the cake with his nose.’

Indefiniteness is marked by the indefinite article *een* which appears before the noun, as shown in (17). When a new participant (e.g., *a mouse*) or a new entity (e.g., *a cake*) is introduced for the first time in the discourse, it is usually marked by *een*.

The article *een* ‘ART.INDF’ developed historically from the numeral *één* ‘one’ (Philippa, Debrabandere, Quak, Schoonheim & van der Sijs, 2011.). The two forms, however, are no longer homophonous,⁴⁴ although they are homographous to some extent. The use of the indefinite article *een* with first-mention nouns is obligatory in Dutch (with

⁴³ *De* ‘ART.DEF’ is used before singular common gender nouns and *het* ‘ART.DEF’ is used before neuter nouns (Haeseryn et al., 1997, 4.2; Donaldson, 2008, p. 27).

⁴⁴ The acute accent distinguishes the numeral *één* ‘one’ from the indefinite article *een* ‘a’ (Donaldson, 2008, p. 12).

the exception of nouns indicating professions, nationalities, and after certain prepositions; see Donaldson, 2008, p. 25).

- (17) *Er kom-t een muis aangelop-en naar een*
 there come-3SG ART.INDF mouse walk-INF to ART.INDF
- hele grote taart en ruik-t met zijn neus*
 very big cake and smell-3SG with 3SG.POSS.M nose
- aan de taart*
 on ART.DEF cake
- ‘A mouse comes walking toward a very big cake and smells the cake with his nose.’

3.2.3 Interim summary

We have seen that in the domain of nominal modification, there are similarities as well as differences between Ambon Malay and Dutch. These are summarized in Table 3.1.

Table 3.1: Summary of the word order patterns of nominal modifiers in Ambon Malay and Dutch.

	AMBON MALAY	DUTCH
Demonstrative	DEMONSTRATIVE-NOUN	DEMONSTRATIVE-NOUN
	NOUN-DEMONSTRATIVE	-
Numeral	NUMERAL-NOUN	NUMERAL-NOUN
	NOUN-NUMERAL	-
Adjective	-	ADJECTIVE-NOUN
	NOUN-ADJECTIVE	-
(In)Definite Marker	-	(IN)DEFINITE-NOUN
	NOUN-DEFINITE	-

With respect to linear order, both languages allow the pre-nominal order for demonstratives and numerals, although Ambon Malay seems to prefer the post-nominal order. The languages differ insofar as Ambon Malay possesses strictly post-

nominal adjectives and definite marker, while in Dutch these classes are strictly pre-nominal. In sum, for the demonstratives and the numerals, Ambon Malay and Dutch share the pre-nominal option, whereas for the adjectives and the (in)definite markers, the two languages do not share any option.

With respect to the function of the modifiers, the two languages differ: in Ambon Malay (in)definite marking is optional and largely dependent on speakers' preferences and on animacy (i.e. nouns referring to animate referents are more likely to be marked). Definiteness can be expressed by the demonstrative *itu* 'D.DIST' and by *=nya* 'DEF' (and marginally by *ini* 'D.PROX'). *Itu* has only an anaphoric function, while *=nya* can have both an anaphoric and a non-anaphoric function. Indefiniteness can be expressed by the numeral *satu* 'one'. In Dutch definite and indefinite marking is obligatory and is expressed by the articles *de/het* 'ART.DEF' and *een* 'ART.INDF', respectively.

3.3 The study

This section discusses the objective (3.3.1), the research questions (3.3.2) and the methodology (3.3.3) of the present study.

3.3.1 Objective

The purpose of the present study is to examine nominal modifiers in the Ambon Malay speech of Dutch-Ambon Malay heritage speakers to test whether the partial overlap observable in the word order patterns of the two languages could trigger reordering, and whether grammatical re-analysis could also trigger reordering (specifically in the case of the demonstratives and of the numeral *satu* 'one').

3.3.2 Research questions and hypotheses

Three main questions are addressed in this study. The first question asks whether the partial overlap between Ambon Malay and Dutch is likely to trigger re-ordering of nominal modifiers (see the Alternation Hypothesis in Section 1.4.1). With respect to the linear order of the demonstratives and the numerals, Ambon Malay and Dutch share the pre-nominal option. Following the Alternation Hypothesis, the general prediction is that surface overlap in word order between Ambon Malay and Dutch

creates the conditions for cross-linguistic influence, which will manifest itself as an increase in the frequency of the pre-nominal order in heritage speakers.

The second question, related to the first, asks whether the lack of variability (that is, no perceivable overlap or complete identity between two languages) is likely to block cross-linguistic influence (Silva-Corvalán, 1994, p. 134). With respect to the linear order of the adjectives and the definite marker, Ambon Malay and Dutch do not overlap: in Ambon Malay the position is strictly post-nominal, while in Dutch is pre-nominal. The prediction is that the lack of variability in Ambon Malay and the consequent lack of overlap with Dutch will block cross-linguistic influence. Furthermore, in the case of the Ambon Malay definite marker =*nya* 'DEF', re-ordering is further blocked by the fact that =*nya* is an enclitic; bound morphemes, in fact, are relatively more resistant to cross-linguistic influence than free forms (Matras, 2009).

The third question, which concerns the demonstratives *ini* 'D.PROX' and *itu* 'D.DIST' and the numeral *satu* 'one', asks whether the change in form correlates to a change in function. In other words, if the demonstratives and the numeral 'one' occur more often in pre-nominal position, is this a pure word-order change, or is this change related to the grammaticalization of these morphemes on the model of Dutch articles? In Dutch, nouns are obligatorily marked for definiteness. Following the Functional Convergence Hypothesis (see Section 1.4.1), we may expect bilingual heritage speakers to re-map the features of definiteness and indefiniteness onto Ambon Malay morphemes. We know that in other Malay varieties, namely Baba Malay (Thurgood, 2001) and Sri Lanka Malay (Nordhoff, 2009), the demonstratives and the numeral 'one' have grammaticalized to definite and indefinite articles respectively. In Baba Malay, *ini* 'D.PROX' and *itu* 'D.DIST' have developed into articles and also become increasingly frequent (Thurgood, 2001, p. 481).⁴⁵ In Sri Lanka Malay the indefinite article *hattu* (<**satu*) has become obligatory under the influence of Sinhala (Nordhoff, 2009, p. 59). Since current definitions of grammaticalization include increasing obligatoriness of a morpheme and functional extension (Traugott & Heine, 1991; Heine & Kuteva, 2005; see also Section 1.3.1.3), one may expect the morphemes expressing (in)definiteness (*ini* 'D.PROX', *itu* 'D.DIST', =*nya* 'DEF', and *satu* 'one') to become more obligatory (thus more frequent) in the heritage language, and to extend their range of use to include a higher number of inanimate referents (see Section 3.2.1).

⁴⁵ Thurgood (2001, p. 481) shows that with its 38% of NPs marked by an article, Baba Malay is similar to languages like Italian (39%) and French (39%).

3.3.3 Participants, task and responses

In order to answer the questions posed in the previous section, we⁴⁶ tested the use of nominal modifiers in the three groups of Ambon Malay speakers: one test group of heritage speakers and two control groups. The test group consists of 32 heritage speakers. The first control group is formed by 27 homeland speakers. The second control group is formed by six first generation speakers of Ambon Malay in the Netherlands. Dutch data are not included in the analysis because the word order of nominal modifiers is fixed in Dutch (see Section 2.1 for more information about the participants).

The dataset for the present study comprises the video descriptions elicited via the simultaneous video description task (see Section 2.2, see also Appendix 1). This task was selected because the videos display independent stories where some characters and items are presented only once (and thus classify as first-mention nouns), while some others re-occur later on in the story (and thus classify as already mentioned nouns). In the video descriptions, thus, demonstratives are mostly used to track referents in the story (i.e., anaphorically). Furthermore, by taking the videos as our elicitation material, we were able to control for specificity.⁴⁷ When performing a video description task, participants mainly refer to the entities in the videos, and thus to entities that have specific properties (cf. Zdorenko & Paradis, 2008, p. 238).

The coding proceeded as follows. All the noun phrases were coded for the order of the modifier(s), i.e. pre-nominal or post-nominal modifier. We included all adjectives, demonstratives, numerals and instances of definite =*nya* found in the video descriptions. We excluded combinations of demonstratives, such as *ini* + NOUN + *ni*, or *itu* + NOUN + *tu* because they do not allow for an investigation on

⁴⁶ Betty Litamahuputty took part in the coding of the data. Her participation in the coding process and the many discussions we had have been of great help for the development of this study.

⁴⁷ Definite nouns and specific nouns carry distinct semantic features. Specific nouns are nouns that refer to an entity possessing some noteworthy properties, this entity can be known to the speaker only (and, thus, be indefinite), or it can be identifiable to the speaker and the hearer (and, thus, be definite). Since, definiteness and specificity play a significant role in determining the use of articles (Ionin, Ko, & Wexler, 2004), excluding non-specific nouns allow us for an observation of the effects of definiteness only.

word order.⁴⁸ The quantifier *samua* ‘all’ was also excluded because this form belongs to the class of quantifiers in Ambon Malay (van Minde, 1997, p. 73) and to the class of adjectives in Dutch. We do not know whether bilingual speakers maintain the homeland conceptualization of *samua* as a numeral, or whether they equate it to Dutch *alle* ‘all’ and treat it as an adjective. The inclusion of *samua* in either the analysis of the adjectives or the numerals may have skewed the results and for this reason we decided to exclude it. Finally, we excluded also instances of possessive =*nya*.⁴⁹

Nouns were also coded as first-mention or already mentioned depending on whether the referent was mentioned for the first time, or whether it had been mentioned earlier. Finally, we considered also animacy: noun phrases referring to humans and animals were coded as animate, noun phrases referring to objects and concepts were coded as inanimate.

Not all noun phrases were included in the analysis, as displayed in Table 3.2.

Table 3.2: Summary of valid and excluded responses in the three Ambon Malay groups.

GROUP	n	RESPONSES	
Heritage Ambon Malay speakers	32	Valid	909
		Excluded	254
Homeland Ambon Malay speakers	27	Valid	1137
		Excluded	225
First generation Ambon Malay speakers	6	Valid	146
		Excluded	30

We excluded non-specific noun phrases, such as ‘he is playing alone, he doesn’t have a *friend* to play with’ (non-specific in italics) and nouns with indeterminate

⁴⁸ These combination occurred 12 times in the homeland group (ten times *ini* + NOUN + *ni* and two times *itu* + NOUN + *tu*), and 13 times in the heritage group (seven times *ini* + NOUN + *ni* and six times *itu* + NOUN + *tu*) and one time in the first generation group (one time *ini* + NOUN + *ni*).

⁴⁹ In some cases, =*nya* clearly functions as a third person singular possessive marker, such as with body parts (*kepala=nya* ‘his/her head’). In other cases, it is difficult to establish with certainty whether =*nya* functions as possessive or as definite marker’, as with *bapa=nya*, which can mean either ‘the man’ or ‘his/her father’. These unclear cases were also excluded from the analysis.

information status, namely nouns that could not be classified as first-mention or already mentioned. Some participants started each video as an independent story, while some others made it explicit that the characters were the same as in the preceding videos. Some other participants, however, did not give any linguistic indication as whether they perceived the characters as ‘new’ or ‘old’. In these cases, we coded the first mention of the characters as ‘unclear’ and exclude it from the dataset.

3.4 Results and discussion

This section presents and discusses the results of the analysis. In each subsection I first discuss the results, and then propose explanations for the observed patterns.

3.4.1 Frequency of word order patterns

This section focuses on the frequency of word order patterns in the three Ambon Malay groups. An independent *t*-test (in SPSS) is used to compare the mean frequency of word orders in the homeland and in the heritage groups (see Section 2.4). The number of first generation speakers is too small to allow a reliable statistical analysis.

3.4.1.1 Results

This section reports the results regarding the frequency of word order patterns in the three Ambon Malay groups. The data for the demonstratives and the numeral *satu* ‘one’ confirm the prediction that partial overlap between Ambon Malay and Dutch creates the conditions for transfer (see Section 3.3.2), which is instantiated as a change in frequency for the option also possible in Dutch.

Demonstratives. The order of the demonstratives *ini* ‘D.PROX’ and *itu* ‘D.DIST’ and the noun is different in the three groups, as displayed in Figure 3.1 on the next page. Heritage speakers show a significantly higher frequency of pre-nominal *ini* (black bar) ($M=.4917$, $SD=.445$) when compared to homeland speakers ($M=.0104$, $SD=.037$) ($(t(11.079)=3.734$, $p=.003$, $r=.74$, equal variances not assumed), and a lower frequency of post-nominal *ini* (dark gray bar) ($M=.5092$, $SD=.446$) in comparison to the homeland group ($M=.9896$, $SD=.037$). Heritage speakers also show a significantly higher frequency of pre-nominal *itu* (light gray bar) ($M=.59.17$, $SD=43.77$) when compared to homeland speakers ($M=.0715$, $SD=.134$)

($t(35.196)=6.180$, $p<.001$, $r=.72$, equal variances not assumed) and a lower frequency of post-nominal *itu* (white bar) ($M=.4087$, $SD=.437$) in comparison to the homeland group ($M=.9285$, $SD=.134$). First generation speakers fall in between homeland speakers and heritage speakers. They still have a strong preference for post-nominal *ini* and *itu* (like homeland speakers) but they use the pre-nominal option more frequently than their homeland peers (like heritage speakers).

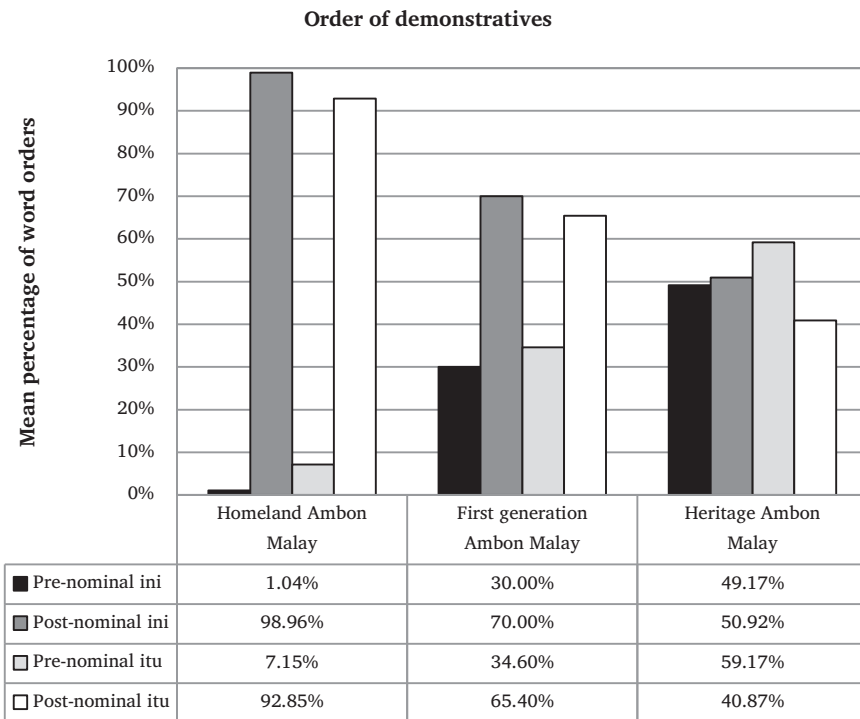


Figure 3.1: Order of the demonstratives in the three Ambon Malay groups.

Numerals. The order of the numerals (except *satu* ‘one’) is the same in the three groups, as displayed in Figure 3.2 on the next page. Overall, pre-nominal numerals (black bar) and post-nominal numerals (dark gray bar) seem to be equally available options in the three groups. Pre-nominal numerals occur with roughly the same frequency in the homeland group ($M=.5467$, $SD=.349$), in the first generation group ($M=.4375$, $SD=.125$) and in the heritage group ($M=.5656$, $SD=.378$). In contrast, the order of the numeral *satu* ‘one’ is different in the three groups. Heritage speakers show a significantly higher frequency of pre-nominal *satu* (light gray bar)

($M = .4324$, $SD = .416$) when compared to homeland speakers ($M = .2167$, $SD = .303$) ($t(51) = 2.115$, $p = .039$, $r = .28$, equal variances assumed), and a lower frequency of post-nominal *satu* (white bar) ($M = .5676$, $SD = .416$) in comparison to the homeland group ($M = .7833$, $SD = .303$). First generation speakers pattern with heritage speakers in showing a higher incidence of pre-nominal *satu*.

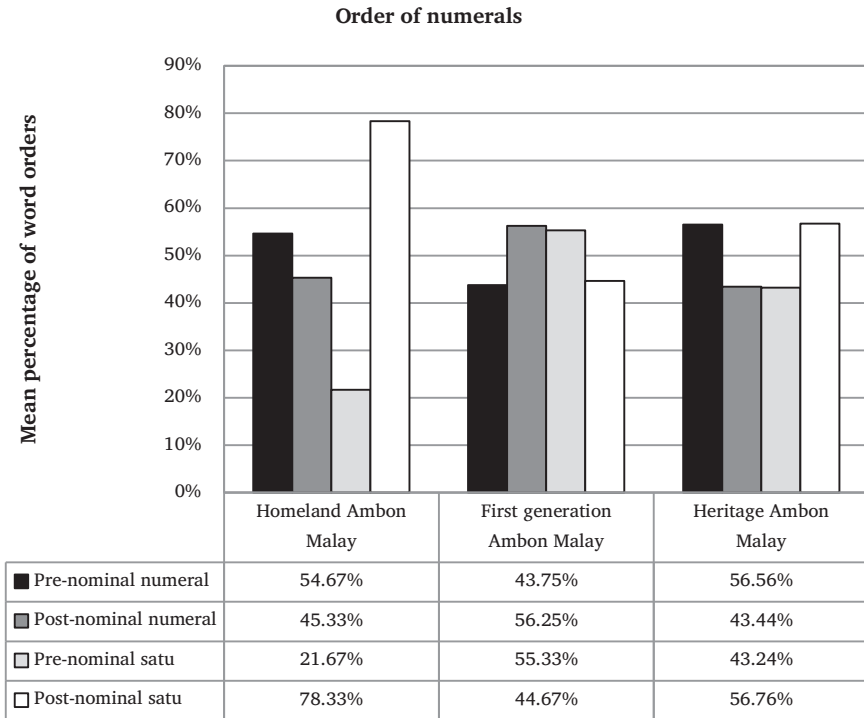


Figure 3.2: Order of the numerals and *satu* ‘one’ in the three Ambon Malay groups.

Let us now move on to the order of the adjectives and the definite marker = *nya*. The prediction that lack of overlap will block transfer effect (see Section 3.3.2) is borne out by the data.

Adjectives. The order of the adjective and noun is the same in the three groups, as shown in Figure 3.3 on the next page.

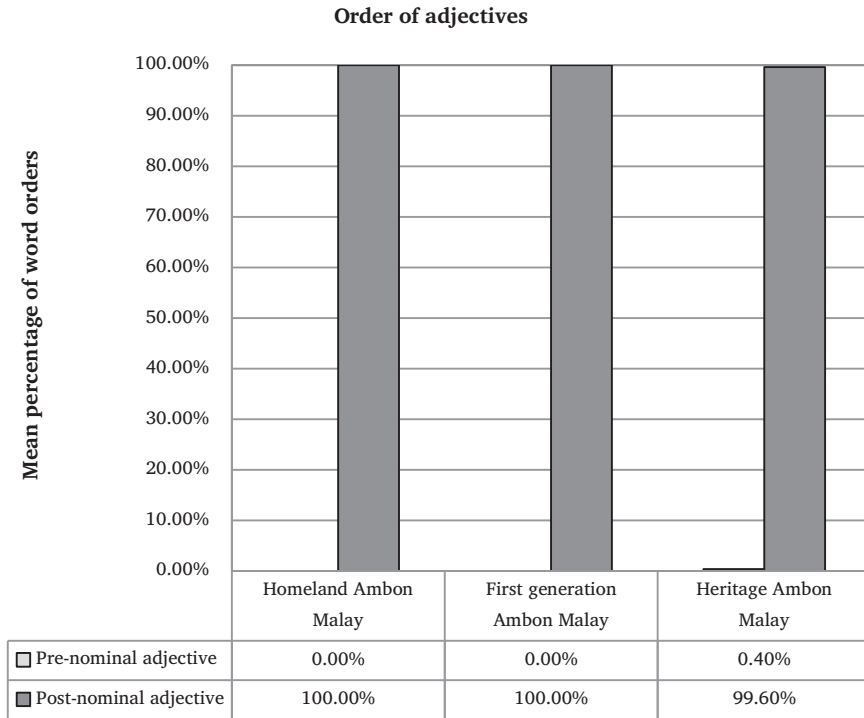


Figure 3.3: Order of the adjective in the three Ambon Malay groups.

All speakers strictly adhere to the NOUN-ADJECTIVE order,⁵⁰ and there is no statistical difference.

The definite marker =nya. The order of the definite marker =*nya* and the noun is the same in the three groups, as shown in Figure 3.4 on the next page. There is no instance of =*nya* used before the noun.

⁵⁰ There is only one token of ADJECTIVE-NOUN order in the whole heritage dataset. This is illustrated in (i).

- (i) *Ini ada ehm lebe b...besar tak*
 D.PROX EXIST ehm more b...big branch (Dutch word)
 ‘Now (he) has a longer branch. [The speaker describes a boy carrying a long tree branch, longer than the one he carried in the previous scene.]’

The reliability of this token is debatable because the speaker makes a rather long pause (about 600 ms) between the adjective *lebe besar* ‘bigger’ and the (Dutch) noun *tak* ‘branch’. The long pause may indicate that *tak* ‘branch’ is an afterthought.

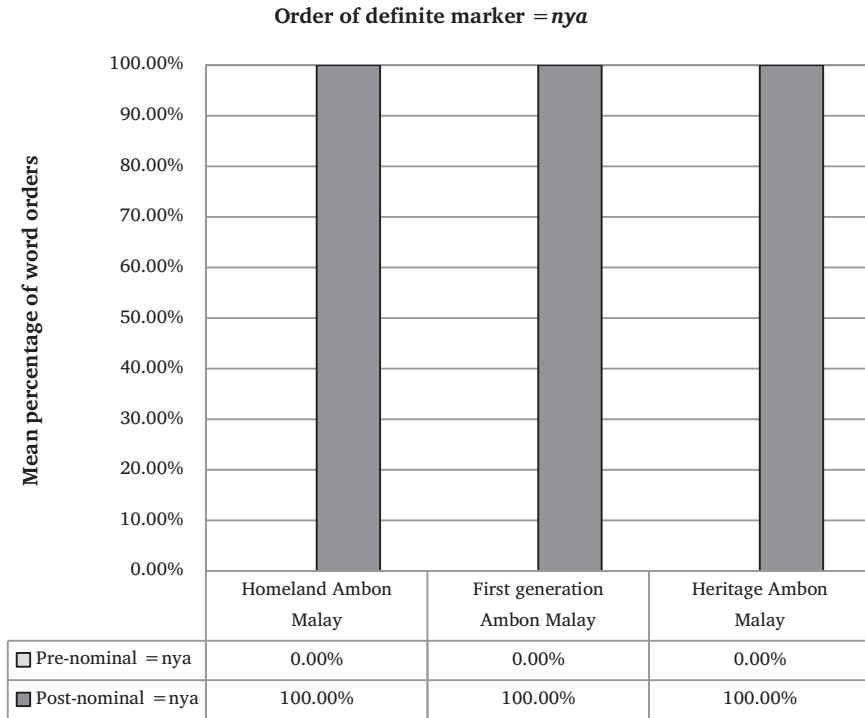


Figure 3.4: Order of the definite marker = *nya* in the three Ambon Malay groups.

3.4.1.2 Discussion

The results illustrated so far allow us to answer the first two questions posed in Section 3.3.2, namely (i) whether variability (partial overlap between the two languages) is likely to trigger cross-linguistic influence, and (ii) whether the lack of variability (no perceivable overlap or complete identity between two languages) is likely to block cross-linguistic influence.

The comparison between the linear order of demonstratives confirms that when the heritage language allows an alternation between two options (i.e., pre-nominal and post-nominal demonstrative), heritage speakers show a stronger preference for the option also possible in Dutch, in this case, the pre-nominal option (DEMONSTRATIVE-NOUN order). This change in preference is accounted for by cross-linguistic influence from Dutch, which is driven by surface overlap between the two languages. In this type of contact-induced change, the influence from the dominant language is ‘indirect’ (Silva-Corvalán, 2008) because the change does not involve the creation of a new word order pattern, but rather a change in the

frequency of the word order patterns that were already available in the heritage language (see Section 1.3.1.1).

This type of cross-linguistic influence is better explained as a case of cross-activation of parallel structures in the dominant and in the heritage languages, rather than as actual ‘transfer’ of linguistic material (Moro & Irizarri van Suchtelen, forthcoming)

The change in frequency observed for *ini* ‘D.PROX’ and *itu* ‘D.DIST’ probably reflects the different entrenchment levels of word order patterns (Backus, 2004; Bybee, 2006; Onar Valk, 2015). It is generally assumed that the more frequent a word order pattern is, the more entrenched it is in the speaker’s repertoire, which in turn increases the likelihood that the speaker will select it in future speech acts. We can speculate that the DEMONSTRATIVE-NOUN order is more entrenched, and therefore more likely to be selected, because this order receives its degree of activation from two languages (Ambon Malay and Dutch), while the NOUN-DEMONSTRATIVE order receives its degree of activation only from Ambon Malay (see Section 7.2.2). In order to establish with certainty whether a decrease in frequency corresponds to a decrease in the entrenchment level, we would need other types of evidence, such as judgment tasks or forced choice tasks which tap into the competence of heritage speakers.

Before turning to numerals, it is first worth making few observations on the language of the first generation group. First generation speakers also show a higher frequency of the DEMONSTRATIVE-NOUN order when compared to homeland speakers (see Figure 3.1). There are two possible explanations for this pattern: (i) first generation speakers are late bilinguals and as such, they are also subject to cross-linguistic influence from Dutch. It is plausible to assume that their language shows signs of attrition due to the long time they have spent in the Netherlands ($M=46$ years); (ii) the language variety brought to the Netherlands by first generation speakers (see Section 1.5.2.1) was characterized by a higher rate of pre-nominal demonstratives. Furthermore, it is likely that Tangsi Malay also had a high rate of DEMONSTRATIVE-NOUN order, like Baba Malay. This second possibility is supported by the fact that, according to Paauw (2008, p. 299), the DEMONSTRATIVE-NOUN order was a typical feature of eastern Malay varieties, including Ambon Malay. The most probable scenario that we can assume at this point is the following. Originally Ambon Malay had a preference for the DEMONSTRATIVE-NOUN order, but under the influence of Standard Indonesian (which became the official language of the Republic of Indonesia in 1945), Ambon

Malay shifted to the NOUN-DEMONSTRATIVE order. First generation speakers, who arrived in the Netherlands in the 1951, brought along a more ‘conservative’ variety of Ambon Malay, one without a strong Indonesian flavor (see Section 1.5.1.1). Heritage speakers, who acquired the language from first generation speakers, found evidence for the DEMONSTRATIVE-NOUN order in the language of their parents, but also in Dutch, their dominant language. As a result of double activation (from the parent’s language and from Dutch), the DEMONSTRATIVE-NOUN order has become increasingly entrenched in the mind of heritage speakers.

In sum, the preference for the DEMONSTRATIVE-NOUN order in heritage speakers is accounted for by cross-linguistic influence from Dutch and by the qualitatively different input heritage speakers were exposed to (see Section 1.3.4). The different preferential tendencies of heritage Ambon Malay (i.e., DEMONSTRATIVE-NOUN) and homeland Ambon Malay (i.e., NOUN-DEMONSTRATIVE) represent an internal change exacerbated by contact: homeland Ambon Malay is subject to the influence of Standard Indonesian which only allows the NOUN- DEMONSTRATIVE order (see Section 3.2.1), while heritage Ambon Malay is subject to the influence of Dutch, which only allows the DEMONSTRATIVE-NOUN order.

The comparison between the linear order of numerals only partially confirms the prediction that heritage speakers will show a stronger preference for the option also possible in Dutch (i.e., NUMERAL-NOUN order). This prediction is confirmed only by the numeral *satu* ‘one’, but not by the other numerals. The reason why homeland speakers and heritage speakers do not differ with respect to the linear orders of numerals is due to the fact that the homeland variety and the heritage variety are changing in the same direction (i.e., NUMERAL-NOUN), but the former due to Standard Indonesian (see Section 3.2.1), and the latter due to Dutch. One observation supports this hypothesis, namely that first generation speakers, whose language was not subject to the influence of Standard Indonesian because they emigrated in 1951, display a higher incidence of the NOUN-NUMERAL order when compared to the other two groups (see Figure 3.2). This may confirm Paauw’s (2008, p. 297) claim that the ‘original’ order in Ambon Malay (and in the other eastern Malay varieties) was indeed NOUN-NUMERAL and that the NUMERAL-NOUN order is a relatively recent development due to Standard Indonesian and other western Malay varieties, such as Colloquial Jakartan Indonesian.

With respect to the numeral *satu* ‘one’, the homeland variety and the heritage variety do differ. In the heritage variety, the position of *satu* is changing at a faster

rate in than in the homeland variety. As for the demonstratives (see above), cross-linguistic influence from Dutch, driven by surface overlap, is the factor accounting for this change. The frequent and obligatory use of the *one*-NOUN order in Dutch is likely to reinforce the parallel structure in the repertoire of heritage speakers who use it also when speaking Ambon Malay. It is probable that in the homeland variety, *satu* changes at a slower rate because everyday use and routine ‘protect’ this numeral from cross-linguistic influence from Standard Indonesian (cf. Matras, 2011, p. 213).

The comparison between the linear orders of the adjectives and the definite marker =*nya* confirms the prediction that lack of variability in Ambon Malay and the consequent lack of overlap with Dutch blocks cross-linguistic influence (see Section 3.3.2). Silva-Corvalán (1994, p. 134) observed that only the linguistic features that are compatible with the structure of the heritage language will be adopted and diffused into the community. In other words, the typological dissimilarity between the heritage language and the dominant language limits structural transfer. In the case of Ambon Malay, the lack of clear structural similarity between the languages makes heritage Ambon Malay impermeable to Dutch influence and disallows heritage Ambon Malay to converge toward Dutch (see Section 7.2).

We have seen that the re-ordering of the demonstratives and the numeral *satu* ‘one’ is arguably related to the partial overlap between Ambon Malay and Dutch; heritage speakers select the pre-nominal order because this order matches the word order in the dominant language, Dutch. At this point, however, I cannot exclude the possibility that simplification, intended as a change from flexible-word order to rigid word order (Polinsky & Kagan, 2007; see Section 1.3.5), also plays a role. In other words, Dutch influence may be reinforced by the internal pressure toward a fixed word order.

In the next section, I examine whether grammatical re-interpretation of the demonstratives *ini* ‘D.PROX’ and *itu* ‘D.DIST’ and the numeral *satu* ‘one’ on the model of Dutch articles is also responsible for their re-ordering.

3.4.2 Grammaticalization of (in)definite markers

This section zooms in on the overall frequency and distribution of (in)definite markers (*ini* ‘D.PROX’, *itu* ‘D.DIST’, =*nya* ‘DEF’, and *satu* ‘one’). A generalized mixed effects model is used to assess the effect of group (homeland, first generation, and heritage) and type of noun (inanimate and animate) on the use the (in)definite

markers, with speaker as a random effect (see Section 2.4 for an explanation of the generalized mixed effects model).

3.4.2.1 Results

This section reports the results regarding a possible effect of grammatical re-interpretation on the demonstratives *ini* 'D.PROX' and *itu* 'D.DIST', the numeral *satu* 'one', and the marker =*nya* 'DEF'. Since grammaticalization increases obligatoriness (and thus frequency) of a morpheme (see Section 3.3.2 and Section 1.3.1.3), I tested grammaticalization by measuring the overall frequency of these morphemes in the homeland and in the heritage variety. The frequency of *ini* 'D.PROX', *itu* 'D.DIST' and =*nya* 'DEF' was measured on already mentioned nouns, the nouns that are most commonly marked as definite, while the frequency of *satu* 'one' was measured on first-mention nouns, the nouns that are more commonly marked as indefinite. In both measurements, nouns were further subdivided into animate and inanimate.

Ini 'D.PROX' is used to mark already mentioned nouns with approximately the same frequency in the three groups, although it is less frequent in the heritage group, as shown in Table 3.3. The data also show that, in the three groups, animate nouns are more likely to be marked by *ini* than inanimate nouns.

Table 3.3: Mean frequency of *ini* 'D.PROX' on all already mentioned nouns, on inanimate (already mentioned) nouns, and on animate (already mentioned) nouns.

	GROUP	n	MEAN %	SD
All already mentioned nouns	Homeland	27	4.4%	.047
	First generation	5	1.8%	.019
	Heritage	32	2.1%	.042
Inanimate nouns	Homeland	27	3.1%	.032
	First generation	5	1.9%	.019
	Heritage	32	1.5%	.041
Animate nouns	Homeland	27	6.2%	.0610
	First generation	5	2.7%	.039
	Heritage	32	3.2%	.096

The generalized mixed effects model reveals that animate nouns are more frequently marked by *ini* than inanimate nouns in all groups ($\beta=0.34266$, $SE=0.07329$,

$p < .001$), and that heritage speakers are slightly less likely to use *ini* than homeland speakers ($\beta = -0.52504$, $SE = 0.19891$, $p = .008$). The overall model is significant ($\chi^2(3) = 30.429$, $p < .001$), when compared to a null model with only speaker as a random effect, meaning that the group to which the speaker belongs and the type of noun do have an influence on the use of *ini*.

The overall frequency of *itu* 'D.DIST' is roughly the same in the three groups, as shown in Table 3.4. The data also show that, in the homeland and in the first generation group, inanimate nouns are more likely to be marked by *itu* than animate nouns, while heritage speakers do not follow this tendency and use *itu* regardless of animacy.

Table 3.4: Mean frequency of *itu* 'D.DIST' on all already mentioned nouns, on inanimate (already mentioned) nouns, and on animate (already mentioned) nouns.

	GROUP	n	MEAN %	SD
All already mentioned nouns	Homeland	27	12.8%	.112
	First generation	5	11.8%	.074
	Heritage	32	10.3%	.149
Inanimate nouns	Homeland	27	17.1%	.131
	First generation	5	15.8%	.077
	Heritage	32	10.7%	.138
Animate nouns	Homeland	27	9.4%	.144
	First generation	5	5.3%	.062
	Heritage	32	9.9%	.192

The generalized mixed effects model reveals that inanimate nouns attract *itu* more than animate nouns in all groups ($\beta = -0.25017$, $SE = 0.04847$, $p < .001$). There is no difference in the use of *itu* among the three groups. The overall model is significant ($\chi^2(3) = 30.09$, $p < .001$), when compared to a null model, meaning that the type of noun (but not the group) has an effect on the use of *itu*.

The enclitic =*nya* 'DEF' is used more by heritage speakers than by homeland speakers, as shown in Table 3.5 on the next page. The data also show that first generation speakers pattern with heritage speakers in displaying a relatively high frequency of =*nya*. In neither of the groups, the use of =*nya* is subject to animacy.

Table 3.5: Mean frequency of =*nya* ‘DEF’ on all already mentioned nouns, on inanimate (already mentioned) nouns, and on animate (already mentioned) nouns.

	GROUP	n	MEAN %	SD
All already mentioned nouns	Homeland	27	0.7%	.015
	First generation	5	16.3%	.146
	Heritage	32	11.1%	.144
Inanimate nouns	Homeland	27	0.4%	.015
	First generation	5	17.2%	.114
	Heritage	32	11.5%	.143
Animate nouns	Homeland	27	0.9%	.021
	First generation	5	12.8%	.191
	Heritage	32	10.0%	.178

The generalized mixed effects model reveals that =*nya* is more frequent in heritage speakers ($\beta=1.54826$, $SE=0.28811$, $p<.001$) and in first generation speakers ($\beta=1.91306$, $SE=0.46807$, $p<.001$) than in homeland speakers. There is no difference among animate and inanimate nouns. The overall model is significant ($\chi^2(3)=34.45$, $p<.001$), when compared to a null model with only speaker as a random effect, meaning that the group to which the speaker belongs is a good predictor for the use of =*nya*, with heritage and first generation speakers using it more often.

Satu ‘one’ is used on first-mention nouns with approximately the same frequency in the three groups, as shown in Table 3.6 on the next page. The results also show that, in the three groups, animate nouns are more likely to be marked by *satu* than inanimate nouns. The generalized mixed effects model reveals that animate nouns attract *satu* more than inanimate nouns ($\beta=1.11097$, $SE=0.06979$, $p<.001$), but there is no difference among the three groups. The overall model is significant ($\chi^2(3)=263.71$, $p<.001$), when compared to a null model with only speaker as a random effect, meaning that the type of noun (but not the group) has an effect on the use of *satu*.

Table 3.6: Mean frequency of *satu* ‘one’ on all first-mention nouns, on inanimate (first-mention) nouns, and on animate (first-mention) nouns.

	GROUP	n	MEAN %	SD
All first-mention nouns	Homeland	27	8.7%	.073
	First generation	5	6.9%	.095
	Heritage	32	5.6%	.055
Inanimate nouns	Homeland	27	4.1%	.048
	First generation	5	5.4%	.090
	Heritage	32	3.0%	.040
Animate nouns	Homeland	27	28.1%	.219
	First generation	5	13.9%	.143
	Heritage	32	16.3%	.158

3.4.2.2 Discussion

The results presented in Section 3.4.2.1 show that (i) the demonstratives and the numeral ‘one’ are not used more frequently in heritage speakers – a result that we would expect if they were becoming more grammaticalized; (ii) the demonstratives and the numeral ‘one’ are not extended to new contexts (*ini* ‘D.PROX’ and *satu* ‘one’ are not extended to inanimate nouns, and *itu* ‘D.DIST’ is not extended to animate nouns) - another result that we would expect if they were becoming more grammaticalized. The definite marker =*nya*, on the other hand, is used increasingly frequently by heritage speakers, a symptom of embryonic contact-induced grammaticalization (see Section 1.3.1.3). We can therefore say that the results do not confirm the prediction that *ini* ‘D.PROX’, *itu* ‘D.DIST’ and *satu* ‘one’ are proceeding a step further in their grammaticalization process under the influence of Dutch. Nevertheless, the prediction that the Dutch category of definiteness will be replicated in heritage Ambon Malay is partially confirmed by the results regarding =*nya* ‘DEF’, which has strengthened its role as definite marker (see Table 3.5).

There are two factors that may explain why we do not see the expected high frequency and distribution of the demonstratives and the numeral ‘one’. One possible reason is that the contact between Ambon Malay and Dutch has been too short and too moderate to allow such development. After all, Ambon Malay has been in contact with Dutch for about 60 years, and we know that contact-induced grammaticalization is a long and gradual process that usually extends over hundreds

of years (see Section 1.3.1.3). The psycholinguistic process that subsumes contact-induced grammaticalization is labeled ‘functional convergence’ by Sánchez (2004, 2006; see Section 1.4.1). In this process, speakers re-associate salient features of their dominant language to forms of their heritage language. It is possible, then, that a process of functional convergence toward Dutch is taking place but its effects are still too weak to surface, at least in the task described in the present chapter.

In the case of *ini* ‘D.PROX’ and *itu* ‘D.DIST’, grammaticalization may be hindered by the presence of the definite marker =*nya* ‘DEF’. The reason why we do not find effects for *ini* and *itu* might be that they are mapped onto Dutch demonstratives, and definiteness marking is selected via =*nya*. *Itu* ‘D.DIST’, in particular, is in competition with =*nya* ‘DEF’, as shown by the fact that, in heritage speakers, these two forms occur with approximately the same frequency (10.3% and 11.1%, respectively) and the same distribution among animate and inanimate nouns. Some heritage speakers show a neat preference for either of these two markers, while other speakers use both markers equally. For instance, speaker H12 has a clear preference for *itu* (50.8%) over =*nya* (0.06%), speaker H28 strongly prefers =*nya* (42.6%) over *itu* (0.01 %); whereas speaker H1 uses *itu* (8.2%) and =*nya* (8.2%) with equal frequency. Note that H12, H28, and H1 are all male speakers, of approximately the same age (55, 59, 59 years old) and with similar biographical characteristics, namely they all grew up in Moluccan camps, and they reported speaking mainly Ambon Malay with both parents (see Table 2.1 in Chapter 2). The differences among these speakers may be explained in terms of the language variety they were exposed to. Since =*nya* was not part of the Ambon Malay grammar (van Minde, 1997), heritage Ambon Malay must have inherited this form from Tangsi Malay (see Section 1.5.2). It is, therefore, probable that the heritage speakers who frequently use =*nya* (e.g., H28) were mainly exposed to Tangsi Malay. I come back to this issue below.

It seems that in some speakers =*nya* ‘DEF’ may be undergoing contact-induced grammaticalization at the expense of *itu* ‘D.DIST’. These heritage speakers have recruited linguistic material from their heritage language (i.e., the definite marker =*nya*) to overtly express a category that is grammaticalized in their dominant language (i.e., definiteness). A very explanatory example is presented in (18), where the heritage speaker leaves the noun unmarked when it is a first-mention noun, but then marks it consistently with the definite marker in the following occurrences (in bold), as a speaker of Dutch or English would do.

(18) *Ana dua ada lia dong pung pakaiang di atas pohong*
 child two EXIST see 3PL POSS cloth at above tree
 ‘Two children are looking at their cloth on a tree,

[...] *dorang balumpa tetapi pakaian = nya talalu tinggi*
 3PL jump but cloth = DEF too high
 they jump but **the cloth** is too high,

ana satu ada ambe barang, dia datang dengang peti
 child one EXIST take thing 3SG come with box
 a child takes something, he comes with a box,

pada peti = nya mungkin talalu kacil
 but box = DEF maybe too small
 but **the box** is probably too small,

sampe dong susa ambe pakaian = nya, pakaian = nya
 until 3PL difficult take cloth = DEF cloth = DEF
 so they can’t take **the cloth, the cloth**

talalu tinggi sekarang anak = nya ada ambe krosi
 too high now child = DEF EXIST take chair
 is too high, now **the child** takes a chair

dia taro krosi = nya di tana
 3SG put chair = DEF at ground
 he puts **the chair** on the ground...’

Although it would be tempting to conclude that the high incidence of = *nya* ‘DEF’ in heritage speakers is a case of embryonic contact-induced grammaticalization, we need to consider that = *nya* is highly frequent in first generation speakers as well. This fact indicates that = *nya* was part of the language brought to the Netherlands by the Moluccans who arrived in 1951, and therefore, was part of the input heritage speakers have received from their parents. As pointed out above, the presence of = *nya* in heritage Ambon Malay is probably a trace of Tangsi Malay influence. Again, the most probable scenario that we can assume is a multicausal one, whereby

an internal change is accelerated by contact (see Section 1.3.4). The frequency of =*nya*, then, is partially due to the fact that this form was present in the language that heritage speakers acquired, and partially due to contact with Dutch, which has further entrenched =*nya* in the heritage variety and among first generation speakers, as a result of attrition. We will see in Chapter 7 (Section 7.3.1), that two factors predict a high use of =*nya* among heritage speakers: (i) sequential bilingualism, and (ii) living outside a Moluccan ward. Notably, having spoken Ambon Malay consistently in early childhood and being immersed in a Dutch speaking environment in adulthood have an effect on the use of =*nya*. These findings support the above mentioned conclusion, that the high incidence of =*nya* depends on both cross-linguistic influence from Dutch (speakers living outside a Moluccan ward are more exposed to Dutch) and language usage (sequential bilinguals had higher exposure to and use of the heritage language).

In the case of *satu* ‘one’, grammaticalization may be hindered by the lack of complete equivalence between *één* ‘one’ and *een* ‘ART.INDF’ in Dutch. The fact that the numeral and the indefinite article are no longer homophonous in Dutch may slow down the re-interpretation of the numeral *satu* to an indefinite article.⁵¹

To conclude, since the demonstratives and *satu* ‘one’ are not undergoing contact-induced grammaticalization, the higher incidence of pre-nominal demonstratives and pre-nominal *satu* reported in Section 3.4.1.1 can only be due to cross-linguistic influence from Dutch, driven by surface overlap, and not to the re-interpretation of these morphemes on the model of Dutch articles. If Cleary-Kemp (2007) is right in her view that Ambon Malay has a system similar to that of Baba Malay, where the pre-nominal position of *ini* ‘D.PROX’ and *itu* ‘D.DIST’ is associated with the demonstrative function, and the post-nominal position with the article function, then heritage speakers in the Netherlands have extended the pre-nominal position to articles, but they have not increased their obligatoriness.

⁵¹ One way to test for the role of homophony would be to study heritage speakers of Ambon Malay with a dominant language where the numeral ‘one’ and the indefinite article are homophonous, such as German (*ein* ‘one’ and *ein* ‘ART.INDF’). If homophony in the dominant language plays a role, one may expect this hypothetical group to proceed faster in the grammaticalization of *satu* ‘one’.

3.5 Conclusions

This chapter has investigated word order changes in noun phrases in heritage Ambon Malay in light of two possible causes: variability (alternation between two patterns) and grammatical re-interpretation.

The results of the first study show that variability does have an effect on word order. In fact, when Ambon Malay allows an alternation between two word order options, heritage speakers show a stronger preference for the option also possible in Dutch, but when Ambon Malay lacks such an alternation, heritage speakers do not differ from homeland speakers. In light of this finding, I conclude that cross-linguistic influence from Dutch manifests itself as a change in frequency. The shift in frequency toward the Dutch-aligned word order will ultimately lead to greater syntactic convergence between Dutch and heritage Ambon Malay (Winford, 2003; Backus, 2004; Matras, 2009; Muysken, 2013).

The results of the second study show that grammatical re-interpretation does not seem to have an effect on word-order changes. In fact, the demonstratives *ini* 'D.PROX', *itu* 'D.DIST' and the numeral *satu* 'one' are not functionally converging toward Dutch. These morphemes are used with a different order by heritage speakers, but with apparently the same function as in the homeland language. Nevertheless, the prediction that the Dutch category of definiteness will be replicated in heritage Ambon Malay is partially confirmed by the results regarding *=nya* 'DEF'. The increase in the frequency of this morpheme seems to be a symptom of incipient contact-induced grammaticalization.

Both studies also show that divergence between the homeland language and the heritage language is better accounted for by two sources: cross-linguistic influence from the dominant language and the different type of input heritage speakers were exposed to. Internal differences between the homeland variety and the heritage variety already existed when first generation speakers arrived in the Netherlands in 1951. Over the past 40-50 years cross-linguistic influence from Dutch onto heritage Ambon Malay and from Indonesian onto homeland Ambon Malay has accelerated such internal changes.

CHAPTER 4

Aspectual distinctions in heritage Ambon Malay⁵²

4.1 Introduction

Aspect is one of the grammatical phenomena that have attracted most attention in language contact studies, such as in bilingualism research (Flecken, 2010; Bylund & Jarvis, 2011, Schmiedtová et al., 2011), L2 acquisition studies (Andersen & Shirai, 1994; Bardovi-Harlig, 2000; Ma, 2006) and heritage language studies (Polinsky, 2008b; Montrul, 2009; Laleko, 2010). One of the reasons is that the category of aspect has a strong conceptual dimension and as such it is expected to remain stable in language contact situations even though the means for expressing it may change. As pointed out by Polinsky (2008b, p. 280), “[a]spectual distinctions are universal belonging with the conceptual representation of events. What varies is the actual linguistic encoding of these distinctions, but not the underlying distinctions themselves” (see also van Hout, de Swart, & Verkuyl, 2005). If on the one hand the conceptual category of aspect is universal and hence stable, the expression of aspectual contrasts has been shown to be rather unstable in language contact settings, especially in heritage languages. Changes have been attested both in the expression of grammatical aspect (Koontz-Garboden, 2004; Montrul, 2009; Laleko, 2010; Shi, 2011) and lexical aspect (Polinsky, 2008; Laleko, 2010; Shi, 2011). I now give a brief overview of these two components of verbal aspect.

Grammatical aspect is the grammatical expression of the internal temporal constituency of a situation. The basic distinction in Comrie’s theory (1976) is between perfective, imperfective and perfect. Perfective aspect presents the situation as a complete whole, without further subdivision into temporal phases, (e.g., *John read the book yesterday*). Imperfective aspect views the situation from inside and focuses on the internal phase, as in *John was reading* (imperfective), *when I entered* (perfective). The category of imperfective is further subdivided into a number of sub-types, including the progressive (e.g., *John is singing*) and the iterative (e.g., *He*

⁵² This chapter is partially based on Moro, F. R. (in press). Aspectual distinctions in Dutch-Ambon Malay bilingual heritage speakers. *International Journal of Bilingualism*.

keeps on eating and eating) (for an overview see Bybee, Perkins, & Pagliuca, 1994, p. 122). The perfect indicates the continuing relevance of a past situation, as in *I have lost my penknife* (perfect) vs. *I lost my penknife* (perfective) (p. 52). As such, the perfect is not an aspect proper because it is not concerned with the internal temporal properties of a situation. Most accounts of aspect, however, do list the perfect as a tense/aspect category (Comrie, 1976; Givón, 1982; Bybee et al., 1994). I follow this convention here also because the perfect is one of the categories that will be discussed in the present study (see Section 4.3). Finally, some languages also have dedicated expressions for prospective aspect, which relates a state to some subsequent situation, such as English *to be going to* or *to be about to* (p. 64).

Lexical aspect, also referred to as inherent aspect, refers to the temporal properties that are encoded in the lexical meaning of the verb. In fact, there are verbs that have a built-in endpoint, such as *fall*, and verbs that do not have an inherent endpoint, such as *love*. Vendler (1967) classified such verbs into four classes: state, activity, accomplishment and achievement (see Filip, 2011 for a complete overview of different theories relative to aspectual classes). States do not involve a process and have no endpoint, as such they can persist invariable over an indefinite amount of time (e.g., *know, love, have* etc.). Activities involve a process but no specific endpoint, hence they can be protracted indefinitely (e.g., *run, walk, play, talk*). Accomplishments also involve a process but with an inherent endpoint, beyond which the process cannot continue (e.g., *make a chair, build a house, read a novel*). Achievements are not processes, but rather time instants with an inherent endpoint (e.g., *reach the summit, find an object*). The main criticism to Vendler's classification is that aspect is not exclusively a verbal matter, but is determined compositionally through the interaction of the verb with its arguments (Verkuyl, 1989; Laleko, 2010). For instance, the verb *run* is an activity in the sentence *he is running*, but it is an accomplishment in *he ran a mile*. Adding a direct object to the verb contributes to its telicity because the spatial delimitation of the object translates into the temporal delimitation of the verb (the verb needs to have an endpoint) (Laleko, 2010; pp. 145-150). For the sake of convenience, I adopt Vendler's terms, but I use them to refer to situations rather than to verbs.

Although grammatical aspect and lexical aspect are theoretically two distinct notions, they often correlate in language usage (Comrie, 1976, pp. 41-51). For instance, activities and accomplishments are more likely to be marked grammatically by progressive aspect because they involve duration (Flecken, 2010, pp. 134-137). Since progressive aspect presents the internal phase of a situation,

situations that can be divided into phases are more suitable to be marked by progressive. On the other hand, situations that have an endpoint, and thus have been brought to completion, are more likely to be marked with the perfective or the perfect.

Studies on aspect in language contact have investigated various issues related to this topic, including the expression of grammatical and lexical aspect in bilinguals, and the relation between the two in language usage (see the Aspect Hypothesis in Bardovi-Harlig, 2000). In bilingualism research, aspect is considered to be a key category for the study of cognitive processes in bilinguals because aspectual distinctions reflect language-specific way of conceptualizing events (see the Conceptual Transfer Hypothesis in Section 1.4.2). Studies like those of von Stutterheim and Nüse (2003), Flecken (2010), Bylund and Jarvis (2011) and Schmiedtová et al. (2011) have shown that bilingual speakers differ significantly from their monolingual peers in the choice of aspectual distinctions, and that cross-linguistic influence from the dominant language is the main factor to account for such differences. Bilingual speakers, in fact, tend to adopt the time-schemas of the dominant language and map them onto the expression of grammatical aspect in the other language. In the field of L2 acquisition, aspect has been investigated to find support for the Aspect Hypothesis, which predicts that in the interlanguage of L2 learners grammatical aspect is influenced by the lexical aspect of the verb.⁵³ According to the hypothesis, L2 learners tend to select specific aspect markers which reduplicate the inherent aspect expressed by the situation, so that perfective past is used only to mark accomplishments and achievements (telic situations), while progressive is used for activities. The main criticism to the Aspect Hypothesis is that the use of tense-aspect markers in L2 learners simply reflects the skewing of distribution found in the target language (the Distributional Bias Hypothesis; see Bardovi-Harlig, 2000, pp. 424-425; Wible & Huang, 2003; Laleko, 2010, pp. 114-115).

In this chapter, I explore the effects of contact on the aspectual system of heritage Ambon Malay, the variety of Ambon Malay spoken by Dutch-Ambon Malay bilinguals in the Netherlands. Dutch and Ambon Malay represent an interesting language pair because Dutch obligatorily marks a past/non-past contrast; however the degree to which aspect is (periphrastically) marked is variable, whereas Ambon

⁵³ The Aspect Hypothesis is a family of hypotheses that make specific predictions about the order of emergence of grammatical markers (Bardovi-Harlig, 2000, p. 227). Here, however, we are concerned only with the general tenet of the hypothesis.

Malay lacks a grammaticalized expression of tense, but has a number of optional aspect markers. Furthermore, Dutch has a clear finite/non-finite contrast, with tense morphology (in addition to agreement morphology) instantiating finiteness. This study investigates contact-induced changes in the aspectual system of heritage Ambon Malay by focusing on four main (tense)aspect markers, namely *ada* (progressive), *su* (perfect), *mau* (volitional, prospective) and reduplication (iterative). The study investigates whether the frequency and the usage of these aspect markers differ between heritage bilingual speakers in the Netherlands and homeland speakers in Ambon and, and whether the aspectual system of Ambon Malay is undergoing restructuring under the influence of Dutch.

This chapter is organized as follows. The next section gives an overview of the various factors responsible for the restructuring of the aspectual system of contact varieties. Section 4.3 describes the main aspectual contrasts expressed in Ambon Malay (Section 4.3.1) and Dutch (Section 4.3.2), the two languages of heritage speakers. Section 4.4 illustrates the design of the present study. The results are presented and discussed in Section 4.5. Section 4.6 summarizes the conclusions.

4.2 The restructuring of aspect in heritage languages

Contact phenomena attested in the aspectual systems of heritage languages are generally of two kinds: decrease in frequency or loss of aspectual distinctions (as reported for heritage Russian in the U.S.A. by Polinsky, 2008 and Laleko, 2010; for heritage Spanish in the U.S.A. by Silva-Corvalán, 1994 and Montrul, 2009), and overextension of progressive forms to mark imperfective aspect (as reported for heritage Spanish in the U.S.A. by Koontz-Garboden, 2004; for Pennsylvania Dutch in the U.S.A. by Brown & Putnam, 2015; for heritage Mandarin in the Netherlands by Shi, 2011). Recent approaches to heritage languages have identified a number of factors that play a role in shaping heritage grammars (Koontz-Garboden, 2004; Polinsky, 2008; Laleko, 2010; Bylund & Jarvis, 2011; O’Grady et al., 2011).

The first factor, which is directly related to the dominant language, is cross-linguistic influence. As seen in Section 1.3.1, cross-linguistic influence can lead to an increase or to a decrease in the use of specific aspect markers, depending on what is obligatorily encoded in the dominant language of heritage speakers. Studies such as Sánchez (2004, 2006), Flecken (2010) and Bylund and Jarvis (2011) have demonstrated that bilingual speakers tend to overtly express the categories that are

grammaticalized in their dominant language (see the Functional Convergence Hypothesis and the Conceptual Transfer Hypothesis in Section 1.4). For instance, Koontz-Garboden (2004) shows that Spanish-English bilinguals use progressive aspect more frequently than their monolingual peers, and he argues that this is due to the influence of English, a language in which progressive aspect is highly grammaticalized. By contrast, Bylund and Jarvis (2011) show that Spanish-Swedish bilinguals use fewer progressive forms than their monolingual peers, and they argue that this is due to the fact that bilinguals are affected by the Swedish-like tendency to attend to the telicity rather than the ongoingness of events. Sánchez (2004, 2006) shows that intense cross-linguistic influence can lead to convergence in the TMA system of the two languages because “TMA systems are sensitive to new associations between abstract functional features from one language to overt morphological forms from another language” (2006, p. 289). Intense cross-linguistic influence can ultimately lead to contact-induced grammaticalization, a well-known process whereby bilingual speakers replicate a prominent (obligatory) category of the dominant language (e.g., tense) using the ‘linguistic material’ of the heritage language (e.g., an aspect marker) (Heine & Kuteva, 2005; see also Section 1.3.1.3).

There are other factors shaping heritage languages that are not directly related to the dominant language, but are related to the effects of bilingualism and to limited exposure to the heritage language (especially after adolescence). These factors, which are sometimes referred to as language ‘internal’ factors, are indeterminacy, frequency and acoustic salience (Montrul, 2009; O’Grady et al., 2011; Laleko & Polinsky, 2013). Indeterminacy refers to forms that have a non-transparent form-meaning mapping because they are optional and functionally ambiguous (Laleko & Polinsky, 2013). For instance, Montrul (2009) shows that heritage speakers of Spanish make more errors with the imperfective than with the preterite, and argues that this is due to the fact that “the imperfect [...] represent[s] [a] relation of one form to several meanings, and [is] thus more inherently complex because the mappings are not always transparent” (p. 266). Another example is that reported by Ma (2006), who shows that the functional-semantic ambiguity of the Mandarin perfective aspect marker *-le* is problematic for English L2 learners. The author argues that L2 learners both overuse and underuse *-le* in different contexts, they overuse it when they reanalyze this form as a past tense marker (on the model of English), and underuse when it serves as a discourse marker. Indeterminacy, thus, can lead to destabilization of both the original frequency and the function of a form. In heritage languages, indeterminate forms tend to become unstable because their

successful acquisition depends on frequent exposure. But frequent exposure is precisely what heritage speakers lack(ed), as they grew up acquiring another (dominant) language since early childhood (O'Grady et al., 2011).

Unlike indeterminacy, input frequency and acoustic salience contribute to the stability of forms because they increase availability in the input (O'Grady et al., 2011). In other words, if a form is always there and is highly audible, there are more chances that it will be retained. O'Grady and Hattori (2012) present data showing that heritage speakers of Korean perform better in a comprehension task when the acoustic salience of case markers is manipulated so as to increase the volume, the duration, and the pitch. Another example is Aalberse and Moro (2014), who argue that the Ambon Malay aspect marker *ada* is stable in the heritage variety, despite being semantically indeterminate (see the next section for a detailed discussion of *ada*). They propose that, when semantic indeterminacy is compensated for by frequency and phonological salience (*ada* carries stress and contains two high sonority vowels), the form may undergo functional extension rather than loss.

Finally, other factors may play a role in the restructuring of the aspectual system. For instance, Polinsky (2008) considers loss of morphology and the reduced lexical competence of heritage speakers as two possible causes affecting the expression of aspect in heritage Russian. Heritage speakers of Russian retain just one member of an aspectual pair (perfective-imperfective), independent of aspect. It is hard to predict which form will be retained, but factor such as frequency and telicity seem to play a role. Reduced lexical competence may account also for the limited number of reduplicated verbs in heritage Mandarin, as reported by Shi (2011). Thus, morphological complexity and lexical knowledge can also have repercussions for the heritage verbal system.

To summarize, cross-linguistic influence from the dominant language can either lead to loss or to the overextension of forms in the heritage language, depending on which tense-aspectual notions are grammaticalized. Language-internal factors, such as indeterminacy, frequency and salience, can also contribute to the (in)stability of the frequency and the function of aspectual forms. . Finally, there seems to be a general tendency for progressive forms, such as *estar + ndo* in Spanish (Klein, 1980; Koontz-Garboden, 2004), *zai* in Mandarin (Shi, 2011), *ada* in Ambon Malay (Aalberse & Moro, 2014) and *am + INF sein* (Brown & Putnam, 2015) to undergo overextension. One possible reason for this phenomenon is that progressive aspect is very open to reanalysis (see Bybee et al., 1994, pp. 127-144; Hengeveld, 2011, p. 590; see also Section 4.5); another reason could be that, in the case of Spanish and

possibly of Pennsylvania Dutch, progressive forms allow the speaker to avoid verbal inflection, which is one of the most difficult components for heritage speakers (Benmamoun et al., 2010).

4.3 Tense-aspect in Ambon Malay and Dutch

This section presents the Ambon Malay grammatical markers under investigation (Section 4.3.1). In order to provide the reader with an overview, I also briefly describe how tense-aspectual distinctions are marked in Dutch, the dominant language of heritage speakers (Section 4.3.2). An interim summary is presented in Section 4.3.3. Unless otherwise specified, the examples presented in this section and elsewhere in this chapter are all from the dataset collected for this dissertation (see Section 2.2 and Section 2.3).

4.3.1 Tense-aspect in Ambon Malay

Ambon Malay does not mark tense, but it marks grammatical aspect analytically by means of three aspect particles, *ada* (progressive), *su* (perfect), *mau* (prospective, volition) and reduplication (iterative, intensifier).⁵⁴ Aspect marking is optional in Ambon Malay and the correct temporal interpretation of the utterance is often inferred from the (extra-) linguistic context (van Minde, 1997, p. 189). Pure tense is expressed by time adverbs if required (e.g., *beso* ‘tomorrow’ and *kamareng* ‘yesterday’).

The existential verb *ada* ‘be (somewhere)’ can function as a progressive aspect marker when it precedes a verb, for this reason it is glossed ‘EXIST’. The pre-verbal marker *ada* mainly marks progressive and continuous aspect (van Minde, 1997, p. 191), but in some contexts it can also have a perfect meaning. With predicates describing a process, such as activities and accomplishments, *ada* indicates the ‘ongoingness’ of the event, as illustrated in (1).

⁵⁴ Other aspect markers are the adverbs *balong* ‘not yet’, a combination of negation and aspect, and *masi* ‘still’. The verbs *suka* ‘to like’, *jaga* ‘to guard’, and the noun *tukang* ‘craftsman’ may be used to express habitual aspect (van Minde, 1997).

- (1) *Tikus ada mamasa di dapur*
 mouse EXIST cook at kitchen
 ‘A mouse is cooking in the kitchen.’

With stative verbs, such as *saki* ‘be sick’ or *tabuka* ‘be open’, *ada* indicates non-permanent or reversible states (in Indonesian the marker *sedang* also has this function, see Grangé, 2011, p. 46). An illustration is given in (2), where the sickness experienced by the subject is perceived (by the speaker) as a temporary condition.

- (2) *Dia ni ada saki*
 3SG D.PROX EXIST sick
 ‘He is sick.’

When *ada* precedes verbs that have an inherent endpoint as part of their lexical meaning, such as accomplishment and achievements, it can carry a perfect meaning (cf. van Minde & Tjia, 2002, p. 293), as illustrated in example (3).⁵⁵ In (3) *ada* is not a marker of progressive aspect because the event expressed by the serial verb construction (snapping a stick into two pieces) is punctual and it has already happened. In fact, the speaker is able to describe the result of it.

- (3) *Parampuang ada pata kayo jadi dua*
 girl EXIST break stick become two
 ‘A girl has broken a stick into two.’

The marker *su*⁵⁶ is a tense-aspect marker whose functions overlap with those of the English perfect (*have –ed*) and of the adverb ‘already’. Generally speaking, the main function of *su* is to contrast a state of affairs to a previous one. *Su*, however, is more than a tense-aspect marker and some of its functions fall within the realm of modality and discourse (see van Minde & Tjia, 2002). In some contexts, “*su* serves to link the objective propositional content of the clause with the realm of presupposition, expectation, affairs, hope, and other subjective, speaker-oriented states” (van Minde & Tjia, 2002, pp. 295-296). An example of *su* is given in (4).

⁵⁵ Examples of *ada* with a (resultative) perfect meaning have also been reported for Singapore Chinese Bazaar Malay by Lee, Ping, & Nomoto (2009, p. 308) and Donohue (2011, p. 418).

⁵⁶ *Su* is the shortened form of *suda*. Although the short form is more common, the full form can also occur pre-verbally to mark aspect.

- (4) *Dong su makang deng balong galap lai*
 3PL PRF eat with not.yet dark also
 ‘They have (already) eaten and it is not dark yet.’ (van Minde, 1997, p. 229)

The preverbal marker *mau/mo* mainly marks mood (volition ‘want’), as shown in (5). However, it can also mark prospective or future aspect, in which case it may be translated as ‘about to (V)’ or ‘going to (V)’ (van Minde, 1997, p. 192), as illustrated in (6).

- (5) *Dia mau pegang duriang tapi dia pung tangang saki*
 3SG want hold durian but 3SG POSS hand sick
 ‘He wants to grab a durian (k.o fruit) but he hurts his hands (due to the thorns).’

- (6) *Ini mau ujang*
 D.PROX want rain
 ‘It’s going to rain (now).’ (van Minde, 1997, p. 193)

Verbal reduplication in Ambon Malay has several functions, which include marking iterative aspect, increased degree and plurality (van Minde, 1997, pp. 119-130). With dynamic verbs, such as activities, reduplication mainly marks iterative aspect, as shown in (7).

- (7) *Tikus toki-toki kue*
 mouse ITER-knock cake
 ‘A mouse (repeatedly) knocks on the cake.’

With stative intransitive verbs (roughly corresponding to adjectives in European languages) reduplication can add the meaning of ‘increased degree’, or ‘plurality’, as illustrated in (8).

- (8) *Tikus ambe tikus talingang, akang panjang-panjang*
 mouse take mouse ear 3SG.N INTENS-long
 ‘The mouse grabs his (own) ears, they are very long.’

4.3.2 Tense-aspect in Dutch

This section gives an overview of the main tense-aspect distinctions encoded in Dutch. This description is meant to orientate the reader in the understanding of possible cross-linguistic influence from Dutch onto heritage Ambon Malay. For the sake of convenience, this section is organized according to function.

In Dutch, present and past tense marking is expressed by means of verbal inflection. Tense inflection and verbal agreement instantiate also the feature of finiteness, which is obligatorily marked in Dutch. The present tense is marked by verbal inflection, and it is used even more extensively than in English; in fact it is the most usual way of expressing an action that is still continuing into the present, as in *Ik woon* (PRESENT) *al tien jaar hier* ‘I have lived (PERFECT) here for ten years’ (Donaldson, 2008, p. 184).

Dutch has two ways to mark past tense, the ‘imperfectum’ and the perfect. The imperfectum is expressed by means of verbal inflection and it is used for describing a series of events in the past (Donaldson, 2008, p. 185). The perfect is encoded by means of a verbal auxiliary and the past participle, as illustrated in (9). The auxiliary is ‘be’ for unaccusative verbs (e.g., *to arrive*) and ‘have’ for unergative verbs (e.g. transitive verbs, such as *to read*, and intransitive verbs, such as *to laugh*; see Broekhuis, Corver & Vos, 2015, p. 30). Dutch perfect is compatible with telic predicates (accomplishments and achievements) and with atelic ones (activities and states).

- (9) *John is gevallen. Max heeft hem geduwd*
 John be.3SG fall.PST.PTCP Max have-3SG 3SG.M.ACC push.PST.PTCP
 ‘John has fallen. Max has pushed him’ (Boogart, 1999, p. 66)

Progressive aspect is marked by the auxiliary verb ‘be’ and the *aan het* + infinitive construction (Boogart, 1999, pp. 167-204; Flecken, 2010, pp. 189-195) as illustrated in (10).

- (10) *Ik ben aan het lezen*
 1SG be.1SG at ART.INDF read-INF
 ‘I am reading’ (Flecken, 2010, p. 82)

Additionally, posture verbs such as *zitten* ‘sit’, *staan* ‘stand’, *liggen* ‘lie’, and the adjective *bezig* ‘busy’, can occur with an infinitive to convey a progressive meaning, as shown in (11).

- (11) *Een man staa-t te viss-en aan de water kant*
 ART.INDF man stand-3SG to fish-INF at ART.DEF water front
 ‘A man is fishing at the waterfront.’ (Flecken, 2010, p. 189)

Progressive aspect in Dutch is mostly restricted to activities and accomplishments, and it is used only marginally with states and achievements. Boogart (1999, p. 195) points out that when an achievement is marked by progressive aspect, “the achievement then gets a non-standard reading in the sense that the situation presented no longer seems to be an achievement at all” (as in *granddad was dying*).

Dutch marks prospective or future aspect with the verb *gaan* ‘go’ + infinitive (Donaldson, 2008, pp. 180-181), as shown in (12).⁵⁷ Unlike Ambon Malay, where the same verb means ‘want’ and prospective aspect, in Dutch *gaan* + infinitive expresses prospective aspect, while *willen* ‘will’ indicates volition and desire.

- (12) *Gaa-t het regen-en*
 go-3SG ART.INDF rain-INF
 ‘Is it going to rain?’ (Donaldson, 2008, p. 181)

Finally, Dutch lacks a specific marker for iterative aspect and uses adverbs, such as *telkens weer* ‘repeatedly’,⁵⁸ to convey this meaning.

⁵⁷ The other auxiliary used in Dutch future expressions is *zullen* ‘shall’. This auxiliary is not considered further here for two reasons. First, it is used only three times in the video descriptions, two of which by the same speaker. Second, because there is no consensus about its status as future auxiliary. Broekhuis et al. (2015, p. 130) analyze it as an epistemic modal verb and explain the future reading normally attributed to this verb as being due to pragmatics.

⁵⁸ Some Dutch verbs ending in *-eren* or *-elen* have an inherent iterative meaning, such as *klapperen* ‘to flap’ as opposed to *klappen* ‘to applaud’ (Philippa et al., 2011). However, the process that led to the formation of these verbs is no longer productive.

4.3.3 Interim summary

The main differences between Ambon Malay and Dutch are summarized in Table 4.1. These involve the expression of tense. Dutch obligatorily marks tense (past/non-past), whereas Ambon Malay does not. Furthermore, Dutch obligatorily encodes finiteness on the verb by means of tense-aspect marking and verbal agreement. In both languages, progressive marking is optional. The overt expression of iterative aspect is also optional in both languages; however, Ambon Malay marks iterativity by means of reduplication, while Dutch lacks a specific marker for iterative aspect.

Table 4.1: Summary of the main tense-aspectual distinction of Ambon Malay and Dutch.

	AMBON MALAY	DUTCH
Present	-	Obligatory, expressed by the stem, stem + t, or the infinitive
(Past) Imperfectum	-	Obligatory, expressed by inflection
(Past) Perfect	Optional, expressed by <i>su</i> , and in some cases by <i>ada</i>	Obligatory, expressed by 'be/have' and the past participle
Progressive	Optional, expressed by <i>ada</i>	Optional, expressed by the <i>aan het</i> + INF construction, and posture verbs plus infinitive
Prospective, Future	Optional, expressed by <i>mau</i> 'want'	Optional, expressed by <i>gaan</i> + INF
Iterative	Optional, expressed by reduplication	Optional, not expressed grammatically

4.4 The study

This section discusses the objective (4.4.1), the research questions (4.4.2) and the methodology (4.4.3) of the present study.

4.4.1 Objective

The purpose of the present study is to examine the effects of Dutch – a language that obligatorily marks past/non-past and finiteness – on the tense-aspect system of heritage Ambon Malay, a language lacking a grammaticalized expression of these distinctions.

4.4.2 Research questions and hypotheses

We have seen in the previous sections that Ambon Malay and Dutch differ with respect to the encoding of tense-aspect distinctions. Now we may ask more broadly how heritage speakers of Ambon Malay, who are bilingual in Dutch and Ambon Malay, but whose dominant language is Dutch, deal with these two sub-systems. The specific research questions that I address in this study are: (i) Does the aspectual system of heritage Ambon Malay feature innovations? (ii) If it does, what are the innovations? Can they be characterized in terms of overextension and/or loss? And (iii) Do external and/or internal factors, such as cross-linguistic influence from the dominant language and input properties, account for the innovations?

Two predictions can be formulated on the basis of previous studies (see Section 4.2). First, heritage speakers will tend to overtly express the contrasts that are grammaticalized in Dutch (e.g., tense, finiteness), when speaking Ambon Malay, as they are used to marking these contrasts in the dominant language. If a category is grammaticalized (and it is therefore used systematically and frequently) in the dominant language, the Functional Convergence Hypothesis and the Conceptual Transfer Hypothesis predict that bilingual speakers will express it also in the heritage language by using linguistic material of the heritage language (Heine & Kuteva, 2005; Matras, 2009; Backus et al., 2011). Second, forms that are semantically and functionally ambiguous will pose a challenge to heritage speakers (Ma, 2006; O’Grady et al., 2011; Laleko & Polinsky, 2013). The difficulty in dealing with these forms may be exacerbated if the forms do not have a readily equivalent in the dominant language. The prediction is that the modal component of expectation of *su* and the polysemy of reduplication contribute to make these two forms non-transparent and hence more difficult to master.

In order to answer these questions and test the predictions, this study investigates how the markers *ada*, *su*, *mau* and reduplication are used by heritage speakers of Ambon Malay when compared to homeland speakers in Ambon,

Indonesia, first generation speakers in the Netherlands and Dutch speakers. It is important to emphasize that aspect marking is optional in Ambon Malay, and that the use of aspect markers is also variable across homeland and first generation speakers. Hence, the quantitative comparison focuses on whether the heritage speakers' use of aspect markers falls within the range of variability observed for homeland speakers and not on whether their use of aspect marker is infelicitous⁵⁹ to a monolingual homeland speaker.

4.4.3 Participants, task, and responses

Four groups of speakers participated in the study: one test group of heritage speakers and three control groups. The test group consists of 32 heritage speakers. The first control group is formed by 27 homeland speakers. The second control group is formed by five first generation speakers of Ambon Malay in the Netherlands. The third control group is formed by ten native speakers of Dutch (see Section 2.1 for more information about the participants).

Every participant performed two tasks: a simultaneous video description task and a video-clip retelling task (see Section 2.2). In the simultaneous video description task, the participants were asked to describe 14 short videos while watching them (see Appendix 1 for the list of videos). In the video-clip retelling task, the participants were asked to watch 29 short video-clips (see Section 2.1 in Appendix 2 for the list of video-clips). They watched two video-clips per time, and then described what they had just seen. In the simultaneous video description task it was not possible to control for how the participant conceptualized the event (did she perceive the situation as an activity or an accomplishment?). The data obtained by means of this task were used to calculate the percentage rate of aspect markers for every speaker on the basis of the total number of predicates. If, for instance, the number of occurrences of *ada* 'EXIST' in the video description is 24 and the total number of predicates is 120, then the rate of *ada* is 0.2 (or 20%). An independent sample *t*-test was conducted to compare the speakers' rates and to determine differences between the homeland group and the heritage group with respect to aspect marking (see Section 2.4).

The problem of controlling for the type of situation described was overcome in the video-clip retelling task because in this task the participants described the video-

⁵⁹ Aspect markers in Ambon Malay are never ungrammatical from a purely syntactic point of view.

clips *after* having watched them. The 29 video-clips were categorized as clearly displaying an activity (5 video-clips), an accomplishment (10 video-clips) or an achievement (14 video-clips).⁶⁰ Unfortunately, the stimulus material did not contain video-clips eliciting states. The activity video-clips displayed events that were ongoing and did not have an endpoint (e.g., swimming). The accomplishment video-clips displayed events that involved duration but had a clear endpoint (e.g., cut off the branch of a tree). The achievement video-clips displayed punctual events with a clear endpoint (e.g., kicking a ball once) (see Section 2.1 in Appendix 2). The data were coded for the presence or absence of *ada*, *su*, *mau* and reduplication in the target descriptions. A mixed effects logistic regression was used to assess the effect of group (homeland, first generation, and heritage) and video-clip type (activity, accomplishment, and achievement) with speaker as a random effect (see Section 2.4 for an explanation of the generalized mixed effects model).

4.5 Results and discussion

This section presents and discusses the results of the experiment. Section 4.5.1 reports the results of the simultaneous video description task, while section 4.5.2 illustrates the results of the video-clip retelling task. In section 4.5.3, I discuss the results of both tasks and propose explanations for the patterns observed.

4.5.1 Results of the simultaneous video description task

Overall, homeland speakers and first generation speakers are more homogeneous in their output, showing a similar rate of *ada*, *su*, *mau* and reduplication, whereas heritage speakers are skewed toward *ada* (see Figure 4.1 on the next page). Given the small sample size of the first generation group, the comparison with this group will be qualitative in nature.

An independent sample t-test revealed that *ada* ‘EXIST’ (black bar) is used significantly more often by heritage speakers ($M=13.75$, $SD=9.32$) than by homeland speakers ($M=3.74$, $SD=2.82$) ($t(37.575)=5.768$, $p<.001$, $r=.68$, equal variances not assumed). In contrast, *su* ‘PRF’ (dark gray bar) is used significantly less frequently in the heritage group ($M=3.43$, $SD=4.43$) than in the homeland group

⁶⁰ The video-clips were intermingled with 39 fillers, for a total of 68 video-clips.

($M=7.22$, $SD=7.48$) ($t(40.693)=-2.308$, $p=.026$, $r=.36$, equal variances not assumed). Reduplication (white bar) also shows a significant decrease in the heritage group ($M=1.46$, $SD=1.96$) compared to the homeland group ($M=4.44$, $SD=2.60$) ($t(47.758)=-4.816$, $p<.001$, $r=.57$, equal variances not assumed). The frequency of the marker *mau* ‘want’ (light gray bar) is approximately the same in the three groups and no statistical difference occurs.

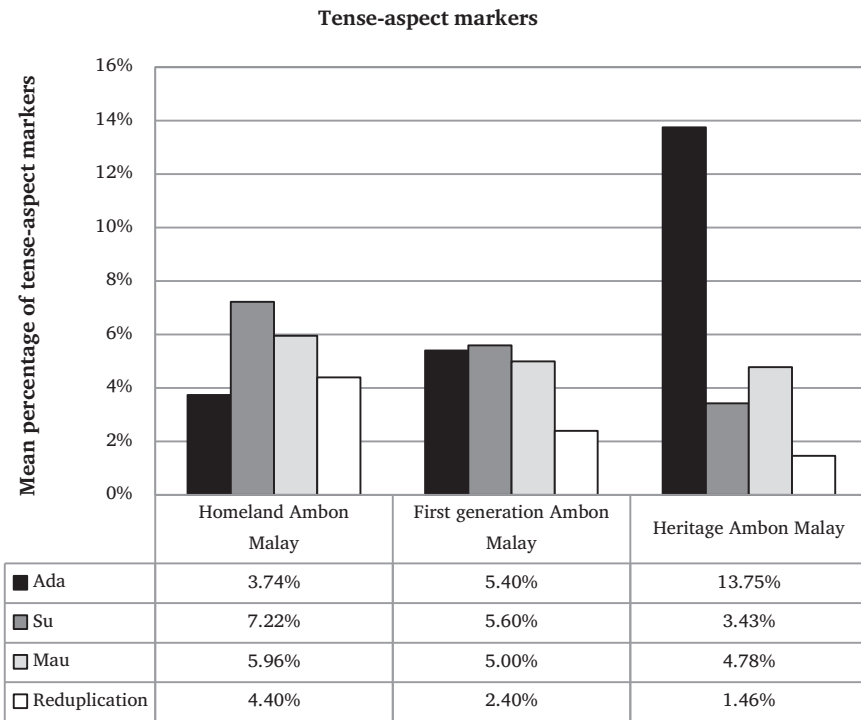


Figure 4.1: Percentage rate of aspect markers in the simultaneous video description of the three Ambon Malay groups.

The Dutch group is extremely homogenous in showing a neat preference for present tense ($M=87.07$, $SD=5.44$), followed by prospective/future aspect expressed by *gaan* ‘go’ + infinitive ($M=5.25$, $SD=2.71$) and by progressive aspect ($M=3.31$, $SD=1.87$). The other tense-aspect markers (perfect and imperfectum) occur considerably less frequently. The data are summarized in Figure 4.2 on the next page.

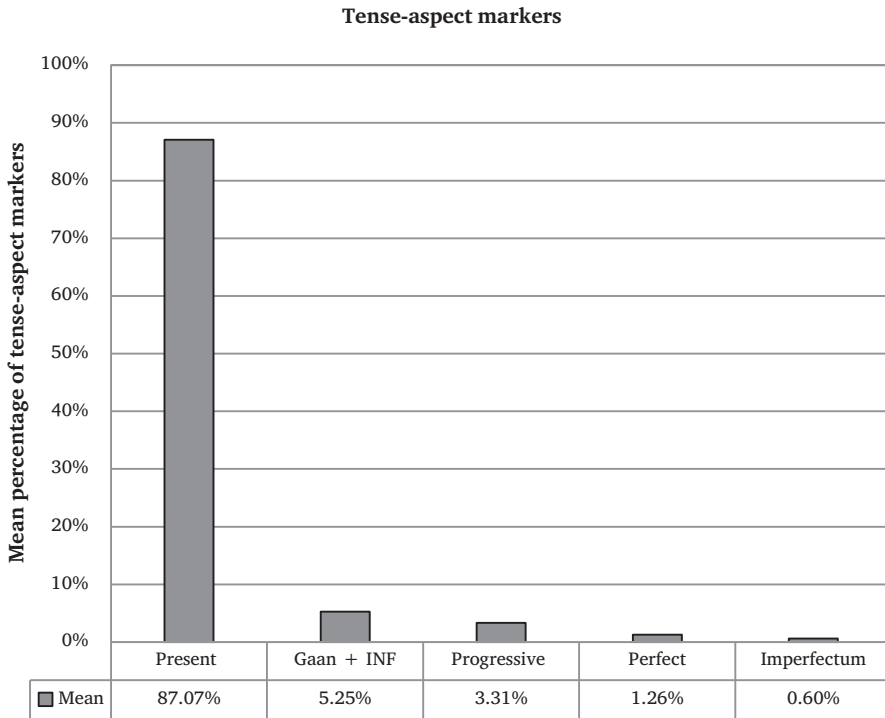


Figure 4.2: Percentage rate of tense-aspect markers in the simultaneous video description of the Dutch group.

In summary, in the simultaneous video description, homeland and first generation speakers use *ada*, *su*, *mau* and reduplication with a similar frequency. In contrast, heritage speakers use *ada* ‘EXIST’ more frequently to the detriment of *su* ‘PRF’ and reduplication. The use of *mau* ‘want’ is the same in the three groups. Dutch speakers mainly use the present tense.

4.5.2 Results of the video-clip retelling task

In the video-clip retelling task, the 74 participants described 29 video-clips. This yielded a total number of 2,146 responses. A total of 140 responses was excluded either because the participant did not describe the target event in the video-clip, or because (s)he described the video-clip by using a non-target predicate. If, for instance, a participant described the achievement video-clip showing a boy who kicks a ball (target: to kick the ball) with an activity predicate (to play with the

ball), the response was excluded. The summary of valid and excluded responses is displayed in Table 4.2.

Table 4.2: Summary of valid and excluded responses in the video-clip retelling task. ACT = activity, ACCO = accomplishment, ACH = achievement.

GROUP	n	RESPONSES	ACT.	ACCO.	ACH.	TOT
Heritage Ambon Malay speakers	32	Valid	155	301	405	861
		Excluded	5	19	43	67
Homeland Ambon Malay speakers	27	Valid	132	265	345	742
		Excluded	3	5	33	41
First generation Ambon Malay speakers	5	Valid	25	49	59	133
		Excluded	0	1	11	12
Dutch speakers	10	Valid	43	95	132	270
		Excluded	7	5	8	20

The data show that the marker *ada* 'EXIST' is more frequent in activities, than in accomplishments and in achievements, see Table 4.3. This trend is the same in the homeland and in the heritage group, but the frequency of *ada* is always higher in the heritage group. Progressive marking is expected with activities and accomplishments because these two situation types have internal duration as they are made of successive phases (see Section 4.1).

Table 4.3: Frequency of *ada* 'EXIST' in activities, accomplishments and achievements. ACT = activity, ACCO = accomplishment, ACH = achievement.

ADA		ACT. (5 CLIPS)	ACCO. (10 CLIPS)	ACH. (14 CLIPS)
Homeland Ambon Malay	Tokens	44	31	22
	% Within group	33.3%	11.7%	6.4%
First generation Ambon Malay	Tokens	12	3	8
	% Within group	48.0%	6.1%	13.6%
Heritage Ambon Malay	Tokens	89	84	64
	% Within group	57.4%	27.9%	15.8%

A generalized linear mixed effects model was used to assess the effect of group and video-clip type on the use of *ada* with speaker as a random effect. The model reveals that heritage speakers use *ada* significantly more than homeland speakers ($\beta=0.65729$, $SE=0.21621$, $p=.002$). There is also an effect of video-clip type. The frequency of *ada* is higher in activity video-clips ($\beta=1.43372$, $SE=0.10874$, $p<.001$) than in achievement video-clips, and, in turn, higher in accomplishment video-clips ($\beta=0.43042$, $SE=0.09518$, $p<.001$) than in achievement video-clips. The overall model is significant ($\chi^2(4)=202.45$, $p<.001$), when compared to a null model with only speaker as a random effect.

The data for *su* 'PRF' shows that, in the homeland group, *su* is more frequent in achievements and in accomplishments than in activities (see Table 4.4). This trend is expected because accomplishments and achievements have a build in endpoint and are therefore more likely to be seen as completed situations in the past (see Section 4.1). It is hard to establish whether the heritage group follows this trend due to the paucity of tokens. What can be said with certainty, is that the frequency of *su* is much lower in the heritage group, where only one token is found, than in the to the homeland group.

Table 4.4: Frequency of *su* 'PRF' in activities, accomplishments and achievements. ACT = activity, ACCO = accomplishment, ACH = achievement.

SU		ACT. (5 clips)	ACCO. (10 clips)	ACH. (14 clips)
Homeland Ambon Malay	Tokens	1	11	16
	% Within Group	0.8%	4.2%	4.6%
First generation Ambon Malay	Tokens	0	1	3
	% Within Group	0.0%	0.2%	5.1%
Heritage Ambon Malay	Tokens	0	0	1
	% Within Group	0.0%	0.0%	0.2%

A generalized linear mixed effects model was used to assess the effect of group and video-clip type on the use of *su* with speaker as a random effect. The model reveals that *su* occurs more often with accomplishments and achievements than with activities ($\beta=-1.2374$, $SE=0.4966$, $p=.01$). There is no difference among the groups. The overall model is significant ($\chi^2(4)=13.075$, $p=.01$), when compared to the null model with only speaker as a random effect.

Mau ‘want’ does not occur in activity video-clips, and it occurs with a low frequency in accomplishment and achievements (see Table 4.5). The low frequency of *mau* is probably a task effect due to the fact that the participants described the videos after having watched them. Descriptions of situations or actions that have already taken place do not trigger the use of prospective aspect. The few tokens of *mau* attested in the dataset are probably instances of *mau* indicating volition rather than prospective aspect.

Table 4.5: Frequency of *mau* ‘want’ in activities, accomplishments and achievements. ACT = activity, ACCO = accomplishment, ACH = achievement.

MAU		ACT. (5 CLIPS)	ACCO. (10 CLIPS)	ACH. (14 CLIPS)
Homeland Ambon	Tokens	0	1	4
	% Within Group	0.0%	0.4%	1.2%
First generation	Tokens	0	1	5
	% Within Group	0.0%	2.0%	8.5%
Heritage Ambon	Tokens	0	3	0
	% Within Group	0.0%	1.0%	0.0%

A generalized linear mixed effects model was used to assess the effect of group and video-clip type on the use of *mau* with speaker as a random effect. No factor has an effect on the use of *mau*, and the model is not significant.

Reduplication is the least frequent of all markers in the three groups; see Table 4.6 on the next page. There is only one token in the heritage group, while it is completely absent in the first generation group. In the homeland group, reduplication seems to be more frequent in the accomplishments. This trend is expected because accomplishments describe a process that extends over a period of time. The general linear model⁶¹ reveals that heritage speakers use less reduplicated verbs than homeland speakers ($\beta = -0.9259$, $SE = 0.3433$, $p = .006$). Furthermore, accomplishments attract more reduplicated verbs than activities and achievements ($\beta = 0.9143$, $SE = 0.3523$, $p = 0.009$).

⁶¹ A generalized mixed effects model with speaker as random effect could not be performed due to the paucity of tokens in the homeland group, a generalized linear model (glm) was used instead.

Table 4.6: Frequency of reduplication in activities, accomplishments and achievements. ACT = activity, ACCO = accomplishment, ACH = achievement.

REDUPLICATION		ACT. (5 CLIPS)	ACCO. (10 CLIPS)	ACH. (14 CLIPS)
Homeland Ambon	Tokens	2	8	1
Malay	% Within Group	1.5%	3.0%	0.3%
First generation	Tokens	0	0	0
Ambon Malay	% Within Group	0.0%	0.0%	0.0%
Heritage Ambon	Tokens	0	1	0
Malay	% Within Group	0.0%	0.3%	0.0%

The results of the Dutch group are summarized together in Table 4.7. Dutch speakers show an overall preference for present tense in all situation types; the progressive is used only with activity predicates, while the imperfectum and the perfect are used mostly with achievement predicates. As mentioned above for Ambon Malay, the task did not elicit prospective aspect, hence the low frequency of *gaan* + infinitive.

Table 4.7: Frequency of Dutch tense-aspect markers in activities, accomplishments and achievements. ACT = activity, ACCO = accomplishment, ACH = achievement.

DUTCH		ACT. (5 CLIPS)	ACCO. (10 CLIPS)	ACH. (14 CLIPS)
Present	Tokens	28	85	118
	% Within Group	65.1%	89.5%	89.4%
<i>Gaan</i> + INF	Tokens	0	1	1
	% Within Group	0.0%	1.1%	0.8%
Progressive	Tokens	10	0	0
	% Within Group	23.3%	0.0%	0.0%
Perfect	Tokens	0	0	2
	% Within Group	0.0%	0.0%	1.5%
Imperfectum	Tokens	0	1	4
	% Within Group	0.0%	1.1%	3.0%

4.5.3 Discussion

The results of the simultaneous video description task show that, when compared to homeland speakers, heritage speaker use *ada* 'EXIST' with a higher frequency, but *su* 'PRF' and reduplication 'ITER, INTENS' with a lower frequency. In contrast, the usage frequency of *mau* 'want' does not differ in the three groups. The video-clip retelling task confirmed these results but it also provided an additional piece of information. The strong association of *ada* with activity predicates tells us that *ada* is still a marker of progressive aspect in the heritage grammar; however, the fact that heritage speakers frequently use *ada* also with accomplishments and achievements tell us that *ada* does not mark progressive aspect only. If, as the Aspect Hypothesis predicts (see Section 4.1), heritage speakers used grammatical markers to duplicate the lexical aspect of the verbs, we would expect *ada* to occur only with activities, and to a lower extent with accomplishments, but definitely not with achievements. The fact that heritage speakers, on the contrary, use *ada* to mark accomplishments and achievements to a greater extent than homeland speakers suggest that when they use *ada* they are not duplicating the lexical aspect, but rather they are encoding something else, possibly tense and/or finiteness. What I would like to argue here is that in heritage Ambon Malay *ada* is changing into a marker of present tense (possibly also encoding finiteness) and of progressive aspect. There are three main reasons for this argument: the empirical data, typology of grammaticalization, and a similar type of change in another Malay variety (Sri Lanka Malay).

The empirical data reported in Section 4.5 show that the overextension of *ada* in heritage Ambon Malay is arguably related to the Dutch present tense, rather than to the progressive or to the perfect. In fact, even though the functions of *ada* and those of the Dutch progressive and perfect overlap, the frequency of *ada* in heritage speakers is much higher than the frequency of either of these two forms in Dutch. In both tasks, Dutch speakers showed an overall preference for present tense, which is indeed the prototypical tense used to describe events (Comrie, 1976, p. 66). Notably, tense and finiteness are highly prominent categories in the dominant language. These categories are, in fact, obligatorily marked on the verb. Even though heritage speakers performed the task in Ambon Malay, the categories of tense and finiteness are highly automatized in their minds, as they encode them when they speak Dutch. Since we know that speakers tend to overtly express the contrasts that are grammaticalized in their languages (see Section 1.3.1.3 and Section 1.4), it is

plausible to assume that heritage speakers have selected the Malay element *ada* to overtly express the tense/finiteness category.

The second argument in support of the reanalysis of *ada* as a present-tense marker comes from grammaticalization theories (Bybee et al., 1994, pp. 127-144; Hengeveld, 2011, p. 590). The semantic change from progressive to present is a well-known grammatical change that involves desemanticization and functional extension (see Section 1.3.1.3). Hengeveld (2011, p. 590) explains that the progressive may be interpreted as present under the following conditions: the progressive describes a situation that occurs at a reference time, when the reference time coincides with the speech moment, the present-tense interpretation is likely to arise. If this interpretation becomes dominant, the form acquires the new present tense meaning (the progressive meaning can either disappear or it can remain available). Desemanticization and functional extension, thus, lead to a higher frequency and distribution of the original progressive marker. I argue that the overgeneralization of *ada* observed in heritage Ambon Malay may be a case of embryonic grammaticalization, whereby the original progressive marker undergoes semantic bleaching and it is expanded to new contexts (achievement predicates). If this development is taking place, this represents a clear instance of contact-induced grammaticalization, whereby a language-internal process (change from progressive to present) is brought about by language contact (Heine & Kuteva, 2005).

Finally, the shift in temporal status and frequency of *ada* is consistent with a change that occurred in another radical heritage variety, namely Sri Lanka Malay. Due to intense contact with Tamil and Sinhala (about 350 years), Sri Lanka Malay has reanalyzed the original progressive marker *ada* as an (almost) obligatory present tense marker (Slomanson, 2006, 2011).⁶² Another potentially significant parallel between these two heritage varieties is that, under the influence of Dravidian languages, Sri Lanka Malay has developed an explicit finite/non-finite contrast that is instantiated by tense morphology (Slomanson, 2006). One could speculate that Ambon Malay in the Netherlands is undergoing a similar development due to the intense contact with Dutch, a language where tense and finiteness are obligatorily encoded on the verb. The difference between Sri Lanka Malay and heritage Ambon Malay is that in Sri Lanka Malay the change has reached completion and the grammatical system presents a neat present/past (and finite/non-finite) contrast, while in heritage Ambon Malay the change is still ongoing and the contrast is not

⁶² Pre-verbal *ada* in non-past contexts is obligatory for most verbs, provided that another functional marker does not appear in pre-verbal position (Slomanson, 2006, p. 143)

clearly expressed. At this point, there are three possible options: (i) *ada* marks the present tense, and there is separate marking for the past tense but I have not elicited it in this study (see below the discussion on *su*); (ii) *ada* encodes the present tense only, and there is not past tense form. *Ada* has the potential to grammaticalize into a present tense marker because the progressive often receives the present tense interpretation (see discussion on grammaticalization theory above); (iii) *ada* only encodes finiteness and is underspecified for tense. The default interpretation for underspecified forms in a neutral context is the present tense (de Hoop, Haverkort, & van den Noort, 2004, p. 1079).

The second finding is that heritage speakers use *su* 'PRF' and reduplication 'ITER, INTENS' with a lower frequency than homeland speakers. I argue that the underuse of these two markers relates to semantic indeterminacy and possibly also to low-acoustic salience (for *su*), and to low-frequency (for reduplication). As we have seen in Section 4.2, non-transparent forms are vulnerable in heritage language grammars. The form *su* is semantically non-transparent because, in addition to the perfect meaning, it carries a modal component of expectation (see Section 4.3.1). In order to correctly use *su*, speakers need to take into account expectations associated with the situation, with the cultural setting or with the common ground of the speaker and the hearer. Furthermore, the form *su* is the least audible of all aspect markers because it contains a low sonority vowel (Gordon, Ghushchyan, McDonnell, Rosenblum, & Shaw, 2012, p. 222) and in fast speech is often reduced to *so* or *s* (van Minde, 1997, p. 228). One may hypothesize that heritage Ambon Malay *su* is losing its modal component while preserving the perfect meaning. Such development has taken place in Sri Lanka Malay where *su* has become a marker of past tense (Slomanson, 2006; Nordhoff, 2009). It is important to note, however, that the elicitation material used in the present study did not elicit sufficient uses of *su* to analyze its function. Further research targeting the use of past forms (e.g., by means of the story re-telling task, see Sánchez, 2004) is necessary to establish the entrenchment level of *su* in heritage Ambon Malay.

Reduplication is indeterminate because it is used to convey several meanings, including iterative aspect, plurality, and intensity (see Section 4.3.1). In addition to being indeterminate, reduplication also seems to be a low frequency form in first generation speakers, those who provided the linguistic input to heritage speakers. This low frequency may either be the result of attrition, or it may be a feature of the language variety first generation speakers brought to the Netherlands in 1950s (see Section 1.5.2.1). Be that as it may, indeterminacy and relatively low frequency in

the input may have acted in a cumulative way hindering the acquisition of this form by heritage speakers. Furthermore, the findings of the present study together with those of Shi (2011) regarding the avoidance of reduplicated verbs in heritage speakers of Mandarin in the Netherlands (see Section 4.2), suggest that reduplication does pose a problem to heritage speakers whose dominant language lacks reduplication. What remains unclear is whether it is the function or the morphological process of reduplication (or both) that are problematic for heritage speakers.

The third, and last, finding concerns *mau* ‘want’. The usage frequency of this marker is the same in the three Ambon Malay groups. A possible explanation for the stability of this form is that *mau* also functions as a modal auxiliary indicating volition, and as such it occurs relatively frequently in the input heritage speakers are and have been exposed to. However, additional data are needed to establish with certainty whether the distribution of this marker is exactly the same in Ambon and in the Netherlands. Interestingly, the frequency of *mau* in the Ambon Malay groups resembles the frequency of *gaan* + infinitive in Dutch. It is not possible at the moment to say whether this is a coincidence or whether the stability of *mau* also depends on the fact that this form is identified with the ‘corresponding’ Dutch equivalent. At the moment, I can only speculate that the stability in the frequency of *mau* depends partially on its function as volition modal and partially on its semantic equivalence with the Dutch *gaan* + infinitive structure.

Before turning to the conclusions, there is one last point that it is worth mentioning. Going back to Sri Lanka Malay, one may ask whether it is possible that heritage Ambon Malay *mau* develops into a marker of non-finiteness, as it did in the Malay variety of Sri Lanka. Sri Lanka Malay, in fact, has developed a finite/non-finite opposition with tense morphology (such as *ada*) marking finiteness and *mə* (>*mau*) marking non-finiteness (Slomanson, 2006). If heritage Ambon Malay will be spoken in the Netherlands for enough time (350 years, as it is the case for Sri Lanka Malay), we may expect it to develop a neat finite/non-finite contrast, with either *mau* or bare verbal forms indicating non-finiteness and tense-aspect markers instantiating finiteness.

To sum up, everything that recurs in someone’s language experience is hypothesized by Backus et al. (2011) to be entrenched in that speaker’s mind. In heritage speakers, patterns belonging to both languages will be entrenched in their minds. Like words and structures, grammatical categories can be expected to be replicated from one language to another. This is exactly what we have observed in

heritage Ambon Malay, where the indigenous progressive aspect marker *ada* has undergone semantic and functional extension and it is now used by heritage speakers to express a category borrowed from the dominant language, namely the (present) tense category. Forms, such as *su* and reduplication, which have a non-transparent form-meaning mapping, are infrequent and have low acoustic salience are difficult for heritage speakers, who tend to avoid them. Furthermore, the fact that these forms lack a readily equivalent in the dominant language may contribute to the decline of their entrenchment in the heritage speakers' repertoire.

4.6 Conclusions

This study examined the effects of Dutch – a language that obligatorily marks past/non-past and finiteness – on the tense-aspect system of heritage Ambon Malay, a language lacking a grammaticalized expression of these distinctions. The analysis of two types of data provided by the same speakers revealed that heritage Ambon Malay is undergoing two types of contact-induced changes: overextension of the progressive marker *ada*, (ii) underuse of the perfect marker *su* and verbal reduplication.

The first innovation concerns the overextension of the progressive marker *ada*. The results of the two video description tasks showed that this tense-aspect marker is used more frequently and more systematically by heritage speakers. On the basis of grammaticalization theory and of similar change that occurred in another heritage Malay variety (Sri Lanka Malay), I argue that the shift in temporal status and frequency of *ada* is consistent with an embryonic process of contact-induced grammaticalization from progressive to present tense. In other words, in the grammar of heritage speakers *ada* is used to convey the present tense function (and possibly also to mark the finiteness contrast), while retaining its (original) progressive function as well.

The second innovation concerns the decrease in usage frequency of the perfect marker *su* and of verbal reduplication, a change that is arguably related to language internal factors. I argue that the underuse of these two markers in the two video description and the video-re-telling tasks relates to semantic indeterminacy and possibly also to low-acoustic salience (for *su*), and to low-frequency (for reduplication). The non-transparent form-meaning mapping of *su* and reduplication, together with low frequency and low salience, renders the conditions for their usage

obscure to heritage speakers, who in turn avoid them. Nevertheless, further research targeting these forms is necessary as the tasks used in the present study did not provide an ideal context for the investigation of past and iterativity.

Finally, the frequency of the marker *mau* 'want' seems to be stable in heritage Ambon Malay, although additional data are needed to establish with certainty whether the distribution of this marker is exactly the same in Ambon and in the Netherlands.

CHAPTER 5

Give-constructions in heritage Ambon Malay⁶³

5.1 Introduction

In language contact situations, grammatical areas which allow variable syntax are often susceptible to change. This is the main tenet of the Alternation Hypothesis and the Vulnerability Hypothesis, which consider variability (alternation of structures) as the locus for cross-linguistic influence (see Section 1.4.1). This has been shown for domains like possessive encoding in Moroccan Arabic (Boumans, 2006), subordinate clauses in Turkish (Onar Valk & Backus, 2013; Onar Valk, 2015), or object marking in Spanish (Montrul & Bowles, 2009). In many languages, the encoding of *give*-events also constitutes such a variable syntactic domain, as the Recipient-like argument (R) and the displaced Theme (T) argument⁶⁴ involved in such events may be ordered in various ways, and receive different encodings — a variation commonly referred to in English as the ‘dative alternation/shift’ (Bresnan, Cueni, Nikitina, & Baayen, 2007; Coleman, 2009; Broekhuis et al., 2015). The terms ‘dative alternation’ or ‘dative shift’ link together the ‘Double Object (DO) construction’, where R and T occur in a fixed order, and are not distinguished by any overt marking (*John gave Mary a book*), and the ‘Dative construction’, also known as ‘Prepositional Object (PO) construction’ where R is differentiated from T by being part of a prepositional phrase (*John gave a book to Mary*). In this chapter, we avoid using the notion ‘dative’, as Ambon Malay does not distinguish the dative case, and we refer to the two constructions as ‘Double Object’ (DO) and ‘Prepositional Object’ (PO) constructions.⁶⁵ In a canonical *give*-event, R and T do not have the same status

⁶³ This chapter is based on Moro, F. R., & Klamer, M. (2015). *Give*-constructions in heritage Ambon Malay in the Netherlands. *Journal of Language Contact*, 8, 263-298.

⁶⁴ We refer to these participants in *give*-events with the capitals R and T, following conventions used in linguistic typology (see Dryer, 2007, p. 254, Malchukov, Haspelmath, & Comrie, 2010, p. 1, Haspelmath, 2011, p. 540).

⁶⁵ Other terms used in the literature to refer to the two constructions include ‘indirective’ versus ‘double object’ construction (Malchukov et al., 2010, p. 18), ‘prepositional dative’ versus ‘double object’ construction (Bresnan et al., 2007, p. 70; Coleman & Bernolet, 2012, p. 88).

in information structure, and this difference is reflected in how they are expressed (cf. Polinsky, 1998; Bresnan et al., 2007). R is the participant that is presupposed to exist independently of the event, which is about the transfer of T. T is the more focus-like, and R constitutes the more topic-like entity. As the previously activated ('known') topic that is accessible to both hearer and speaker, the R is more easily shortened or deleted than the T. In contrast, the T is the element that is typically the new information which the speaker wants to convey.

The variability of argument encoding found in *give*-constructions makes it an interesting domain of inquiry, not only from a monolingual perspective (Colleman, 2009; Theijssen, 2012, among others), but even more so from a language contact perspective (see Schoonbaert et al., 2007 for Dutch-English bilinguals; Yip & Matthews, 2007 for Cantonese-English bilinguals; Şahin, 2015 for Papiamentto-Dutch bilinguals). For studies of language contact, the main interest of the *give*-constructions lies in the issue of what happens when patterns of variable argument encoding that exist in two languages are combined in the same bilingual speaker. Preliminary results of research investigating such combinations indicate that the expression of *give*-events is indeed a vulnerable domain which is subject to cross-linguistic influence. For example, Şahin (2015) has found significant cross-linguistic effects in the production of *give*-constructions by Papiamentto-Dutch bilinguals. Similarly, Irizarri van Suchtelen (2014) has found significant changes in the dative constructions produced by heritage speakers of Spanish (all Dutch-Spanish bilinguals) compared to those produced by their monolinguals peers in Chile.

This chapter seeks to contribute to our understanding of cross-linguistic effects in the production of *give*-constructions by studying another heritage language: the Ambon Malay variety as spoken by heritage speakers in the Netherlands. Heritage speakers of Ambon Malay are second or third generation immigrants to the Netherlands who grew up as simultaneous or sequential Dutch-Ambon Malay bilinguals, with Dutch as the dominant language. The central question addressed in this chapter is: Has there been a restructuring of the *give*-construction in the heritage Ambon Malay of these bilingual speakers, as compared to the Ambon Malay spoken in the homeland? And if restructuring of *give*-constructions in heritage Ambon Malay did take place, what did the change involve?

By comparing *give*-constructions used by the heritage speakers with those used by the homeland and first generation speakers, we find that heritage Ambon Malay has indeed been significantly restructured. What is particularly interesting is that this restructuring is not manifested as a categorical change in the grammar of

heritage speakers, but it rather manifests itself as a significant change in the frequency with which certain constructions that exist in the homeland variety occur in the heritage language (see Section 1.3.1.1). In other words, heritage Ambon Malay is ‘restructuring by changing frequency’. We argue that this is caused by a combination of factors: it is partly due to contact with Dutch, the dominant language, and partly due to universal principles of language development in the context of language disuse (see Section 1.3.5). In addition, we relate the extent of the attested patterns to the amount of exposure that individual heritage speakers had to Ambon Malay in the course of their lifetime.

This chapter is structured as follows. Section 5.2 describes how *give*-events are expressed in the languages of the bilingual heritage speakers: Ambon Malay (section 5.2.1) and Dutch (section 5.2.2). Section 5.3 illustrates the design of the present study. The results are presented and discussed in section 5.4. Section 5.5 summarizes the conclusions.

5.2 *Give-constructions in Ambon Malay and Dutch*

This section presents a descriptive overview of the various ways in which *give*-events are expressed in Ambon Malay and Dutch, the languages that are combined and used by the bilingual heritage speaker of Ambon Malay. We define *give*-events as involving verbs with a meaning of ‘transfer’ or ‘caused possession’, that are translated as ‘give’ or ‘show’ in English, and have three arguments: an Agent-like argument (A), a Recipient-like argument (R), and a Theme argument (R).⁶⁶ This section focuses on describing those structural features that are relevant for answering the question how the *give*-constructions in the heritage language have been restructured as compared to the language of the homeland. The examples presented in this section and elsewhere in this chapter are all from the dataset collected for this dissertation (see Section 2.2 and Section 2.3).

⁶⁶ Verbs for ‘give’ and ‘show’ are among the most typical and frequently found ditransitive verbs cross-linguistically: “It is striking that when a language has a closed class of ditransitive verbs, the same lexemes tend to recur in this class in language after language, most frequently verbs like ‘give’, ‘show’, ‘teach’; sometimes also ‘tell’, ‘send’ and ‘ask’.” (Malchukov et al., 2010, p. 50).

5.2.1 Give-constructions in Ambon Malay

In the domain of *give*-events, Ambon Malay allows five basic constructions. All these five constructions are attested in both homeland and heritage Ambon Malay.

The first construction is the Prepositional Object (PO) construction, in which the Theme (T) is a bare NP, while the Recipient (R) is encoded in a prepositional phrase, as illustrated in (1). As the examples show, different prepositions may be used to introduce the R. The preposition *par* ‘for, to’ (1a) is a lexeme that is often used in PO constructions in Ambon Malay (van Minde, 1997, p. 76). Not being found in any other Malay variety, it can be considered a unique and typical feature of Ambon Malay. Apart from *par* ‘for, to’, Ambon Malay may introduce the R with other prepositions that have similar or identical meanings, such as *buat* ‘for, to’, illustrated in (1b). Other prepositions that may be used are *for* ‘for, to’, *ka* ‘to’, *kepada* ‘to’, *untuk* ‘for’, and *sama* ‘to, with’ (cf. Paauw, 2008, p. 122). *Kepada* and *untuk* are recent loans from Standard Indonesian.

- | | | | | | | | | | |
|-----|----|--------------------------------------|-------------|-------------|---------------|-------------|------------|------------------|-------------|
| | | | V | T | R | | | | |
| (1) | a. | <i>Cowo</i> | <i>kasi</i> | <i>tas</i> | <i>par</i> | <i>cewe</i> | | | |
| | | boy | give | bag | to | girl | | | |
| | | ‘A boy gives a bag to a girl.’ | | | | | | | |
| | | | | | | | | | |
| | | | V | T | R | | | | |
| | b. | <i>Dia</i> | <i>mau</i> | <i>kasi</i> | <i>kemeja</i> | <i>buat</i> | <i>ana</i> | <i>laki-laki</i> | <i>satu</i> |
| | | 3SG | want | give | shirt | to | child | male | one |
| | | ‘He wants to give a shirt to a boy.’ | | | | | | | |

The variable choice of prepositions in PO constructions is further discussed in Section 5.4.2 below, where we see that homeland and heritage speakers use different prepositions in PO constructions.

The second construction used in Ambon Malay *give*-expressions is the Double Object (DO) construction, in which the T and the R are both bare NPs, as illustrated in example (2). In Ambon Malay, PO constructions like those illustrated in (1a-b) are more frequent than DO constructions like (2) (van Minde, 1997, p. 223).⁶⁷

⁶⁷ Haspelmath, Michaelis and the APiCS Consortium (2013) represent Ambon Malay with a pie chart that has 75% PO and 25% DO, but it remains unclear where these percentages come from. The reference they provide for the percentages is van Minde (1997, p. 221), but no

introduced in the first part of the sentence as argument of the verb *pegang* ‘hold’,⁶⁸ while the R *dia pung tamang* ‘his friend’ is introduced with the verb *kasi* ‘give’. In the ‘two predicate constructions’ in the dataset, the first predicate is usually *pegang* ‘hold’, as in (4).

						V ₁	T
(4)	<i>Yang</i>	<i>cowo</i>	<i>satu</i>	<i>ni,</i>	<i>dia</i>	<i>pegang</i>	<i>tas</i>
	REL	boy	one	D.PROX	3SG	hold	bag
			V ₂	R			
	<i>tarus</i>	<i>dia</i>	<i>kasi</i>	<i>par</i>	<i>dia</i>	<i>pung</i>	<i>tamang</i>
	next	3SG	give	to	3SG	POSS	friend
	‘This boy here, he holds a bag, and then he gives (it) to his friend.’						

The first predicate in a ‘two predicate construction’ may also be a different verb, as illustrated in (5). In (5), *buka* ‘open’ introduces the T *buku* ‘book’, while R *laki-laki satu* ‘one boy’ is the argument of the second verb *kasi* ‘give’.⁶⁹

						V ₁	T
(5)	<i>Ada</i>	<i>laki-laki</i>	<i>satu</i>	<i>ni,</i>	<i>dia</i>	<i>buka</i>	<i>buku,</i>
	EXIST	male	one	D.PROX	3SG	open	book
			V ₂	R			
	<i>la</i>	<i>dia</i>	<i>kasi</i>	<i>laki-laki</i>	<i>satu</i>		
	then	3SG	give	male	one		
	‘There is this boy here, he opens a book, and then he gives (it) (to) another boy.’						

⁶⁸ The data contains one instance where the T is introduced as part of the phrase denoting the Agent:

(i)	<i>Pace</i>	<i>dengan</i>	<i>krusli,</i>	<i>mau</i>	<i>kasi</i>	<i>for</i>	<i>itu</i>	<i>mace</i>
	man	with	muesli	want	give	to	D.DIST	girl
	‘A man with (a box of) muesli, (he) wants to give (it) to the girl.’							

This construction patterns with the ‘two predicate construction’ in that the T is introduced first, and is shared (but not repeated) with the second predicate, which introduces the R.

⁶⁹ In ‘two predicate constructions’, R is most often encoded as part of a prepositional phrase, as illustrated in (4). However, our dataset has also instances where R is expressed as a bare NP, as shown in (5).

The ‘two predicate construction’ may consist of two clauses that are connected with a conjunction: *tarus* ‘and then, next’ in (4), *la/lalu* ‘and then’ in (5) are often used, as well as *jadi* ‘so’, *langsung* ‘and then immediately’, or *dan* ‘and’. However, the clauses can also be simply juxtaposed without an overt linker, being separated with a pause, as shown in (6).

		V ₁		T			
(6)	<i>Antua</i>	<i>pegang</i>		<i>kemeja,</i>			
	3SG.FML	hold		shirt			
		V ₂	R				
	<i>mau</i> ⁷⁰	<i>kasi</i>	<i>par</i>	<i>antua</i>	<i>pung</i>	<i>ana</i>	<i>itu</i>
	want	give	to	3SG.FML	POSS	child	D.DIST
	‘He holds a shirt, (he) wants to give (it) to his child.’						

‘Two predicate constructions’ such as those found in Ambon Malay are a typical feature of the Malay varieties that are spoken in eastern Indonesia. In these varieties, complex events tend to be expressed through two or more consecutive verbal predicates (referred to as ‘serial verb constructions’⁷¹ in Paauw, 2008, pp. 232-236; see also Moro, 2014; and Chapter 6 in this dissertation).

The fifth construction attested in the dataset is the ‘argument fronting construction’. In such constructions, one of the two object arguments, mostly T, is fronted to precede A, while R is part of a prepositional phrase following the verb. This is illustrated in (7), where the T *buku* ‘book’ is fronted to the position preceding the A *dia* ‘3SG’.⁷²

⁷⁰ We consider auxiliary verbs such as *mau* ‘want’ (van Minde, 1997, p. 192) to form a complex predicate with the main verb which they precede. In (6), *mau kasi* ‘wants to give’ thus counts as one (complex) predicate.

⁷¹ We do not use this term here, as the ‘two predicate constructions’ in our data include various types of structures along the cline from ‘serial verbs’ to ‘asyndetic parataxis’ to ‘conjoined clauses’, as illustrated above.

⁷² In the dataset, there are nine instances where T is fronted, against one instance where R is fronted.

- (7) *Jadi ini ada se-bua buku,*
 So D.PROX EXIST one-CLF book
 T V R
buku dia kasi tunju par dia pung tamang
 book 3SG give point to 3SG POSS friend
 ‘So here there is a book, (the) book he shows to his friend.’

In sum, the Ambon Malay dataset contains five types of constructions that express *give*-events: (i) the prepositional object (PO) construction, (ii) the double object (DO) construction, (iii) the Recipient omission construction, (iv) the ‘two predicate construction’ and (v) the ‘argument fronting construction’. These constructions are used by all speakers of Ambon Malay, both homeland and heritage. The difference between homeland and heritage *give*-expressions does not lie in the type of constructions used, but rather involves a change in the frequency of certain constructions; we return to this in section 5.4.1. In addition, the expression of *give*-events in Ambon Malay shows variation in the choice of preposition that heads the prepositional phrase in the PO construction, a topic we return to in section 5.4.2.

5.2.2 *Give*-constructions in Dutch

Being the dominant language of Ambon Malay heritage speakers, it is likely that Dutch has influenced the way in which *give*-events are expressed in heritage Ambon Malay. This section presents a summary of the type of *give*-constructions used by native speakers of standard Dutch (see Section 2.1.2.4).

The Dutch *give*-constructions allow alternations that involve a PO construction, as in (8), and a DO construction, as in (9). The preposition used in the prepositional phrase is always *aan* ‘to’.

- (8) *Een man geef-t zijn tas*
 ART.INDF man give-3SG 3SG.POSS.M bag
 R
aan een andere man
 to ART.INDF other man
 ‘A man gives his bag to another man.’

- | | | | | | | | |
|-----|---|---------------|------------|---------------|-----------|---------------|------------|
| | | | V | R | | | |
| (9) | <i>De</i> | <i>ene</i> | <i>man</i> | <i>geef-t</i> | <i>de</i> | <i>andere</i> | <i>man</i> |
| | ART.DEF | one | man | give-3SG | ART.DEF | other | man |
| | T | | | | | | |
| | <i>een</i> | <i>rugzak</i> | | | | | |
| | ART.INDF | backpack | | | | | |
| | ‘The one man gives the other man a backpack.’ | | | | | | |

The R argument may be omitted, as shown in (10). Dutch *give*-events are always expressed with a single verbal predicate.

- | | | | | | | |
|------|-----------------------|------------|--------------|------------|-------------|--------------|
| | | | V | T | | |
| (10) | <i>Een</i> | <i>man</i> | <i>laa-t</i> | <i>een</i> | <i>boek</i> | <i>zie-n</i> |
| | ART.INDF | man | let-3SG | ART.INDF | book | see-INF |
| | ‘A man shows a book.’ | | | | | |

In standard Dutch, different types of factors determine the choice between a PO or a DO construction (cf. Broekhuis et al., 2015, pp. 517-525, and references therein). Semantics plays a role: a PO construction is used when the referent of T undergoes a change of location, whereas a DO construction is used when the referent of R is expected to become the possessor of T. Another important factor in the choice for a DO or PO construction is the size of the object noun phrase: shorter noun phrases are often bare, and feature in DO constructions; while longer noun phrases are often part of prepositional phrases, and are typically placed at the end of the utterance according to the “principle of end weight” (Wasow, 2002; Bresnan et al., 2007). Moreover, the animacy of the referents also plays a role in the DO/PO alternation: canonically, R is animate and T inanimate, so that an inanimate R in a DO construction is less felicitous (*Peter gaf Jan/?de bibliotheek het boek* ‘Peter gave John/?the library the book’). Furthermore, the information packaging of the clause is relevant for the alternation: in a canonical *give*-event, if the R is given and T is new information, the DO construction is used; if T is given and R is new information, then a PO construction is used. And finally, different lexemes of transfer verbs show different biases for one construction over the other. For instance, the analytic causative *laten zien* ‘to show’ (lit. ‘to let see’) clearly prefers a PO construction, as in (11), while the verb *tonen* ‘to show’ does not have such a clear preference.

- | | | | | | |
|------|----------------------------------|------------|---------------|------------|------------------|
| | | V | T | | |
| (11) | <i>Een</i> | <i>man</i> | <i>laa-t</i> | <i>een</i> | <i>jas zie-n</i> |
| | ART.INDF | man | let-3SG | ART.INDF | jacket see-INF |
| | R | | | | |
| | <i>aan</i> | <i>een</i> | <i>jongen</i> | | |
| | to | ART.INDF | boy | | |
| | 'A man shows a jacket to a boy.' | | | | |

In sum, Dutch objects are more likely to appear in DO constructions than in PO constructions, when R is pronominal, definite, presuppositional, short, and animate; and when T is non-pronominal, indefinite, in focus, and long (Bresnan et al., 2007; Broekhuis et al., 2015, p. 524).

But this is not all there is to say about the DO/PO alternation: we know that the genre and context of the utterances also plays an important role. In particular, experimental and corpus data often show different, sometimes opposing, tendencies in frequencies of certain constructions. In the domain of Dutch *give*-expressions, it is reported that the verbs *geven* 'give' and *tonen* 'show' as used in corpora are skewed towards the DO construction (Colleman, 2006, 2009; Colleman & Bernolet, 2012), while the same verbs show a strong preference for PO over DO constructions in de-contextualized experiments (Colleman & Bernolet, 2012, pp. 96, 104).⁷³ In other words, the choice for a DO or PO construction is also context-dependent.

In the Dutch data set that we analyzed, we see significantly more PO constructions than DO constructions (see Figure 5.1 below). We explain this preference for PO constructions as a consequence of the de-contextualized setting of the experiment (described in Section 5.3.3). Such a setting involves an R that is not presuppositional, and it typically requires the R to be expressed with a lexical nominal constituent that is not pronominal, and not short. In fact, many of the Dutch *give*-expressions in the data have an R that is quite long, as for example 'a girl on his right hand side' in (12).

⁷³ In the picture description task reported in Colleman & Bernolet (2012, p. 96), the elicited sentences showed a preference for PO (54.1%) over DO (13.9%).

- (12)
- | | | | | | | | |
|------------|--------------------|------------------|---------------------|-------------|--------------------|---|-------------|
| | | | | | | T | |
| <i>Ik</i> | <i>zie</i> | <i>een</i> | <i>man</i> | <i>die</i> | <i>een</i> | | <i>paar</i> |
| 1SG | see.1SG | ART.INDF | man | REL.PRO | ART.INDF | | pair |
| | | | | V | | | |
| | <i>gestippelde</i> | <i>schoen-en</i> | <i>overhandig-t</i> | | | | |
| | dotted | shoe-PL | hand.over-3SG | | | | |
| | R | | | | | | |
| <i>aan</i> | <i>een</i> | <i>meisje</i> | <i>aan</i> | <i>zijn</i> | <i>rechterkant</i> | | |
| to | ART.INDF | girl | to | 3SG.POSS.M | right.hand.side | | |
- ‘I see a man who hands over a pair of dotted shoes to a girl on his right hand side.’

In all cases, the R is expressed with at least two words, as in (12) above, and in (13) and in (14), and most of the R’s are indefinite noun phrases, as in (12) and (13), though definite ones also occur, as in (14).⁷⁴

- (13)
- | | | | | | | |
|------------|------------|---------------|------------|------------|---|--|
| | | | V | | T | |
| <i>Een</i> | <i>man</i> | <i>geef-t</i> | <i>een</i> | <i>tas</i> | | |
| ART.INDF | man | give-3SG | ART.INDF | bag | | |
| | R | | | | | |
| <i>aan</i> | <i>een</i> | <i>vrouw</i> | | | | |
| to | ART.INDF | woman | | | | |
- ‘A man gives a bag to a woman.’
- (14)
- | | | | | | | |
|------------|---------------|----------------|-------------|------------------|-------------|-------------------|
| <i>Ik</i> | <i>zie</i> | <i>een</i> | <i>man</i> | <i>met</i> | <i>twee</i> | <i>kind-eren,</i> |
| 1SG | see.1SG | ART.INDF | man | with | two | child-PL |
| | | | | V | | T |
| <i>hij</i> | <i>geef-t</i> | <i>een</i> | <i>paar</i> | <i>schoen-en</i> | | |
| 3SG.M | give-3SG | ART.INDF | pair | shoe-PL | | |
| | R | | | | | |
| <i>aan</i> | <i>het</i> | <i>rechter</i> | <i>kind</i> | | | |
| to | ART.DEF | right.hand | child | | | |
- ‘I see a man with two children, he gives a pair of shoes to the child on the right.’

⁷⁴ The responses had 28 indefinite Rs and 15 definite Rs.

Finally, Dutch allows one of the arguments, usually the T, to be fronted to the first position of a main clause, as in (15).⁷⁵

	T					V	A
(15)	<i>De</i>	<i>tas</i>	<i>in</i>	<i>zijn</i>	<i>rechterhand</i>	<i>overhandig-t</i>	<i>hij</i>
	ART.DEF	bag	in	3SG.POSS	right.hand	hand.over-3SG	3SG.M
	R						
	<i>aan</i>	<i>de</i>	<i>man</i>	<i>die</i>	<i>tegenover</i>	<i>hem</i>	<i>staa-t</i>
	to	ART.DEF	man	REL.PRO	opposite	3SG.M.ACC	stand-3SG
	‘The bag in his right hand he hands over to the man who stands in front of him.’						

In sum, Dutch has four constructions to express *give*-events: (i) the prepositional object (PO) construction, (ii) the double object (DO) construction, and (iii) the ‘Recipient omission construction’ and (iv) the ‘argument fronting construction’. The frequency of the various constructions may differ depending on factors such as context and genre (natural corpus data versus elicited experimental data), or the category of T and R (noun or pronoun).

5.2.3 Interim summary

We have seen that in the domain of *give*-expressions, there are many similarities between Ambon Malay and Dutch. Both languages use the PO construction, the DO construction, the ‘R omission’ construction, and the possibility to front T. The ‘two predicate construction’ that is used in Ambon Malay is never used in Dutch. Finally, Ambon Malay allows six different prepositions to encode R, while Dutch only allows one.

⁷⁵ The dataset also contained a response like (i), where the T and the R are part of a subject relative clause:

(i) *Een man die een boek laa-t zie-n aan een andere man*
 ART.INDF man REL.PRO ART.INDF book let-3SG see-INF to ART.INDF other man
 ‘(I see) a man who shows a book to another man.’

Even though T precedes the verb here, we do not count this as an instance of ‘T fronting’, as it reflects the basic constituent order of Dutch subordinate clauses, which is always object-verb.

5.3 The study

This section discusses the objective (5.3.1), the research questions (5.3.2) and the methodology (5.3.3) of the present study.

5.3.1 Objective

The purpose of the present study is to examine how the *give*-construction in the Ambon Malay variety as spoken by heritage speakers in the Netherlands has been restructured as compared to the language of the homeland.

5.3.2 Research questions and hypotheses

The study focuses on the following two research questions: (i) Is the frequency of the attested constructions the same among the heritage speakers and the homeland speakers? (ii) Do the heritage speakers and the homeland speakers select the same preposition to encode R in the PO construction? To answer these questions, we first compare the patterns identified in heritage Ambon Malay with patterns in the Ambon Malay variety spoken by homeland speakers. Then we compare the patterns of heritage speakers to those of first generation speakers in the Netherlands (who are late bilinguals), and to those of Dutch speakers (with no knowledge of Ambon Malay). Given the small sample size of these last two groups, these latter comparisons are mostly qualitative in nature.

The typological similarities and differences between the Ambon Malay and Dutch *give*-expressions described in Section 5.2 allow us to make the following predictions. The first prediction is that heritage Ambon Malay diverges from homeland Ambon Malay with respect to the DO/PO alternation. Although both Dutch and Ambon Malay allow a choice between PO constructions and DO constructions, we saw in Section 5.2 that there are different preferential tendencies: Ambon Malay always prefers PO, while Dutch has a bias for DO in corpus data, and for PO in de-contextualized elicited data. It has been demonstrated (Boumans, 2006; Moro, 2014; Onar Valk, 2015, among others) that heritage speakers tend to adapt the frequency of a construction in the heritage language to the frequency of the ‘corresponding’ construction in their dominant language (see Section 1.3.1.1). In this case, this means that the frequency of PO and DO constructions in heritage Ambon Malay will

be adapted to the Dutch frequencies, and thus they would be different from that of homeland Ambon Malay.

The second prediction is that the ‘two predicate construction’ is problematic for heritage speakers because this construction is not found in Dutch, their dominant language. The choice between a ‘two predicate construction’ and a construction where only a single predicate is used relates to principles of information organization, and these principles are clearly different in Ambon Malay and Dutch. In Ambon Malay, the description of an event is often segmented into two or more predicates (referred to as a ‘serial verb construction’, Tjia, 1997; van Staden & Reesink, 2008; Paauw, 2008), a pattern that we also saw present in the *give*-expressions. In Dutch, however, *give*-expressions involve just a single verbal predicate. We know that principles of information organization are susceptible to considerable transfer effects in bilingual speakers (see Slobin, 1991; Carroll & von Stutterheim, 2003; and Section 1.4.2), so we may expect heritage speakers to follow the Dutch principles and to use constructions with one single predicate more often than homeland speakers.

Finally, given the many different prepositions allowed in the PO construction in Ambon Malay and the different historical trajectories of these prepositions (see Section 1.5 and Section 1.6.2.7), we expect that the choice of prepositions used by the heritage speakers will be different from those used in the homeland variety.

5.3.3 Participants, task, and responses

Four groups of speakers participated in the study: one test group of heritage speakers and three control groups. The test group consists of 32 heritage speakers. The first control group is formed by 27 homelands speakers. The second control group is formed by six first generation speakers of Ambon Malay in the Netherlands. The third control group is formed by ten native speakers of Dutch (see Section 2.1 for more information about the participants and data collection).

The task that the participants were asked to complete was to give an oral description of a series of six video-clips that were shown to them on a laptop screen (see Section 2.2). Three of these video-clips depict a person handing over an object (a pair of shoes or a bag) to another person, and three of them depict a person who is showing an object (a book or a jacket) to another person (see Section 2.2 in Appendix 2). The six video-clips were intermingled with 62 fillers (see Section 2.4 in Appendix 2).

By using these materials, we controlled for three major factors that play a role in the choice between various available *give*-constructions: animacy, discourse accessibility and register (cf. Bresnan et al. 2007, and references cited there). In the experiment, all participants described the same video-clips, so the animacy values of the arguments are kept constant: A and R are both animate (humans), while T is inanimate (e. g. bag, shoes, or book). Additionally, the video-clips are canonical for a non-abstract use of ‘give’ and ‘show’. Second, the discourse accessibility is kept constant for all the arguments: none of them was previously introduced, as utterances elicited as responses to a video-clip in a test situation always lack a natural discourse context. Third, the register is kept constant, since the data were all oral retellings of video-clips. Any variation we find in the expression of ‘give’-events we can thus interpret as a consequence of factors other than animacy value, discourse accessibility or register.

Every participant produced six responses. All the responses were transcribed and entered into a separate database in Excel (see Section 2.4). Not all responses were included in the analysis, as laid out in Table 5.1. The criterion for including a response for analysis was that it contained an adequate description of the action of ‘giving’ or ‘showing’ a THEME.

Table 5.1: Summary of valid and excluded responses in the four groups.

GROUP	<i>n</i>	RESPONSES	
Heritage Ambon Malay speakers	32	Valid	181
		Excluded	11
Homeland Ambon Malay speakers	27	Valid	140
		Excluded	22
First generation Ambon Malay speakers	6	Valid	29
		Excluded	7
Dutch speakers	10	Valid	46
		Excluded	14

In the Ambon Malay dataset, we included responses with verbs of giving and showing: the most frequently used verbs are *kasi* ‘give’ (used 262 times), *kasi tunju* ‘show’ (lit. ‘give show’) (used 60 times), *kasi lia(t)* ‘show’ (lit. ‘give see’) (used 21 times) and *tunju* ‘show’(used nine times). Responses that were excluded contained verbs with a completely different meaning, such as *kembali* ‘return’, *tukar* ‘change,

exchange', *dapat* 'receive', *jual* 'sell', as well as reciprocals such as *baku-kasi* 'give each other', *baku-tukar* 'exchange with each other', and *baku-tawar* 'offer to each other'.

In the Dutch *give*-constructions we elicited, the most frequent verbs were *geven* 'to give' (used 26 times) and *laten zien* 'to show' (lit. 'to let see') (used 14 times). Additionally used verbs were *tonen* 'to show' (used one time), *overhandigen* 'to hand (over)' (used three times), and *doorgeven* 'to pass on' (used two times). As these verbs are all part of the class of *Geven-werkwoorden* 'Give-verbs' in Coleman's study of Dutch (2006, p. 437, Table 6.2), they were all included. Responses that were excluded contained verbs that are not in Coleman's *Geven-werkwoorden* class: *aangeven* 'to hand (to)', *toedraaien* 'to turn towards' and *aanbieden* 'to offer'.

5.4 Results and discussion

This section presents and discusses the results of the experiment. In each section, we discuss the results and propose explanations for the patterns observed.

5.4.1 Frequency of *give*-constructions

In this section we analyze the similarities and differences between all four test groups with respect to the various types of constructions which they use to express *give*-events.

5.4.1.1 Results

Figure 5.1 (on the next page) lays out the results regarding the frequency of the five types of *give*-constructions that are attested in the Ambon Malay data (see Section 5.2.1). It shows that heritage Ambon Malay lies in between homeland Ambon Malay and Dutch.

All four groups show a strong preference for the structure involving a single predicate in a PO construction (dark gray bar). Thus, in preferring PO constructions to DO constructions, the four groups behave alike. Another similarity across all four groups is that 'Recipient omission' occurs with approximately the same frequency (white bar). This pattern may be explained by considering the discourse status of participants in a *give*-event, where R is a previously activated ('known') topic that is

accessible to both hearer and speaker, and is thus more easily deleted than T, the element that conveys the ‘new information’.

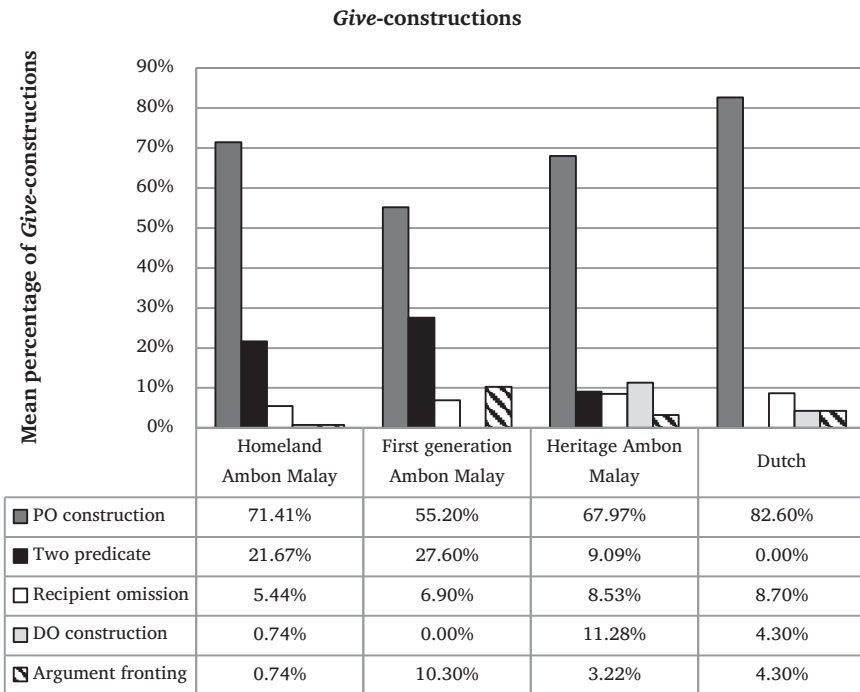


Figure 5.1: The mean percentages of the types of *give*-constructions attested in the four datasets.

Turning now to the differences between the heritage group and the homeland group, Figure 5.1 shows that these revolve around the ‘DO construction’ (light gray bar) and the ‘two predicate construction’ (black bar). Heritage Ambon Malay displays a significantly higher incidence of DO constructions ($M=.1128$, $SD=.163$) when compared to homeland Ambon Malay ($M=.074$, $SD=.384$) ($t(35.037)=3.537$, $p=.001$, $r=.51$, equal variances not assumed). In the heritage group, we find 20 tokens of DO constructions provided by 14 different speakers, while there is only one token of the DO construction in the homeland group. Interestingly, in the Dutch group, we find only two tokens of DO constructions, both provided by the same speaker.

Heritage speakers also show a significantly lower incidence of ‘two predicate constructions’ ($M=.909$, $SD=.135$) when compared to homeland speakers

($M = .2167$, $SD = .206$) ($t(43.361) = -2.708$, $p = .010$, $r = .38$, equal variances not assumed) and to first generation speakers. In the heritage group, only 12 speakers out of 32 provided at least one ‘two predicate construction’, while in the homeland group 17 speakers out of 26 did so, and in the first generation group all six speakers provided at least one ‘two predicate construction’. Thus, while the ‘two predicate construction’ is used by heritage speakers, nevertheless, they use it much less than the homeland and first generation speakers.

Finally, with respect to the ‘argument fronting construction’ (diagonal lines bar), heritage speakers and first generation speakers behave alike. (In the Dutch group only two tokens of this construction are attested). Increased use of a fronting strategy may be related to speakers having difficulties in accessing lexical items. It is known from the literature that speakers who have word finding problems (such as heritage speakers or elder speakers) tend to front well-known words in order to gain time when producing an utterance (see Aalberse & Muysken, 2013, p. 11; see also Section 1.3.5). However, given the overall paucity of this construction in our data, it will not be considered further here.

5.4.1.2 Discussion

Three main results emerge from the quantitative analysis of the data: (i) PO constructions are equally frequent in the homeland and heritage groups (as well as in the first generation and in the Dutch groups); (ii) DO constructions are used significantly more in the heritage group; (iii) ‘two predicate constructions’ are used significantly less in the heritage group. We discuss these three findings here in turn.

We suggest that PO constructions are stable in the heritage language because they are the preferred strategies in both Ambon Malay and Dutch (see Section 5.2). It seems that Ambon Malay always prefers PO constructions, independent of context. In contrast, Dutch has a clear bias for PO constructions only in experimental settings. The preference for PO in the heritage Ambon Malay data can thus be seen as a reflection of the overall preference to use PO constructions in both languages spoken by heritage speakers, in a de-contextualized experimental setting like ours. In a study on a Malay-Portuguese creole, Baxter (1990, p. 182) states that “the best chance for a feature to become dominant in a creole is where there is a conspiracy between more than one source: superstrate / substrate / creole universals”. If we extend this claim from creoles to heritage languages, we can say that the best chance for a feature to maintain its status quo in a heritage language is when there is a conspiracy between the homeland language and the dominant language.

Although the PO construction is dominant in Ambon Malay, the language also uses DO constructions, see Figure 5.1. Recall that in natural speech, DO constructions are more likely to occur with a pronominal R (see Section 5.2.1). In our experimental setting, however, responses depict de-contextualized events, where the R is not presupposed and not definite. Hence, reference to R is less likely to occur with a pronominal form in a DO construction. In other words, our setting predicts a bias for PO constructions over DO constructions in Ambon Malay and in Dutch. In this light, it is interesting to observe that heritage speaker still use DO constructions with a significantly higher frequency when compared to homeland speakers and to Dutch speakers. How can we explain this?

We propose that the increase in the frequency of DO constructions in heritage speakers result from two forces that apply at the same time: (i) indirect transfer from the dominant language Dutch (see Section 1.3.1), and (ii) universal principles of language development in the context of language disuse (see Section 1.3.5). These two forces are probably interrelated, but for the sake of clarity we discuss them separately here (see Section 1.3.6).

We have seen in Section 1.3.1 that, in the heritage contact scenario, syntactic changes nearly always involve an adaptation in the frequency distribution of patterns that already existed in the heritage language (Silva-Corvalán, 1993, 1994, 2008; Backus, 2004; Onar Valk, 2015). This type of change has been referred to as ‘indirect transfer’ by (Silva-Corvalán, 1994), and as ‘frequential copying’ by Johanson (2002). In this type of change, frequency patterns of the dominant language are copied onto the corresponding patterns of the heritage language (see Section 1.3.1.1). In usage-based terms, we could say that when speakers find evidence for a construction in both their dominant language and their heritage language, this construction becomes more entrenched and more productive (see Backus, 2004). As a construction that is grammatical in both Dutch and Ambon Malay, heritage speakers perceive that the DO construction in the dominant language, Dutch, has a structural counterpart in the recipient heritage language, Ambon Malay, and consequently, they use this construction more frequently. Recall that DO constructions are rather common in Dutch conversational data (Section 5.2.2)

However, cross-linguistic influence from Dutch cannot be the only source of divergence as heritage Ambon Malay seems to stand on its own with a rate of DO constructions higher than that of Dutch, the source language. So, apart from cross-linguistic influence from Dutch, there may be another factor responsible for the

higher incidence of DO constructions in heritage Ambon Malay: universal principles or regression processes under reduced input conditions (as attested in the case of creole genesis) (Benmamoun et al., 2011, p. 53). We know that DO constructions spontaneously emerge in contact varieties such as creoles. According to Michaelis (2014, p. 31), typical creole feature should not be found in the superstrate and substrate language, should be more common in creoles than in non-creoles, and should be identified by looking at languages which are relatively independent from each other in their historical origin. The WALS reports that DO constructions are found in 124 out of 378 languages (32.8%), while the APiCS reports that DO constructions are found in 60 out of 76 creoles (78.9%), and Bruyn, Muysken, & Verrips (1999, p. 330) show that they are found *even* in those creoles whose lexifier languages have no DO constructions. Thus, with respect to the first two parameters, DO may qualify as a typical creole feature. It has been suggested (Michaelis & Haspelmath, 2003), however, that the substrate may have played a role in the genesis of DO constructions in creoles, as DO constructions are found widely in the languages of West Africa. An argument against the substrate influence is that we know of no other West African structural feature that has had such a categorical pan-creole effect, going against the categorical word order in their lexifiers in the case of Romance-lexicon creoles (Pieter Muysken, personal communication; but cf. Michaelis, 2014). In addition, a number of studies report that DO constructions are overgeneralized by children during the course of L1 and L2 acquisition (see Mazurkewich & White, 1984; Whong-Barr & Schwartz, 2002, among others). Data from creoles and from language acquisition suggests that there is something special or iconic about DO constructions that makes this feature likely to be selected in language development under limited input.

To sum up, the innovative use of DO constructions in heritage Ambon Malay is the result of two forces: a process of ‘indirect transfer’ from Dutch and universal principles of language contact which favor DO constructions. These two forces are not easy to tease apart and it is reasonable to assume that they act in a cumulative way (Polinsky & Kagan, 2007, p. 382; see also Section 1.3.1.1 and Section 1.3.1.2).⁷⁶

⁷⁶ One way to tease apart these two forces would be to study heritage speakers of Ambon Malay with a dominant language that does not allow DO constructions, such as Italian. If this hypothetical group also shows a higher rate of DO constructions compared to homeland speakers, then we can conclude that universal principles are the main source of divergence. If, conversely, heritage Malay speakers with Italian as dominant language do not show a higher

The third finding we discuss here is that ‘two predicate constructions’ are used less frequently by heritage speakers than by homeland (and first generation) speakers, as shown in Figure 5.1. In other words, homeland speakers are more likely to describe the visual stimuli by using two predicates. This pattern reflects a way of segmenting the flow of information which is typical of Ambon Malay (Tjia, 1997; van Staden & Reesink, 2008). The change in preference where heritage speakers strongly prefer constructions with only one predicate suggests that they apply different information organization principles (Slobin, 1991; Carroll & von Stutterheim, 2003). We have seen in Section 1.4.2 that one of the hypotheses in bilingualism research is that speakers of different languages display differing patterns of event construal, and that these patterns have the potential to transfer across languages (the Conceptual Transfer Hypothesis). The decrease in the use of ‘two predicate construction’ in heritage Ambon Malay lends support to this hypothesis, and suggests that heritage speakers organize information via their dominant language, Dutch, where *give*-events are prototypically expressed with a single verbal predicate (see Section 5.2.2).

Further evidence for the claim that heritage speakers express events using the organizing principles of Dutch rather than of Ambon Malay comes from data on another type of semantically complex events - resultative constructions -collected from the same speaker groups (Moro, 2014; Chapter 6 in this dissertation). In a video-clips retelling task eliciting resultative constructions (e.g. ‘break a stick (in two)’, ‘tear a piece of cloth (in two)’), homeland speakers used a construction involving two predicates (either verb serialization or a ‘two predicate construction’) in 61.8% of the responses, while heritage speakers did so only in 15.5% of the responses. These data indicate that heritage speakers are shifting towards a Dutch-like way of organizing information using one predicate instead of two, and that this shift involves various syntactic domains, including the expression of *give*-events and of resultative events.

In arguing about the degree of restructuring of heritage speakers’ grammars, it is important to consider individuals’ data, since heritage speakers are known to have variable language backgrounds. We therefore investigated whether there is a relation between the amount of exposure that individuals had to Ambon Malay in the course of their lifetime and their use of DO and ‘two predicate constructions’. The investigation carried out here is mostly qualitative, for a quantitative analysis

rate of DO, then we can conclude that, in the case of heritage Ambon Malay speakers in the Netherlands, transfer from Dutch is the major force at work.

the reader is referred to Chapter 7 (Section 7.3) We divided the heritage speakers into two groups, according to the amount of exposure to Ambon Malay: speakers who had either predominantly Dutch parental input or mixed input (Ambon Malay and Dutch) belong to the ‘LOW exposure’ group, while speakers whose parents spoke predominantly Ambon Malay belong to the ‘MEDIUM-HIGH exposure’ group, see Table 5.2. The division according to parental input roughly corresponds to the type of bilingualism: most simultaneous bilinguals (AoA from birth) belong to the ‘LOW exposure’ group, while sequential bilinguals (AoA since the age of four) belong to the ‘MEDIUM-HIGH exposure’ group.

Table 5.2: The sociolinguistic background of the heritage Ambon Malay group. Speakers using DO constructions are shaded gray, speakers using the ‘two predicate construction’ are printed in bold.

	Sp	AoA	L Mo	L Fa	L Sb	L Pa	LIVES	GREW	HC
	H30	0	Dutch	Dutch	Dutch	Dutch	city	city	2
	H11	0	Dutch	Dutch	Dutch	Dutch	city	city	4
	H32	0	Dutch	Dutch	Dutch	Dutch	city	city	6
	H31	0	AM	Dutch	Dutch	Dutch	city	city	0
	H14	0	Dutch	Dutch	Dutch	Dutch	city	city	4
	H33	0	AM	Dutch	Dutch	Dutch	city	city	2
	H17	0	Dutch	AM	Dutch	Dutch	city	city	2
	H27	0	Dutch	Dutch	Dutch	Dutch	<i>wijk</i>	city	0
	H23	0	Dutch	Dutch	Dutch	Dutch	<i>wijk</i>	city	3
	H21	0	Dutch	Dutch	Dutch	Dutch	<i>wijk</i>	city	4
	H13	0	Dutch	Dutch	Dutch	Dutch	<i>wijk</i>	<i>wijk</i>	0
	H29	0	Dutch	AM	Dutch	Dutch	city	city	5
	H25	0	Dutch	AM	Dutch	Dutch	<i>wijk</i>	<i>wijk</i>	0
	H19	0	Dutch	Dutch	Dutch	Dutch	<i>wijk</i>	<i>wijk</i>	1
	H20	0	mixed	Dutch	Dutch	-	<i>wijk</i>	city	6
	H22	0	Dutch	Dutch	Dutch	Dutch	<i>wijk</i>	camp	3
	H8	0	Dutch	Dutch	Dutch	mixed	<i>wijk</i>	city	6
	H7	0	mixed	mixed	Dutch	Dutch	<i>wijk</i>	<i>wijk</i>	1
	H6	0	Dutch	AM	mixed	mixed	<i>wijk</i>	city	-

LOW EXPOSURE TO AMBON MALAY

The table continues on the next page

	Sp	AoA	L Mo	L Fa	L Sb	L Pa	LIVES	GREW	HC
MEDIUM-HIGH EXPOSURE TO AMBON MALAY	H26	0	AM	AM	mixed	Dutch	<i>wijk</i>	city	3
	H15	0	AM	AM	Dutch	mixed	<i>wijk</i>	city	3
	H3	0	AM	Dutch	mixed	AM	<i>wijk</i>	camp	2
	H28	> 4	AM	AM	mixed	Dutch	city	camp	4
	H24	> 4	AM	AM	mixed	Dutch	city	camp	10
	H2	> 4	AM	AM	mixed	Dutch	<i>wijk</i>	camp	-
	H9	> 4	AM	AM	mixed	Dutch	<i>wijk</i>	<i>wijk</i>	5
	H16	> 4	AM	AM	mixed	Dutch	<i>wijk</i>	camp	2
	H18	> 4	AM	AM	AM	Dutch	<i>wijk</i>	<i>wijk</i>	3
	H5	> 4	AM	AM	mixed	Dutch	<i>wijk</i>	camp	6
	H4	> 4	AM	AM	mixed	mixed	<i>wijk</i>	camp	3
	H1	> 4	AM	AM	AM	Dutch	<i>wijk</i>	camp	10
H12	> 4	AM	AM	mixed	AM	<i>wijk</i>	camp	8	

From the information presented in Table 5.2, three observations can be made. First, the majority of heritage speakers who use DO constructions (shaded gray) belong to the ‘LOW exposure’ group. They grew up as simultaneous bilinguals and mainly spoke Dutch in the family. Half of them live in cities (outside a Moluccan *wijk* ‘ward’), and are thus completely immersed in a Dutch speaking environment.

Second, most of the speakers who use the ‘two predicate construction’ (printed in bold) belong to the ‘MEDIUM-HIGH exposure’ group. These speakers report that they mainly spoke Ambon Malay in the family during their childhood, and nowadays most of them live in a Moluccan *wijk*, thus surrounded by many Moluccan neighbors. Third, overall, speakers who use DO constructions do not use ‘two predicate constructions’, and *vice versa* (speakers H16, H2, H32 are the only exceptions). The fact that one and the same speaker typically does not use both of the constructions confirms that DO constructions are an innovation that occurs in speakers who had relatively low exposure to Ambon Malay, while the ‘two predicate constructions’ are a typical Ambon Malay feature maintained by speakers with a medium-high exposure to Ambon Malay.

In sum, the information on the language history of the speakers suggests that low exposure to Ambon Malay corresponds to the use of DO constructions as an innovation in the heritage variety. In contrast, high exposure to Ambon Malay corresponds to the use of the ‘two predicate construction’, a feature typical of the homeland variety.

5.4.2 Preposition selection

In the Ambon Malay data, seven prepositions are used to encode the PO (see section 5.2.1). This section reports the differences in preposition selection, and offers an explanation for the attested variation.

5.4.2.1 Results

The use of prepositions in the PO construction by the three Ambon Malay speaker groups is presented in Table 5.3. Seven prepositions are used, with similar meanings: *par* ‘for, to’, *for* ‘for, to’, *buat* ‘for, to’, *untuk* ‘for’, *ka* ‘to’, *kepada* ‘to’ and *sama* ‘to, with’ (see Section 5.2.1 for illustrations). We only compared, using an independent *t*-test, the means of the homeland and the heritage group, as the number of first generation speakers is too small to allow a reliable statistical analysis. Table 5.3 shows that with respect to preposition selection, first generation speakers follow the pattern of heritage speakers, *not* of homeland speakers. While homeland speakers show a very strong preference for *par* ‘to’ (65.48%), the other two groups prefer the preposition *buat* ‘for, to’ and *untuk* ‘for’. The only preposition attested with a similar frequency across the three groups is *for* ‘for, to’. The remaining three prepositions *ka* ‘to’, *kepada* ‘to’ and *sama* ‘to, with’ are not very frequent. *Ka* ‘to’ occurs both in homeland speakers and in heritage speakers, while *kepada* ‘to’ is found only in heritage speakers (nine tokens). Finally, there is only one token of *sama* ‘to, with’ which was produced by a first generation speaker.

Table 5.3: The mean percentages of the different prepositions attested in the Ambon Malay datasets. Asterisks indicate statistically significant differences.

PREPOSITION	GROUP	<i>n</i>	MEAN %	SD	SIG. (2 TAILED)
<i>par</i>	Homeland	27	65.48%	.366	
	Heritage	32	2.59%	.085	.000***
	First generation	6	17.83%	.222	
<i>buat</i>	Homeland	27	1.89%	.054	
	Heritage	32	32.50%	.348	.000***
	First generation	6	30.67%	.427	

The table continues on the next page.

PREPOSITION	GROUP	<i>n</i>	MEAN %	SD	SIG. (2 TAILED)
<i>for</i>	Homeland	27	22.07%	.349	
	Heritage	32	20.81%	.362	.893
	First generation	6	18.33%	.285	
<i>untuk</i>	Homeland	27	0.00%	.000	
	Heritage	32	14.69%	.275	.008**
	First generation	6	16.67%	.408	
<i>ka</i>	Homeland	27	3.78%	.126	
	Heritage	32	4.38%	.125	.856
	First generation	6	0.00%	.000	
<i>kepada</i>	Homeland	27	0.00%	.000	
	Heritage	32	4.88%	.133	.063
	First generation	6	0.00%	.000	
<i>sama</i>	Homeland	27	0.00%	.000	
	Heritage	32	0.00%	.000	-
	First generation	6	2.83%	.069	

5.4.2.2 Discussion

We propose that heritage speakers show a preference for *buat* ‘for, to’ because *buat* was the most frequent form in the input they picked up from their parents. In other words, the Ambon Malay variety spoken by the first generation differs in choice of preposition from the one currently spoken by homeland speakers in Ambon, and the heritage speakers reflect the preposition choice of the first generation. In this area of the grammar, the divergence between heritage Ambon Malay and homeland Ambon Malay is, thus, ascribed to a different type of input that heritage speakers received, and not to other factors, such cross-linguistic influence or incomplete acquisition.

The incongruence between choice of preposition by first generation and homeland speakers has an obvious historical explanation. As discussed in Section 1.5.2, the majority of the Moluccan KNIL-soldiers and their families who were shipped to the Netherlands spoke Tangsi Malay or Ambon Malay with a strong Tangsi flavor. It is likely that *buat* ‘for, to’ was used in Tangsi Malay, as it is the preposition that is prototypically used to mark a PO in *give*-constructions in numerous eastern varieties of Malay, as well as in colloquial Indonesian (van Minde,

1997; Paauw, 2008). Heritage speakers thus find evidence for *buat* ‘for, to’ not only in their parental language but also when they interact with Malay speakers from other parts of the Indonesian archipelago.

Par ‘for, to’, on the other hand, is a lexeme unique to Ambon Malay. It was brought along by Moluccan immigrants to the Netherlands, and was thus part of the input of heritage speakers. In addition, *par* has also spread in the Ambon Malay community of the Netherlands because heritage speakers visit Ambon and new Ambon Malay speaking immigrants continue to arrive in the Netherlands.

Unlike *par* and *buat*, the prepositions *untuk*, *ka*, *kepada* and *sama* are not prototypically used in Ambon Malay *give*-constructions: *ka* means ‘to’ and marks a direction or a goal, *kepada* introduces an animate recipient in very formal contexts, *untuk* indicates a beneficiary rather than a recipient, and *sama* ‘be equal to, to, with’ is a multifunctional preposition that usually has a comitative meaning. *Kepada*, and most probably *untuk* as well, are recent loan from Standard Indonesian. It is probable that heritage speakers picked them up in the Moluccan churches in the Netherlands (where services are conducted in Standard Indonesian), through interactions with Indonesian speaking people in the Netherlands (relatives, or clergypersons), or when visiting Indonesia. In Section 1.3.4, I discussed the fact that heritage speakers show difficulties in understanding and mastering the complete range of registers and styles available to monolingual homeland speakers. This may account for their use of more ‘formal’ prepositions such as *kepada* in the *give*-constructions elicited by the video-clips.

In sum, in their choice of prepositions in the PO construction, heritage speakers resemble first generation speakers, and diverge from homeland speakers of Ambon Malay. The choice of prepositions in the heritage language shows traces of Tangsi Malay, and also of interactions with speakers of Standard Indonesian. As such, it reflects the different social histories of heritage and homeland speakers.

5.5 Conclusions

Ambon Malay has five constructions to express *give*-events: (i) the prepositional object (PO) construction, (ii) the double object (DO) construction, (iii) the Recipient omission construction, (iv) the ‘two predicate construction’ and (v) the ‘argument fronting construction’. These constructions are used in both the homeland and the heritage varieties. However, heritage Ambon Malay diverges from the homeland

variety in a number of ways. First, heritage Ambon Malay shows an innovative higher incidence of DO constructions. We propose that this change is the result of two forces, acting in a cumulative way: a process of ‘indirect transfer’ from Dutch, and universal principles in situations of language disuse which favor DO constructions. The language history of the individual heritage speakers confirms our analysis that a higher incidence of DO constructions corresponds to low exposure to Ambon Malay and high exposure to Dutch.

Heritage Ambon Malay also differs from the homeland variety in that it shows a lower incidence of ‘two predicate constructions’. This is also seen as influence from Dutch, where *give*-events are prototypically expressed with a single verbal predicate. The language history of individual heritage speakers indicates that speakers with a high exposure to Ambon Malay use this typical Ambonese construction more often than heritage speakers with a history of low exposure to Ambon Malay.

Thus, the divergence between heritage and homeland Ambon Malay *give*-constructions does not involve a categorical change, but rather it manifests itself as a change in the frequency of already existing constructions. This ‘restructuring by changing frequency’ is partly due to the path of language acquisition of heritage Ambon Malay speakers, and partly due to contact with Dutch. Apart from quantitative differences, there are also qualitative differences between the *give*-constructions of heritage and homeland speakers of Ambon Malay: the different prepositions both groups choose for the PO construction reflect their different social histories and the different type of input they were exposed to.

In conclusion, this study has shown once more that domains where languages allow variable syntax are susceptible to cross-linguistic effects. When the variable argument encodings of *give*-events in two languages are combined in the same bilingual speaker, new frequency patterns emerge.

CHAPTER 6

Resultative constructions in heritage Ambon Malay⁷⁷

6.1 Introduction

Resultatives are an interesting area for the study of language contact phenomena because we know from previous research that the domains where languages have two (or more) competing syntactic constructions expressing the same meaning are problematic for bilingual (heritage) speakers. This observation constitutes the main tenet of the Alternation Hypothesis, as discussed in Section 1.4.1. The Alternation Hypothesis predicts that, if the heritage language allows an alternation between two constructions, heritage speakers will tend to use more frequently the construction which is present in the dominant language. This type of contact-induced change, defined as ‘change in frequency’ in Section 1.3.1.1, is likely to lead to convergence between the heritage language and the dominant language. In other words, bilingual speakers use the overlap between their two languages to bring the two languages closer together. Change in frequency between two equally possible options is probably among the most common types of changes in heritage languages (Backus, 2004; Muysken, 2005; Silva-Corvalán, 1994, 2008; Onar Valk, 2015; see examples in Section 1.3.1.1).

In many languages, including Ambon Malay and Dutch, the domain of resultative events allows variable syntactic encoding. Resultative events, like motion events, represent a type of complex events, and languages differ with respect to the elements (i.e., verb, prepositional phrase, adjective) expressing the various meaning components (i.e., MANNER, RESULT, MOTION, PATH) (Kaufmann & Wunderlich, 1998; Talmy, 2000; Croft, Barðdal, Hollmann, Sotirova, & Taoka, 2010). More specifically, resultative constructions contain: (i) a verb form denoting an activity (the MANNER component in Levin & Rappaport Hovav, 2005), (ii) an argument that undergoes a change of state as a result of the activity (the THEME) and (iii) a component denoting the resultant state that is caused by the activity (the RESULT

⁷⁷ This chapter is partially based on Moro, F. R. (2014). Resultative constructions in heritage Ambon Malay in the Netherlands. *Linguistics in the Netherlands*, 31(1), 78-92.

component in Levin & Rappaport Hovav, 2005; see also Goldberg, 1995; Boas, 2003; Goldberg & Jackendoff, 2004). For instance, in the English sentences *Tony broke the piggy bank to pieces* and *Tony broke the piggy bank open* (Levin, 1993, p. 242) the main verb *broke* expresses the MANNER component, the *piggy bank* is the THEME, and the prepositional phrase *to pieces* in the first case, and the adjective *open* in the second case, express the RESULT component.

Languages display considerable variability in the way the meaning components (MANNER and RESULT) are mapped onto syntactic structures. The main typological distinction is between verb framing languages and satellite framing languages (Talmy, 2000; but cf. Croft et al., 2010).

The verb framing category subsumes all languages in which the RESULT component is expressed by a verb. Strategies of verb framing languages include compounding, verb serialization and coordination (cf. Croft et al., 2010). These strategies differ with respect to the degree of integration (co-predication in a single clause or separate clauses) but they are all symmetric because the MANNER and RESULT are expressed by verbal elements. For instance, in the Ambon Malay resultative construction *dia pata akang jadi dua*, literally ‘she snaps it becomes two’, the MANNER *pata* ‘snap’ and the RESULT *jadi dua* ‘become two’ are encoded by two elements that can independently function as verbs. Symmetric strategies, such as serial verb constructions, are attested in creoles (Veenstra, 1996, 2003), as well as in Austronesian languages, including several Malay varieties (Tjia, 1997; Donohue, 2011; Jacob & Grimes, 2011), and in Papuan languages (Senft, 2008).

The satellite framing category subsumes all languages in which the RESULT component is expressed by a satellite element, such as a (P)repositional (P)hrase or and (A)djectival (P)hrase. Satellite framing constructions are asymmetric because the MANNER is encoded by a verb, whereas the RESULT is encoded by a non-verbal part, usually a particle, a PP or an AP, as in *Tony broke (MANNER) the piggy bank to pieces/open (RESULT)*. Germanic languages, such as Dutch, English and German, are of the satellite framing type. Languages can also have mix patterns and make use of various strategies, such as verb serialization and prepositional phrases, alongside each other.

Given the variability of encoding strategies in the domain of resultative events, it is likely that (two) languages in contact share (at least) one type of construction, or/and that they have constructions that resemble each other (at the surface level). We know that partial overlap of structures between languages poses the conditions for cross-linguistic influence (see the Alternation Hypothesis in Section 1.4.1).

Bilingual speakers, in fact, tend to select the construction that is shared by both languages more frequently, thus increasing the structural similarity between the two languages, a process known as ‘convergence’ (see Section 1.3). Convergence by changing frequency can lead, in long-term contact situations, to a profound restructuring of the grammar, to the extent that a language can change its typological profile. For instance, Tetun Dili, an Austronesian serializing language in contact with a non-serializing European language, Portuguese, is changing from a serializing-type language to a preposition- type language. Hajek (2006) explains that due to the long-term contact with Portuguese, Tetun Dili speakers are moving toward the encoding preferences of Portuguese by replacing serial verb constructions with lexical verbs (loans from Portuguese) and prepositions.

This chapter contributes to our understanding of cross-linguistic effects in the domain of complex events by studying the expression of resultative events in heritage Ambon Malay. The central question addressed in this chapter is: have resultative constructions in heritage Ambon Malay been restructured as compared to the Ambon Malay spoken in the homeland? And if restructuring of resultative constructions in heritage Ambon Malay did take place, what did the change involve? The results show structural convergence between heritage Ambon Malay and Dutch, with heritage speakers moving toward the encoding preferences of Dutch speakers. Convergence is instantiated as a significant change in the frequency with which certain constructions that exist in the homeland variety occur in the heritage language. I argue that this ‘restructuring by changing frequency’ is mainly caused by contact with Dutch, although other factors, such as universal principles of language development in contact settings also play a role.

This chapter is structured as follows. Section 6.2 describes how resultative events are expressed in the languages of bilingual heritage speakers: Ambon Malay (section 6.2.1) and Dutch (section 6.2.2). Section 6.3 illustrates the design of the present study. The results are presented and discussed in section 6.4. Section 6.5 summarizes the conclusions.

6.2 Resultative constructions in Ambon Malay and Dutch

This section presents a descriptive overview of the various ways in which resultative events are expressed in Ambon Malay and Dutch, the two languages of bilingual heritage speakers of Ambon Malay. This chapter focuses on resultative constructions

involving verbs for *cutting*, *breaking* and *hitting* because these events were depicted in the video-clips used as elicitation material (see Section 6.3.3). Interestingly, due to their semantic differences, these verbs are likely to trigger different types of resultative constructions. Verbs for *cutting* and *breaking* (hereafter CB) entail a change of state in the entity affected, whereas verbs for *hitting* do not necessarily entail that the contact has any effect on the entity (Levin, 1993, p. 150). CB verbs usually select a resultative PP because PPs simply specify the change of state that is already encoded by the base verb, without introducing independent information (Kaufmann & Wunderlich, 1998, p. 14). In contrast, verbs for *hitting* tend to select a resultative AP, because APs add independent information, namely a new sub-event (become X) which is not implied by the base verb (Levin, 1993; Kaufmann & Wunderlich, 1998; Boas, 2003). For instance, in the Dutch dataset used for the present study, the CB verb *scheuren* ‘to tear’ was used almost always with a resultative PP (*scheuren in tweeën* ‘to tear into two’) (80%); while the verb *slaan* ‘to hit’ was used almost always with a resultative AP (*slaan een vaas kapot*, lit.: ‘to hit a vase broken’) (80%).

The remainder of this section focuses on describing those structural features that are relevant for answering the question how resultative constructions in the heritage language have been restructured as compared to the language of the homeland. Unless otherwise specified, the examples presented in this section and elsewhere in this chapter are from the dataset collected for this dissertation (see Section 2.2 and Section 2.3).

6.2.1 Resultative constructions in Ambon Malay

In the domain of resultative events, Ambon Malay allows five basic constructions. All these five constructions are attested in both homeland and heritage Ambon Malay.

The first construction is the Serial Verb Construction (SVC), in which the MANNER and the RESULT are expressed by two verbs sharing at least one argument, usually the THEME (T) (Tjia, 1997). The shared argument can be the object of the first verb (V_1) and the subject of the second verb (V_2), as illustrated (1a) where *kaeng* ‘cloth’ is simultaneously the object of *robe* ‘tear’ and the subject of *jadi* ‘become’. The two verbs can also share both their subject and their object, as shown in (1b) where the verb *bage* ‘divide’ shares the subject *dia* ‘3SG’ and the object *akang* ‘3SG.N’. When V_1 and V_2 share the same object, the object argument may intervene between the

two verbs (V_1 O V_2), as in (1b) or follows the verb complex (V_1 V_2 O). In resultative events involving *cutting*, *breaking* and *hitting* verbs, V_2 is usually *jadi* ‘become’, as illustrated in (1a), or *bage* ‘divide’, as illustrated in (1b), or *bala* ‘split’.

- (1) a.

			V_1	T	V_2		
	<i>Parampuang</i>	<i>robe</i>	<i>kaeng</i>	<i>jadi</i>	<i>dua</i>		
	girl	tear	cloth	become	two		

‘A girl tears a piece of cloth into two (lit.: tears a piece of cloth becomes two).’
- b.

	<i>Ada</i>	<i>nona</i>	<i>pegang</i>	<i>satu</i>	<i>kaeng</i>	<i>ijo</i>	
	EXIST	girl	hold	one	cloth	green	
			V_1	T	V_2		
	<i>lalu</i>	<i>dia</i>	<i>robe</i>	<i>akang</i>	<i>bage</i>	<i>dua</i>	
	then	3SG	tear	3SG.N	divide	two	

‘There is a girl holding a piece of green cloth, and then she tears it up into two (lit.: tears it up divides two).’

The second type of construction found in Ambon Malay is one that I refer to as the ‘two predicate construction’.⁷⁸ We have already encountered this construction in Chapter 5 (see Section 5.2.1), when discussing *give*-events. The ‘two predicate construction’ expresses resultative events (and *give*-events) using two predicates in a single sentence, as illustrated in (2). The MANNER is expressed by the first predicate - a transitive verb- (e.g., *potong* ‘cut’); while the RESULT is expressed by the second predicate, - an intransitive verb- (e.g., *tabage* ‘get divided’). Unlike the verbs of a SVC, the verbs of a ‘two predicate construction’ do not share any argument: the argument is either repeated, as in (2a, *ikang* ‘fish’, *ikang* ‘fish’), referred to with a pronoun, as in (2b, *wortel* ‘carrot’, *akang* ‘3SG.N’) or dropped, as in (5b) below. The ‘two predicate construction’ may consist of two clauses that are connected with a conjunction: *tarus* ‘and then, next’, as in (2a), *la/lalu* ‘and then’, or *sampe* ‘until’, as in (5b). The clauses can also be simply juxtaposed without an overt linker, separated by a pause, as shown in (2b).

⁷⁸ In Moro (2014), this construction is referred to as ‘Coordination’. For the sake of comparability, here I adopt the term ‘two predicate construction’.

- (2) a.

<i>Dia</i>	V ₁	<i>potong</i>	T	<i>ikang</i>	<i>tarus</i>	T	<i>ikang</i>	<i>ta-bage</i>	<i>tiga</i>
3SG		cut		fish	then		fish	ACL-divide	three

‘She cuts a fish, and the fish gets divided into three.’
- b.

<i>Orang</i>	<i>ada</i>	V ₁	<i>potong</i>	T	<i>wortel,</i>	T	<i>akang</i>	V ₂	<i>jadi</i>	<i>dua</i>
person	EXIST		cut		carrot		3SG.N		become	two

‘A person has cut a carrot, it became two (pieces).’

SVCs and ‘two predicate’ constructions can be subsumed under the category ‘symmetrical construction’, because the MANNER and RESULT components of the event are expressed by two verbs (Croft et al., 2010, p. 206). ‘Symmetrical constructions’ are the most frequent constructions in the (homeland) Ambon Malay dataset and account for about 62% of all resultative constructions (see Figure 6.1 below). As we have already seen in Chapter 5, segmenting the flow of information over two or more predicates is a typical feature of the Malay varieties that are spoken in eastern Indonesia (Tjia, 1997; van Staden & Reesink, 2008; Paauw, 2008, pp. 232-236; see also Section 5.2.1).

The third construction found in the Ambon Malay dataset is the one in which the MANNER is expressed by the main verb and the RESULT is expressed by means of a Prepositional Phrase (PP), as illustrated in (3).

- (3) a.

<i>Ada</i>	<i>nona</i>	<i>yang</i>	<i>robe</i>	<i>kaeng,</i>
EXIST	girl	REL	tear	cloth
	V	T	PP	
<i>dia</i>	<i>robe</i>	<i>akang</i>	<i>par</i>	<i>dua</i>
3SG	tear	3SG.N	to	two

‘There is a girl who tears a piece of cloth, she tears it up into two’
- b.

<i>Nona</i>	<i>ada</i>	V	<i>robe</i>	T	<i>kaeng</i>	PP	<i>dalang</i>	<i>dua</i>	<i>bagean</i>
girl	EXIST		tear		cloth		in(side)	two	part

‘A girl has torn a piece of cloth into two parts.’

As the examples (3a and 3b) show, different prepositions may be used to introduce the resultative phrase. The preposition *par* ‘for, to’ (3a) is a lexeme that is often used in Ambon Malay (van Minde, 1997, p. 76; see also Section 5.4.2). Not being found in any other Malay variety, it can be considered a unique and typical feature of Ambon Malay. In the Ambon Malay dataset, we find that *par* ‘for, to’ is only used by homeland speakers, while heritage speakers use other prepositions, such as *dalam/dalang* ‘in, inside’ (in 3b), *ka* ‘to’, or *di* ‘at, in’. The variable choice of prepositions in the PP construction is further discussed in Section 6.4.2.

In the fourth construction, the RESULT is expressed by a stative intransitive verb in a subtype of serial verb construction, as illustrated in (4), where the resulting state is encoded by the stative verb *pica* ‘broken’. Ambon Malay stative intransitive verbs (e.g., *pica* ‘broken’, *pende* ‘short’) express properties that are normally expressed by adjectives in European languages (see also Chapter 3 in this dissertation).

- | | | | | | | | | |
|-----|--|---------------|-------------|-----------|---------------|-------------|-------------|-----------|
| | | | | | V | T | AP | |
| (4) | <i>Oe</i> | <i>setang</i> | <i>sapa</i> | <i>su</i> | <i>lempar</i> | <i>kaca</i> | <i>pica</i> | <i>tu</i> |
| | EXCL | ghost | who | PRF | pelt | glass | broken | D.DIST |
| | ‘Hey, who the hell pelted the glass (until it got) broken?’ (Tjia 1997:56) | | | | | | | |

I refer to the construction in (4) as the (A)djectival (Phrase) construction, rather than as SVC for two reasons. First, the AP construction in Ambon Malay is identical, at the surface level, to the AP construction in Dutch (see Section 6.2.2). We know that (perceived) similarities between constructions are likely to trigger cross-linguistic influence, so it is convenient to label them in a comparable way (see Section 1.3.1.1). Second, AP constructions are rare in the homeland Ambon Malay dataset. There is only one instance of RESULT expressed solely by an AP (see Figure 6.1 below). It is much more common to find the adjective (or stative verb) in a SVC, where it is introduced by the causative *kasi* ‘give’, as in (5), or in a ‘two predicate construction’, where it is introduced by the conjunction *sampe* ‘until’, as in (5b).

‘through’ particle, such as *afhakken* ‘to chop off’ or *doorbreken* ‘to break through’, as illustrated in example (7a-7b).

- (7) a. *Een man hak-t een tak van*
 ART.INDF man chop-3SG ART.INDF branch of
 VP
een boom af
 ART.INDF tree off
 ‘A man cuts off a branch of a tree.’
- b. *Iemand breek-t een touw door*
 someone break-3SG ART.INDF rope through

met een beitel
 with ART.INDF chisel
 ‘Someone breaks a rope in half with a chisel.’

The second possibility for expressing the RESULT is to use the adverb (ADV) *doormidden* ‘through, in half’⁷⁹ or some of its less grammaticalized variants, such as the PP *door de midden*⁸⁰, as illustrated in (8). *Doormidden* ‘through, in half’ specifies the location where the separation takes place (through the middle), thus implying that the object is divided into two more or less equal halves.

- (8) *Een vrouw scheur-t een stuk stof doormidden*
 ART.INDF woman tear-3SG ART.INDF piece cloth in.half
 ‘A woman tears a piece of cloth in half.’

⁷⁹ *Doormidden* is kept separate from verb particles, because unlike the latter it cannot cluster in embedded clauses with the verb after the auxiliary (Geert Booij, personal communication, November 20th, 2015).

⁸⁰ In the Dutch dataset, there are 14 instances of *doormidden* ‘through/in half’, five instances of the less grammaticalized variant *door de midden*, and one instance of *door het midden*.

Another possibility is to encode the RESULT by means of a PP. In this type of construction, the main verb describes the activity or the MANNER, while the PP specifies the change of state that results from the activity. Two typical PPs that were used in the description of the stimuli are *in tweeën* ‘into two’ or *in twee/drie stukken* ‘in two/three pieces’, as shown in (9). The PP *in tweeën* specifies the number of parts in which the object is divided.⁸¹ When the number of pieces is specified, as in the PPs *in tweeën* ‘into two’ or *in twee/drie stukken* ‘in two/three pieces’, the preposition is always *in* ‘in’. In the remainder of this chapter, I use the label ‘PP’ or ‘PP construction’ only to refer to the PP *in tweeën* ‘into two’ or *in twee/drie stukken* ‘in two/three pieces’.

- | | | | V | T | | PP |
|-----|--|--------------|-----------------|------------|-------------|------------------|
| (9) | <i>Een</i> | <i>vrouw</i> | <i>scheur-t</i> | <i>een</i> | <i>doek</i> | <i>in tweeën</i> |
| | ART.INDF | woman | tear-3SG | ART.INDF | cloth | in two |
| | ‘A woman tears a piece of cloth into two.’ | | | | | |

The fourth possibility is to encode the RESULT by means of an AP, as illustrated in (10). In the AP construction, the main verb describes the MANNER, while the AP specifies the RESULT. According to Kaufman and Wunderlich (1998, pp. 18-19), the AP introduces a change of state that is not implied by the base verb, but it nevertheless depends crucially on it. In the Dutch dataset, APs occur only with the verb *slaan* ‘to hit’. As noted by Levin (1993, p. 150), ‘hit’ verbs “describe moving one entity in order to bring it into contact with another entity, but they do not necessarily entail that this contact has any effect on the second entity”. In order to encode the effect, an AP is then required.

- | | | | V | T | | AP |
|------|---|--------------|---------------|------------|-------------|--------------|
| (10) | <i>Een</i> | <i>vrouw</i> | <i>slaa-t</i> | <i>een</i> | <i>vaas</i> | <i>kapot</i> |
| | ART.INDF | woman | hit-3SG | ART.INDF | vase | broken |
| | ‘A woman hits a pot (until it got) broken.’ | | | | | |

Finally, Dutch speakers can decide to leave out any further specification of the RESULT, as illustrated in (11).

⁸¹ I wish to thank Margot van den Berg, Lotte Hogeweg and Monique Flecken for the insightful discussion on the difference between *doormidden* and *in tweeën*.

		V		T	
(11)	<i>Er</i>	<i>word-t</i>	<i>een</i>	<i>wortel-tje</i>	<i>gebroken</i>
	there	become-3SG	ART.INDF	carrot-DIM	break.PST.PTCP
	'A carrot is broken.'				

6.2.3 Interim summary

We have seen that in the domain of resultative constructions, Ambon Malay and Dutch, represents two different types of languages. Ambon Malay prefers symmetrical strategies where the RESULT is expressed by a verbal element (SVC and 'two predicate construction'), while Dutch makes use of satellite elements, such as verb particles, the adverb *doormidden* 'through, in half', PPs and APs. The two languages, however, share the PP and AP constructions, because these two options are available in both languages. They also share the possibility of leaving the RESULT unspecified. The differences and similarities are summarized in Table 6.1.

Table 6.1: Summary of resultative constructions in Ambon Malay and in Dutch.

	AMBON MALAY	DUTCH
Symmetrical	Serial Verb Construction	-
	'Two Predicate Construction'	-
Satellite	-	Verb Particle
	-	<i>Doormidden</i> 'through, in half'
	PP	PP
	AP	AP
	Unspecified Result	Unspecified Result

Finally, in the Ambon Malay dataset, we find different prepositions to encode the RESULT: *par* ‘to, for’ (used by homeland speakers), and *dalang* ‘in(side)’ and *ka* ‘to’ (used by heritage speakers); in the Dutch dataset, we find *in* ‘in’ and *tot* ‘to, until’.⁸²

6.3 The study

This section discusses the objective (6.3.1), the research questions (6.3.2) and the methodology (6.3.3) of the present study.

6.3.1 Objective

In this chapter, I investigate the frequency of competing resultative constructions, namely SVC, ‘two predicate’, PPs and APs in semi-spontaneous speech from heritage speakers of Ambon Malay in the Netherlands. I compared resultative constructions in heritage Ambon Malay to those of homeland Ambon Malay in order to detect signs of structural divergence between the two varieties, and to those of Dutch in order to find out signs of structural convergence between the heritage language and the dominant language.

6.3.2 Research questions and hypotheses

This study asks whether resultative constructions in heritage Ambon Malay have been restructured as compared to the Ambon Malay spoken in the homeland. More specifically, this section focuses on the following two research questions: (i) is the frequency of the attested constructions the same among heritage speakers as it is in homeland speakers? (ii) Do we see divergence only in the frequency of constructions or also in sectional properties, such as in the prepositions used? To answer these questions, I first compare the patterns identified in heritage Ambon Malay with patterns in the Ambon Malay variety spoken by homeland speakers. Then I compare the patterns of heritage speakers to those of first generation speakers in the Netherlands (who are late bilinguals), and to those of Dutch speakers (with no

⁸² There is only one token of *tot* in the dataset. This preposition is found in the phrase *tot kort* ‘short’ (lit. ‘to/until short’), used by one participant to describe the ‘cut-hair’ video-clip (see Section 6.3.3).

knowledge of Ambon Malay). Given the small sample size of these last two groups, these latter comparisons are mostly qualitative in nature. The typological similarities and differences between the Ambon Malay and Dutch resultative constructions described in Section 6.2 allow us to make the following predictions.

Following the Alternation Hypothesis (see Section 1.4.1), I hypothesize that the (partial) parallelism between Ambon Malay and Dutch will create the condition for cross-linguistic influence, and that this will manifest itself as a change in the frequency of resultative constructions in the heritage variety. In other words, heritage speakers will adapt the frequency of a construction in the heritage language to the frequency of the ‘corresponding’ construction in their dominant language. The prediction is that heritage speakers will more frequently use the constructions shared by Ambon Malay and Dutch (PPs and APs), and that they will underuse the constructions attested only in Ambon Malay (SVC and ‘two predicate construction’).

Following the Conceptual Transfer Hypothesis (see Section 1.4.2), which maintains that patterns of event construal are transferrable across languages, I expect the SVC and the ‘two predicate construction’ to pose a challenge to heritage speakers. These symmetrical constructions, in fact, reflect a typical Ambon Malay way of organizing and segmenting information (Tjia, 1997; van Staden & Reesink, 2008; Paauw, 2008), which is not found in Dutch. Now, since principles of information organization are susceptible to cross-linguistic influence (Carroll & von Stutterheim, 2003; Flecken, 2010; Bylund & Jarvis, 2011), I expect heritage speakers to follow the Dutch principles, and use constructions with one single predicate more often than homeland speakers. The findings of Chapter 5 regarding the decrease in use of the ‘two predicate construction’ lend support to this hypothesis.

Finally, I also expect divergence in the choice of prepositions. We have seen in Chapter 5 that the prepositions used by heritage speakers in PO constructions are different from those used in the homeland variety, and that the use of the typical Ambonese preposition *par* ‘to, for’ is decreasing in the Netherlands. Given the findings of Chapter 5, I hypothesize that heritage speakers will diverge from homeland speakers with respect to preposition selection in PP constructions. Furthermore, we have seen in Section 1.3.1 that lexico-semantic calques are very common in heritage languages, and that elements such as prepositions are likely to undergo semantic extension under the dominant language influence (e.g., the meaning of Spanish *para atrás* ‘behind’ was extended on the model of English *back*). In light of this finding, we may expect transfer of sectional properties in heritage

Ambon Malay, such as the mapping of the subcategorization of Dutch prepositions *in* ‘in’ and *tot* ‘to, until’ onto their Ambon Malay counterparts.

6.3.3 Participants, task, and responses

Four groups of speakers participated in the study: one test group of heritage speakers and three control groups. The test group consists of 32 heritage speakers.⁸³ The first control group is formed by 27 homelands speakers. The second control group is formed by six first generation speakers of Ambon Malay in the Netherlands.⁸⁴ The third control group is formed by ten native speakers of Dutch (see Section 2.1 for more information about the participants and data collection).

The dataset consists of descriptions elicited by means of ten video-clips, as shown in Table 6.2 (see Section 2.3 in Appendix 2). The ten video-clips were intermingled with 58 fillers (see Section 2.4 in Appendix 2).

Table 6.2: List of the ten video clips used as elicitation stimuli.

TYPE OF RESULT	LIST OF VIDEO-CLIPS
An object (Theme) is clearly divided into two or three recognizable pieces	Tear cloth into two pieces by hand
	Piece of cloth tears spontaneously into two pieces
	Slice carrot lengthwise with knife into two pieces
	Carrot snaps spontaneously into two pieces
	Cut rope stretched between two tables
	Snap twig into two pieces by hand
An object (Theme) undergoes another type of change of state	Cut fish into three pieces with sawing motion of knife
	Hack branch off tree with machete
	Cut hair with scissors
	Smash flower pot with single blow of hammer

⁸³ Moro (2014) contains 33 heritage speakers. The heritage speaker H10 has been removed from the present study and from the other studies presented in this dissertation because the Ambon Malay variety that he uses shows too much influence from Standard Indonesian.

⁸⁴ In Moro (2014), homeland speakers and first generation speakers are collapsed into one group labeled ‘baseline’. For the sake of comparability with the other chapters of this dissertation, the two groups are kept separate here.

The ten video-clips depicted resultative events of various kinds, where an inanimate object (e.g., a carrot, a twig) undergoes a change of state as a result of an activity (e.g., cut, break). Seven video-clips display a result that can be clearly identified as ‘into two (or there) pieces’, while three video-clips display other kind of results, such as ‘cut the hair short’, or ‘smash a pot’. In eight video-clips there is a human agent performing the activity, while in two video-clips the object changes its state spontaneously or ‘magically’ (e.g., a carrot snaps spontaneously into two pieces).

Every participant produced ten responses. All the responses were transcribed and entered into a separate database in Excel (see Section 2.4). Not all responses were included in the analysis, as laid out in Table 6.3. The criterion for including a response for analysis was that it contained an adequate description of the MANNER and the THEME.

Table 6.3: Summary of valid and excluded responses in the four groups.

GROUP	<i>n</i>	RESPONSES	
Heritage Ambon Malay speakers	32	Valid	312
		Excluded	8
Homeland Ambon Malay speakers	27	Valid	263
		Excluded	7
First generation Ambon Malay speakers	6	Valid	56
		Excluded	4
Dutch speakers	10	Valid	95
		Excluded	5

In the Ambon Malay resultative constructions I elicited, the most frequent MANNER verbs were *potong* ‘cut’ (used 192 times), *robe/rabe* ‘tear’ (used 110 times), *pata* ‘snap’ (used 86 times), *gunting* ‘cut hair’ (used 56 times), *pukol* ‘hit’ (used 49 times), *bela* ‘divide’ (used 26 times), *biking rusa* ‘break’ (lit. ‘make broken’) (used 17 times), *putus* ‘separate’ (used 15 times). The most frequent RESULT verbs (V₂) in SVCs were *jadi* ‘become’ (used 69 times), *bage* ‘split’ (37 times), *bela* ‘divide’ (used seven times) and *kasi pica* ‘break’ (lit. ‘give break’) (used seven times).

In the Dutch dataset, the most frequent verbs were *scheuren* ‘to tear’ (used 15 times), *slaan* ‘to hit’ (used 12 times), *afknippen* ‘to cut off’ (used ten times), *breken* ‘to break’ (used ten times), *snijden* ‘to cut’ (used ten times), *afsnijden* ‘to cut off’ (used seven times), *afhakken* ‘to chop off’ (used six times), *doorbreken* ‘to break through’

(used three times) and *doorsnijden* ‘to cut through’ (used three times). Responses that were excluded did not contain an adequate description or contained change of state verbs, such as Ambon Malay *bengkol* ‘curve’, *biking barsi* ‘clean’ (lit. ‘make clean’), or Dutch *fileren* ‘to remove the bones’, that do not belong to the class of *cutting*, *breaking* and *hitting* verbs (see Levin, 1993).

6.4 Results and discussion

This section presents and discusses the results of the experiment. In each section, I discuss the results and propose explanations for the patterns observed.

6.4.1 Frequency of resultative constructions

In this section, I analyze the similarities and differences between the four groups with respect to the various types of constructions that they use to express resultative events.

6.4.1.1 Results

The set of constructions used by homeland speakers and heritage speakers is the same, but the frequency is different in the two groups. Figure 6.1, on the next page, lays out the results regarding the frequency of the resultative constructions that are attested in the Ambon Malay and the Dutch data (see Section 6.2).⁸⁵ It shows that, with a decrease of symmetrical strategies and an increase of satellite strategies, heritage Ambon Malay lies in between homeland Ambon Malay and Dutch.

SVCs (diagonal lines bar) are used significantly less frequently by heritage speakers ($M=10.03$, $SD=16.84$) than by homeland speakers ($M=.3952$, $SD=.158$) ($t(57)=-6.886$, $p<.001$, $r=.67$, equal variances assumed). In the homeland and first generation groups, SVCs are the preferred strategy and every homeland or first generation speaker provided at least one response containing a SVC. In the heritage group, on the contrary, SVCs occur only in about $\frac{1}{3}$ of the speakers (12 out of 32). A similar picture emerges with respect to the ‘two predicate construction’ (black bar),

⁸⁵ The percentages do not reach 100%, because in some cases speakers described the result of the event by means of other strategies, such as periphrasis or reformulation. These strategies were coded as ‘other’, but are not included in Figure 6.1 due to lack of space.

which is used significantly less frequently by heritage speakers ($M=.0547$, $SD=.082$) than by homeland speakers ($M=.2233$, $SD=.146$) ($t(39.346)=-5.327$, $p<.001$, $r=.64$, equal variances not assumed).

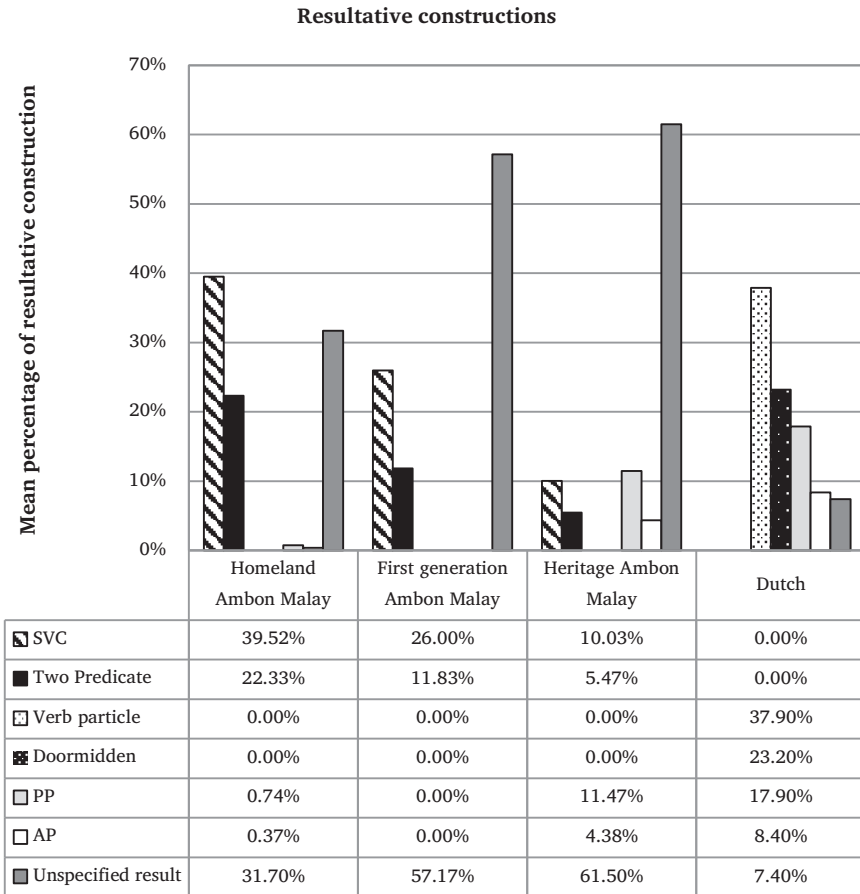


Figure 6.1: The mean percentages of the types of resultative constructions attested in the four datasets.

PPs (light gray bar) and APs (white bar) are more frequent in the heritage group than in the homeland group. In the homeland group, resultative PPs account for 0.74% of all responses ($SD=.038$), and the only two instances attested were provided by the same speaker. In heritage Ambon Malay, the frequency of PPs shows a significant increase ($M=.1147$, $SD=.128$) ($t(37.403)=4.478$, $p<.001$, $r=.59$, equal variances not assumed), and more than half of heritage speakers (17 out of 32)

provided at least one PP construction. APs are also used significantly more often by heritage speakers ($M=.0438$, $SD=.084$) than by homeland speakers ($M=.0037$, $SD=.019$) ($t(34.816)=2.617$, $p=.013$, $r=.40$, equal variances not assumed), while there is only one token of AP in the homeland group, in the heritage group ten speakers used an AP at least once. With respect to the percentage rate of PPs and APS, heritage Ambon Malay is clearly in the middle (11.4% and 4.3%) between homeland Ambon Malay (0.74% and 0.3%) and Dutch (17.9% and 8.4%). Heritage speakers are more likely to leave out any further specification of the RESULT ($M=.6150$, $SD=.209$), when compared to homeland speakers ($M=.3170$, $SD=.160$) ($t(57)=6.037$, $p<.001$, $r=.62$, equal variances assumed). First generation speakers also present a high frequency of ‘unspecified’ results.

Dutch speakers prefer to encode the RESULT by means of verb particles (black dots bar). The adverb *doormidden* ‘through, in half’ (white dots bar) is the second preferred strategy. The relatively high frequency of *doormidden* compared to PP is probably a task effect because in six video-clips out of ten the object is divided into two approximately equal sub-parts (see Section 6.2.2). PPs and APs account for 17.90% and 8.40% of the responses, respectively. Finally, the relatively low frequency of ‘unspecified results’ relates to the use of verb particles. Dutch speakers prefer to use particle verbs (e.g., *af-snijden* ‘to cut off’) rather than a simple verb (e.g., *snijden* ‘to cut’) when describing resultative events. This choice lowers the percentage of ‘unspecified results’ responses in the Dutch dataset.

6.4.1.2 Discussion

The data presented in Section 6.4.1.1 show that the resultative constructions used by heritage speakers are qualitatively similar but quantitatively different from those of homeland speakers. Three main results emerge from the quantitative analysis of the data: (i) symmetrical strategies (SVC and ‘two predicate construction’) are used significantly less by heritage speakers, (ii) satellite strategies (PP and AP) are used significantly more by heritage speakers, (iii) heritage speakers leaves the RESULT unspecified more often than homeland speakers. I discuss these three findings here in turn.

The first finding is that symmetrical constructions are used less frequently by heritage speakers than by homeland (and first generation) speakers, as shown in Figure 6.1. In other words, homeland speakers are more likely to encode the MANNER and the RESULT in two verbal elements. In this respect, homeland and first generation speakers of Ambon Malay resemble speakers of other eastern Malay

varieties currently spoken in Indonesia, all of which allow complex events to be expressed through two or more consecutive verbs (Paauw, 2008, pp. 232-236). Segmenting the flow of information over two or more predicates is a typical pattern of Ambon Malay, and it is used in the expression of various types of events, including resultatives, give-events, and motion events (Tjia, 1997; van Staden & Reesink, 2008; Chapter 5 in this dissertation). Heritage speakers, by contrast, show a much lower incidence of symmetrical constructions, and use constructions with only one predicate (encoding the MANNER) with a higher frequency. In this respect, they resemble more Dutch speakers who also express resultative events by means of a single verbal element. This shift in preference patterns reflects different principles of event construal. The principles of information organization are language-specific and determine how the body of information is segmented and structured and which linguistic means are selected to convey the message (Slobin, 1991; Carrol & von Stutterheim, 2003; Bylund & Jarvis 2011). The results of this study suggest that heritage speakers tend to organize information via their dominant language, Dutch, and this has an effect also on the grammatical means they select (less SVCs, more satellite elements). This finding, together with the finding of Chapter 5 on *give*-constructions (see Section 5.4.1), lend support to the Conceptual Transfer Hypothesis, which maintains that patterns of event construal are transferrable across languages. That heritage Ambon Malay may be undergoing a ‘deserialization’ process is supported by external evidence from another serializing language, Tetun Dili. We have seen in Section 6.1 that Tetun Dili is an Austronesian serializing language in contact with Portuguese, a non-serializing language. Hajek (2006) states that Tetun Dili is subject to an ongoing process of substantial deserialization, and in his view “the most significant factor in deserialization is the effect of long-term contact with Portuguese” (p. 251). The finding of Hajek (2006) and the finding of the present study strongly suggest that the loss of SVCs can be contact-induced (see Aikhenvald, 2006, p. 53).

The second finding is that heritage speakers of Ambon Malay have a stronger preference for the Dutch-aligned ways of expressing resultative events (satellite constructions). The increase in the frequency of resultative PPs and APs in heritage Ambon Malay may be regarded as an innovation due to cross-linguistic influence, because although these structures (‘VERB - THEME - PP/AP’) were already present in the language, they are increasingly extended to the resultative domain. Thus, what is transferred from the dominant language is the frequency of a construction and not its structure. We have already seen in the previous chapters of this

dissertation that this type of ‘indirect transfer’ is the one of the most common mechanisms underlying contact-induced change in heritage bilingual communities (see Section 1.3.1.1). In this type of change, transfer from the dominant language does not involve the creation of a new structure, but rather a change in the frequency of a structure that already exists in the heritage language. Indirect transfer is possible, because heritage speakers establish an equivalence between the Dutch and the Ambon Malay construction. In psycholinguistics terms, this equivalence relation allows cross-language activation in the bilingual mind: every time that a PP or AP construction is activated in the dominant language (Dutch), it automatically reinforces the corresponding construction in the heritage language (Ambon Malay). In other words, while SVCs receives their degree of entrenchment only from one source (the heritage language), PPs and APs receive their degree of entrenchment from two sources (the heritage language and the dominant language), and are therefore more likely to be selected in language production (see Section 7.2.2). In this regard, it may be interesting to note that not only the frequency of PPs and APs is similar between heritage Ambon Malay and Dutch but also their distribution. For instance, the video-clip ‘snap twig into two pieces’ triggered a high number of PPs both in Dutch (50%) and in heritage Ambon Malay (23%), while the video-clip ‘smash flower pot’ triggered a high number of APs both in Dutch (80%) and in heritage Ambon Malay (15%). This is explained by the fact that the ‘snap twig into two pieces’ video-clip was described with the same lemma in Ambon Malay (*pata* ‘snap, break’) and Dutch (*breken* ‘to break’), while the second video-clip was described with the lemma ‘hit’ in Ambon Malay (*pukul* ‘hit’) and Dutch (*slaan* ‘to hit’). Notably, *pata* and *breken* ‘(to) break’, and *slaan* and *pukul* ‘(to) hit’ are translation equivalents. They share not only similar conceptual properties, but also similar grammatical structures (PP/AP). Whenever *pata* or *pukul* are used, their equivalents *breken* and *slaan* are also activated and with them the grammatical structures that they are associated with. Thus, the co-activation of the Dutch equivalent triggers the use of PP/AP with the Ambon Malay lemma.

The third finding is that heritage speakers tend to leave the RESULT unspecified more often than homeland speakers and Dutch speakers. This finding can be accounted for by two processes. The first and the more logical explanation is that heritage speakers leave out the result component more often due to universal regression processes or simplification under reduced input conditions, rather than influence from Dutch. As shown in Section 1.3.5, universal principles of language development in the context of language disuse, such as preference simplification,

regularization, naturalness and universal principles of human communication, represent an important cause of divergence between the heritage grammar and the homeland grammar. However, the increase of ‘unspecified results’ in heritage Ambon Malay may well be indirectly related to Dutch influence. Dutch speakers often encode the result in verb particles or by means of *doormidden* ‘through, in half’. These types of satellite elements are not available in Ambon Malay. Thus, it may be the case that, whereas Dutch influence activates the procedure ‘satellite’ more than the procedure ‘verb’, the instantiation of the satellite is blocked by the fact that the particular elements are not available in the system, consequently the RESULT remains unspecified. This suggests that structural transfer is constrained by the structure of the heritage language and that transfer effects are blocked when structural parallelism between the two languages is lacking (structural factors constraining cross-linguistic influence are further discussed in Section 7.2).

To summarize, the decrease in the use of symmetrical constructions and the increase in the use of satellite elements are indicators of a change in frequency rather than a dramatic categorical change. The Ambon Malay structures that are not shared by Dutch, such as SVCs, are fading away and are replaced by structures shared by both languages, such as PPs. This change in frequency is leading to an increased similarity between heritage Ambon Malay and Dutch, with heritage speakers preferring the Dutch-aligned ways of expressing resultative events.

In arguing about the degree of restructuring of heritage speakers’ grammars it is important to consider individuals’ data, since heritage speakers have variable language backgrounds. I therefore investigated whether there was a relation between the amount of exposure that individuals had to Ambon Malay in the course of their lifetime and their use of the SVC, PP and AP. I divided the heritage speakers into two groups according to the exposure to Ambon Malay: the ‘LOW exposure’ group vs. the ‘MEDIUM-HIGH exposure’ group (see Table 6.4 on the next page). From the information presented in Table 6.4, three observations can be made. First, most of the heritage speakers who use SVCs (shaded gray) belong to the ‘MEDIUM-HIGH exposure’ group. In this group more than half of the speakers (seven out of thirteen) used a SVC at least once, whereas in the ‘LOW exposure’ group only $\frac{1}{4}$ of the speakers did (five out of nineteen). Furthermore, all but one of the speakers who used a SVC live in a Moluccan ward. We see in Chapter 7 (Section 7.3.1) that the place where the speaker lives seems to be the best predictor of SVC use. Second, the speakers who used a PP or an AP construction (printed in bold and underlined, respectively) belong to both groups. About half of the speakers in each group (seven

out of thirteen in the ‘MEDIUM-HIGH exposure’ group and ten out of nineteen in the ‘LOW exposure’ group) used a PP, thus no factor seems to strongly predict the use of PPs. APs occur slightly more in the ‘MEDIUM-HIGH exposure’ group (six out of thirteen) than in the ‘LOW exposure’ group (three out of nineteen). These small numbers, however, do not allow any robust conclusion. The third observation is that overall speakers who use a resultative SVC do not use a resultative PP or AP, and *vice versa* (H5, H12, H13 are the only exceptions for PPs, and H4 and H6 are the exceptions for APs). The fact that the same speaker typically does not use both of the constructions confirms that SVC are a conservative feature which is typical of the homeland variety, while PPs and APs are innovative Dutch-like features.

Table 6.4: The sociolinguistic background of the heritage Ambon Malay group. Speakers using SVCs are shaded gray, speakers using PPs are printed in bold, and speakers using APs are underlined.

	Sp	AoA	L Mo	L Fa	L Sb	L Pa	LIVES	GREW	HC
LOW EXPOSURE TO AMBON MALAY	H30	0	Dutch	Dutch	Dutch	Dutch	city	city	2
	H11	0	Dutch	Dutch	Dutch	Dutch	city	city	4
	H32	0	Dutch	Dutch	Dutch	Dutch	city	city	6
	H31	0	AM	Dutch	Dutch	Dutch	city	city	0
	H14	0	Dutch	Dutch	Dutch	Dutch	city	city	4
	H33	0	AM	Dutch	Dutch	Dutch	city	city	2
	H17	0	Dutch	AM	Dutch	Dutch	city	city	2
	H27	0	Dutch	Dutch	Dutch	Dutch	<i>wijk</i>	city	0
	<u>H23</u>	<u>0</u>	<u>Dutch</u>	<u>Dutch</u>	<u>Dutch</u>	<u>Dutch</u>	<u>wijk</u>	<u>city</u>	<u>3</u>
	<u>H21</u>	<u>0</u>	<u>Dutch</u>	<u>Dutch</u>	<u>Dutch</u>	<u>Dutch</u>	<u>wijk</u>	<u>city</u>	<u>4</u>
	H13	0	Dutch	Dutch	Dutch	Dutch	<i>wijk</i>	<i>wijk</i>	0
	H29	0	Dutch	AM	Dutch	Dutch	city	city	5
	H25	0	Dutch	AM	Dutch	Dutch	<i>wijk</i>	<i>wijk</i>	0
	H19	0	Dutch	Dutch	Dutch	Dutch	<i>wijk</i>	<i>wijk</i>	1
	H20	0	mixed	Dutch	Dutch	-	<i>wijk</i>	city	6
	<u>H22</u>	<u>0</u>	<u>Dutch</u>	<u>Dutch</u>	<u>Dutch</u>	<u>Dutch</u>	<u>wijk</u>	<u>camp</u>	<u>3</u>
	H8	0	Dutch	Dutch	Dutch	mixed	<i>wijk</i>	city	6
	H7	0	mixed	mixed	Dutch	Dutch	<i>wijk</i>	<i>wijk</i>	1
<u>H6</u>	<u>0</u>	<u>Dutch</u>	<u>AM</u>	<u>mixed</u>	<u>mixed</u>	<u>wijk</u>	<u>city</u>	-	

The table continues on the next page.

	Sp	AoA	L Mo	L Fa	L Sb	L Pa	LIVES	GREW	HC
	H26	0	AM	AM	mixed	Dutch	<i>wijk</i>	city	3
	H15	0	AM	AM	Dutch	mixed	<i>wijk</i>	city	3
	H3	0	AM	Dutch	mixed	AM	<i>wijk</i>	camp	2
LOW EXPOSURE TO AMBON MALAY	H28	> 4	AM	AM	mixed	Dutch	city	camp	4
	H24	> 4	AM	AM	mixed	Dutch	city	camp	10
	H2	> 4	AM	AM	mixed	Dutch	<i>wijk</i>	camp	-
	H9	> 4	AM	AM	mixed	Dutch	<i>wijk</i>	<i>wijk</i>	5
	H16	> 4	AM	AM	mixed	Dutch	<i>wijk</i>	camp	2
	H18	> 4	AM	AM	AM	Dutch	<i>wijk</i>	<i>wijk</i>	3
	H5	> 4	AM	AM	mixed	Dutch	<i>wijk</i>	camp	6
	H4	> 4	AM	AM	mixed	mixed	<i>wijk</i>	camp	3
	H1	> 4	AM	AM	AM	Dutch	<i>wijk</i>	camp	10
	H12	> 4	AM	AM	mixed	AM	<i>wijk</i>	camp	8

In sum, the information on the language history of the speakers suggests that high exposure to Ambon Malay corresponds with the use of the SVCs, a feature typical of the homeland variety; the picture that emerges for PPs and APs is less clear, as high vs. low exposure do not seem to strongly correlate with these features (see Section 7.3 for more information).

6.4.2 Preposition selection

In the Ambon Malay data, six prepositions are used to encode the RESULT in a PP construction (as described in Section 6.2.1). This section reports the qualitative differences in preposition selection, and offers an explanation for the attested variation.

6.4.2.1 Results

The use of prepositions in the PP construction by homeland speakers and heritage speakers is presented in Table 6.5, on the next page. First generation speakers are not included because they did not use the PP construction. Homeland speakers seem to use only *par* ‘for, to’ (*par dua* ‘into two’). This claim needs to be taken with caution because there are only two instances of PP construction in the homeland dataset, both provided by the same speaker. Heritage speakers display great variation: *ka* ‘to’ is the preposition selected more frequently (*ka dua* ‘into two’, used

by ten speakers), followed by *dalang* ‘in(side)’ (*dalang dua* ‘into two’, or *dalang dua bagean* ‘into two pieces’, used by six speakers), *di* ‘at, in, on’ (*di dua* ‘into two’, used by two speakers), *in* ‘in’ (Dutch loan) (*in dua* ‘into two’, used by two speakers), and *sampe* (*sampe dua* ‘into two’, used by one speaker). Dutch speakers use only the preposition *in* ‘in’ (17 tokens) and *tot* ‘to, until’ (one token). Although statistical test could not be performed due to the paucity of tokens, Table 6.5 shows that with respect to preposition selection, heritage speakers differ qualitatively from homeland speakers.

Table 6.5: The number of tokens and within group percentages of the different prepositions attested in the PP resultative construction.

PREPOSITION	GROUP	<i>n</i>	TOKEN	% WITHIN GROUP
<i>par</i>	Homeland	27	2	100%
	Heritage	32	0	0
<i>ka</i>	Homeland	27	0	0
	Heritage	32	20	55.5%
<i>dalang</i>	Homeland	27	0	0
	Heritage	32	10	27.8%
<i>di</i>	Homeland	27	0	0
	Heritage	32	3	8.3%
<i>in</i>	Homeland	27	0	0
	Heritage	32	2	5.6%
<i>sampe</i>	Homeland	27	0	0
	Heritage	32	1	2.8%

6.4.2.2 Discussion

Homeland speakers seem to prefer *par* ‘for, to’, to encode resultative events. This finding, although preliminary, is in line with the findings of Chapter 5 which show that homeland speaker also have a strong preference for *par* ‘for, to’ in the expression of *give*-events (see Table 5.3 in Chapter 5). Although this typical Ambon Malay morpheme was brought along by Moluccan immigrants to the Netherlands, and was thus part of the input of heritage speakers, it is not used very frequently in

the Moluccan community in the Netherlands, where other prepositions are preferred (see Section 5.4.2).

Heritage speakers seem to prefer locative prepositions, such as *ka* ‘to’ or *dalang* ‘in(side)’, to introduce the RESULT in a PP (see Table 6.5). In this respect, they resemble Dutch speakers who also use the locative preposition *in* ‘in, into’ and the directional preposition *tot* ‘to, until’, in resultative PPs. In Dutch, however, *in* ‘in, into’ and *tot* ‘to, until’ do not have only a spatial meaning, *in* can also indicate path or part-whole relationships (Cuyckens, 1993), and *tot* can indicate extension of an activity up to a certain point (Zwarts, 1995). We may therefore hypothesize that heritage Ambon Malay prepositions *ka* ‘to’ and *dalang* ‘in(side)’ are undergoing semantic extension under the influence of their Dutch counterparts (*tot* ‘to, until’ and *in* ‘in, into’, respectively), and that their subcategorization is changing from purely locative preposition (van Minde, 1997, p. 75) to more abstract form.⁸⁶ This change, however, is more advanced for *ka* ‘to’ than for *dalang* ‘in(side)’ (see Table 6.5). This finding is somehow counter intuitive at a first glance, because since the Dutch preposition *in* is much more frequent than *tot* (in the dataset), we would expect its lexico-syntactic Ambon Malay equivalents *dalang* ‘in(side)’ or *di* ‘at, in, on’ to be also more frequent. The preference for *ka* ‘to’ over the ‘Dutch-equivalent’ *dalang* ‘in(side)’ is probably accounted for by the fact that *ka* ‘to’ is more prone to semantic extension, while *dalang* ‘in(side)’ still preserves a strong locative meaning.

The innovative use of the preposition *ka* ‘to’ in the expression of resultative events reflects a cross-linguistically common pattern, whereby spatial preposition are extended to indicate non-spatial meanings (Bowerman, 2011). According to Enfield (2003), semantic extensions that are very common cross-linguistically are driven by universal principles that he subsumes under the label ‘conceptual naturalness’ (see Section 1.3.5). Rice and Kabata (2007) have shown that there is a general tendency for allative prepositions to develop other senses, including Beneficiary and Resultative. For instance, the English preposition *to* can have a locative (*walk to school*), dative (*give to him*), and resultative (*break to pieces*) meaning. We have seen in Chapter 5 that the Ambon Malay preposition *ka* ‘to’ has already been extended to indicate various non-spatial senses, such as the Recipient/Beneficiary in *give*-constructions. The semantic extension of *ka* ‘to’ from

⁸⁶ A similar change is described by Schoenmakers-Klein Gunnewiek (1997), who reports that heritage speakers of Dutch in Brazil often impose Brazilian-Portuguese distribution to Dutch functional items. One example is the Dutch preposition *op* ‘on, at’ which has changed its subcategorization to resemble the Brazilian-Portuguese *em* ‘in, on, at’.

locatives to datives and resultatives reflects a natural conceptual extension whereby places, persons and states are all construed as destinations (Rice & Kabata 2007, p. 497).

The semantic extension of *ka* ‘to’ and the gradual shift in the subcategorization of *ka* ‘to’ and *dalang* ‘in(side)’ are language internal processes accelerated by contact with Dutch. By selecting resultative PPs, heritage bilingual speakers allow their heritage language to converge toward the dominant language, but at the same time they do not simply replicate the equivalent Dutch preposition into the Ambon Malay PP construction. According to Matras (2009, p. 243), contact-induced change in the distribution of structural elements, such as the ones we have seen above, “derives from an individual speaker’s scan for an optimal construction through which to communicate the local meanings”. In other words, bilingual speakers scan the entire repertoire to identify the structure of the heritage language that conveys the intended meaning more effectively. Although Matras (2009) perhaps emphasizes too much the role of agentivity of the speaker, he is right in saying that convergence is not a completely ‘mechanic’ process but it involves some degree of creativity. If heritage speakers were only trying to replicate the Dutch PP construction (*in tweeën* ‘into two’), we would mostly find PP introduced by *dalang* ‘in(side)’, but this is not the case. The variation between different prepositions within and among speakers indicates that speakers ‘consciously’ control the choice of prepositions. It seems to me that they try to select the preposition that better matches the Dutch equivalent, while controlling for the meaning (and the possible meaning extension) on the preposition selected. This explains why they find *ka* ‘to’ more suitable than *dalang* ‘in(side)’, although the latter corresponds more closely to Dutch *in* ‘in’.

6.5 Conclusions

The domain of resultative construction in heritage Ambon Malay is undergoing substantial restructuring under the influence of Dutch. The changes in the frequency of SVCs and ‘two predicate constructions’ on the one hand, and of PPs and APs on the other hand, lead to structural divergence from the homeland variety and to structural convergence toward Dutch.

The first change in frequency involves a decrease in the use of symmetrical constructions, where the MANNER and the RESULTS are expressed by means of a SVC or two verbal predicates. I argue that this is due to different information

organization principles. Heritage speakers tend to organize information following the principles of their dominant language, Dutch, where resultative events are expressed by means of a single verbal element. The finding of this study reinforces Hajek's (2006) findings that the loss of SVCs or 'deserialization' in favor of PPs can be a contact-induced process.

The second change in frequency involves an increase in the frequency of resultative PPs and APs, which are becoming more entrenched and therefore more productive. The increase in the frequency of resultative PPs and APs may be regarded as an innovation due to cross-linguistic influence because these structures are increasingly extended to the resultative domain. However, the structure ('VERB - THEME - PP/AP') is not an innovation. Thus, what is 'transferred' from the dominant language is the frequency of the construction and not its structure.

The third, this time qualitative, change described in this chapter concerns the use of different prepositions in the PP construction. While homeland speakers seem to prefer *par* 'to, for' for resultative, as well as for *give*-events, heritage speakers have different preferences. The prepositions *ka* 'to' and *dalang* 'in(side)' are the two most frequent alternatives. The use of these two prepositions in resultative events indicate that in heritage Ambon Malay the subcategorization of *ka* 'to' and *dalang* 'in(side)' is changing from purely locative to more abstract. It is likely that the semantic extension of *ka* 'to' and *dalang* 'in(side)' represents a language internal processes accelerated by contact with Dutch.

Finally, exposure to Ambon Malay seems to have some effects on the use of SVCs, but not on the use of PPs and APs. SVCs are mostly used by speakers of the 'MEDIUM-HIGH exposure' group, while PPs and APs are used by speakers of both the 'MEDIUM-HIGH exposure' and the 'LOW exposure' groups.

CHAPTER 7

Divergence and convergence in heritage Ambon Malay: structural and social factors

7.1 Introduction

What are the structural and social mechanisms and factors responsible for the patterns of divergence (from the homeland variety) and convergence (toward Dutch) observed in various areas of heritage Ambon Malay grammar? This chapter brings together the findings of the previous chapters to answer this question. In the scenario approach to language contact, every change is seen as both structurally and socially embedded (Muysken, 2010b, p. 272). The outcome of language contact is thus determined both by structural and social factors: structural factors, such as the typological profile of the two languages, supply the material for the change and determine the shape that the change is going to take; social factors determine the amount and the destination of the change (Thomason & Kaufman, 1988; Croft, 2000; Thomason, 2001; Winford, 2003; Johanson, 2013, among many others).

Most of the changes which are reported in dissertation are ‘on-going’ changes involving a *change in frequency* between already available structures. *Changes in frequency* are considered a symptom of language contact as much as lexical borrowings, calques or *contact-induced grammaticalization* (Backus, 2004; Boumans, 2006; Johanson, 2002; Alferink, 2015; Onar Valk, 2015). Frequency is, in fact, a crucial component of language change, as it reflects the synchronic variation typical of ‘on-going’ changes. In the language change continuum, the first step or level of change consists of ‘spontaneous innovations’ or ‘momentary cases of interferences’ at the individual level; when the innovation propagates within the community, the change is ‘on-going’ and considerable intra- and inter- speaker variation is found; when the innovation has stabilized and it is shared by all speakers, then the change is completed (Croft, 2000).

Every change starts or has started in the mind of a bilingual individual, after all, languages come into contact in the bilingual mind and not elsewhere. Therefore, every change is, at least in its initial stage, psychologically motivated. Knowing how cross-linguistic influence operates at the psycholinguistic level can help us in

understanding how the mechanisms of language contact (i.e., change in frequency, grammatical re-analysis, contact-induced grammaticalization) operate at the community level (Muysken, 2010b; 2013; Hartsuiker, 2013). Hartsuiker (2013, p. 738) has recently pointed out that:

To the extent that groups of people (e.g., in particular regions) have a similar profile of L1 and L2, such mechanisms at the individual level can affect distribution properties [(e.g., frequency of a particular structure)] at the level of groups of speakers. An interesting prospect, then, is that basic principles from one level of description (the individual speaker) can explain outcomes at a higher level of description (a group of speakers).

With this idea in mind, in the next section, I will discuss the results of the previous chapters by integrating Hartsuiker et al.'s (2004) model of cross-linguistic activation into a broader usage-based account of language contact (see also Moro & Irizarri van Suchtelen, forthcoming).

This chapter is organized as follows. Section 7.2 recounts the major findings of this dissertation and discusses the mechanisms and the (structural) factors which play a role in shaping heritage Ambon Malay grammar. With the help of cluster analysis, the patterns of convergence and divergence are analyzed in order to establish meaningful groups of features. The results of the cluster analysis show that two macro-groups of features can be identified, the 'Dutch-like' or 'innovative' features and the 'Malay-like' or 'conservative' features. This grouping of features lends support to the idea that changes in heritage Ambon Malay grammar result from the same mechanisms. Section 7.3 focuses on social factors. If structural factors determine which structures will be selected among those available, it is social factors that determine which individuals will be the initiators and propagators of change in the community. After giving an overview of the major social factors responsible for change in heritage communities, Section 7.3.1 explores the effects of social factors on the linguistic behavior of heritage speakers. The results of the ANOVA tests show that the place where the speaker lives seems to be the best predictor of language change, followed by the exposure to Ambon Malay during childhood.

7.2 Structural factors and language change

The previous chapters of this dissertation have illustrated and discussed a number of structural changes that heritage Ambon Malay is undergoing in the Netherlands.

These changes are more pervasive than ‘momentary cases of interferences’ but have not reached the final stage of completion yet, as witnessed by the fact that there is considerable variation among speakers. Nevertheless, some major patterns can be individuated.

The main, and perhaps the most important, outcome of language contact between Dutch and Ambon Malay is convergence, which manifests as a *change in frequency or preference* between structures already present in Ambon Malay (see Section 1.3 and Section 1.3.1.1). This *change in frequency* leads to greater structural similarity between Ambon Malay and Dutch, as bilingual speakers select the Ambon Malay structure equivalent to the Dutch structure more often. Other terms used in the literature for this type of change are ‘indirect transfer’ (Silva-Corvalán, 1994), ‘frequential copying’ (Johanson, 2002), ‘redistribution’ and/or ‘shift in proportion’ (Alferink, 2015). We have seen in Section 1.3.1.3 that changes that ‘only’ affect the frequency of a structure are among the most common types of changes attested in bilingual (heritage) communities and that, although these changes do not deeply alter the system, they can lead to significant structural changes (Backus, 2004). Clear cases of *change in frequency* are illustrated in Chapter 3, Chapter 5 and Chapter 6.

Chapter 3 examines possible effects of language contact on the word order of nominal modifiers (demonstrative, numerals, adjectives, and the definite marker). The chapter shows that when Ambon Malay allows an alternation between two word order patterns (demonstratives and numerals can precede or follow the noun), heritage bilingual speakers prefer the pattern shared by Dutch (i.e., pre-nominal demonstratives, and pre-nominal numeral), but when Ambon Malay has no such alternative (adjectives can only follow the noun), cross-linguistic influence from Dutch does not occur. Chapter 3, therefore, reports a *change in frequency or preference*: whereas monolingual in Ambon prefer post-nominal modifiers, heritage speakers in the Netherlands prefer pre-nominal modifiers. Chapter 5 and Chapter 6 report similar results in the domain of *give-* and resultative constructions, respectively. Chapter 5 reports a changes in the frequency of two constructions, the ‘Double Object construction’ (*John gave Mary a book*) and the ‘two predicate construction’ (*John holds the book and gives (it) to Mary*). Both constructions are attested in Ambon Malay, but the former is also found in Dutch, whereas the latter is not. Contrarily to homeland speakers, heritage speakers use the DO construction with a higher frequency, and the ‘two predicate construction’ with a lower frequency. Chapter 6 illustrates a change in the frequency of the constructions

dedicated to the expression of resultative events. Ambon Malay prefers serial verb constructions (*She breaks a stick becomes two*) or ‘two predicate construction’ (*She hits the pot, the bot breaks*), while Dutch prefers verb particles (*She cuts off a branch*), but both languages also allow preposition phrases (*She breaks a stick in two*) and adjectival phrases (*She hits a vase broken*). Heritage speakers use the SVCs and the ‘two predicate construction’ significantly less than homeland speakers, however they use PPs and APs, the constructions shared with Dutch, significantly more. To sum up, Chapter 3, Chapter 5 and Chapter report a shift in heritage speakers away from the preferences of homeland speakers, and towards those of Dutch speakers. This *change in frequency* is slowly but surely leading to greater structural convergence between heritage Ambon Malay and Dutch.

Notably, convergence can be brought about by different preferential tendencies of bilingual speakers, but it can also be the result of grammatical replication and/or *contact-induced grammaticalization*, a well-known process whereby bilingual speakers replicate a prominent (obligatory) category of one language using the ‘linguistic material’ of another language (Heine & Kuteva, 2005; Backus et al., 2011). At the basis of *contact-induced grammaticalization*, there is the psycholinguistic process of ‘functional convergence’ as described by Sánchez (2004, 2006). In this process, a functional feature of the dominant language is re-associated to a morphological unit of the heritage language. Thus, *contact-induced grammaticalization* always entails a certain degree of convergence to the dominant language (see Section 1.4.1). Like *change in frequency* discussed above, *contact-induced grammaticalization* usually involves increase in the frequency of an item (see Section 1.3.1.3), but unlike a ‘pure’ *change in frequency*, it also involves semantic bleaching and functional expansion (Heine & Kuteva, 2005, 2007). Furthermore, unlike *change in frequency*, *contact-induced grammaticalization* can ‘deeply’ alter the system (‘system-altering change’) by adding a new (sub-)category to the language system.

Chapter 3 and Chapter 4 report cases of incipient *contact-induced grammaticalization*, that is, grammaticalization in its embryonic stage. The hypothesis behind the chapters is that bilingual heritage speakers tend to overtly express the linguistic categories that are grammaticalized, and therefore obligatory in their dominant language, when they are also speaking the heritage language (see the Functional Convergence Hypothesis in Section 1.4.2). Chapter 3 reports an increase in the frequency of the definite marker =*nya*, which is arguably due to contact with Dutch. Since the category of definiteness is highly salient in Dutch and (in)definite marking is obligatory, it is expected that heritage speakers will try to replicate this

linguistic category with the available material. Some heritage speakers, in fact, use this form consistently to mark already mentioned nouns, following a pattern typical of Dutch. In this stage, we only observe an increase in the frequency, or obligatoriness, of =*nya* 'DEF', but we cannot say whether the increase in frequency is accompanied by greater semantic generality. Chapter 4 shows that, in heritage Ambon Malay, the progressive marker *ada* matches the frequency and the distribution of Dutch present tense. Given this similarity and an analogue development in Sri Lanka Malay (another heritage variety), the chapter argues that *ada* is undergoing a contact-induced grammaticalization process, whereby heritage speakers have selected the Ambon Malay element *ada* to overtly express the feature of present tense and, possibly, also of finiteness (see Section 7.2.2).

Finally, other factors, such a different type of input heritage speakers have been exposed to (see Section 1.3.4), or universal principles of language development in contexts of language disuse (see Section 1.3.5), can act separately, or cumulatively with cross-linguistic influence, and bring about divergence from the homeland variety and convergence toward the dominant language. Chapter 3 illustrates that two innovations of heritage Ambon Malay (higher frequency of DEMONSTRATIVE-NOUN order and of =*nya* 'DEF') were probably already present in the language variety that heritage speakers were exposed to (the language of first generation speakers). The chapter concludes that the increase in the frequency of DEMONSTRATIVE-NOUN order and of =*nya* 'DEF' in heritage speakers is an internal change (due to a different type of input) accelerated by contact with Dutch. Divergence from the homeland variety caused by a different type of input is also covered in Chapter 5, which illustrates that heritage speakers select different prepositions from homeland speakers to encode the Recipient argument of *give-events*'. This difference is a reflection of the different social histories of the speakers and the different input they received in childhood. Universal principles of language development in contact situations have been found to play a role in the selection of DO constructions, as discussed in Chapter 5. The chapter argues that there is something iconic about DO constructions that contributes to the selection of this feature among heritage speakers.

On the basis of the findings of the previous chapters, the following structural constraints and tendencies seem to hold for heritage Ambon Malay. These tendencies are in line with similar strategies found in bilingual speakers.

- a. Maximizing the compatibility or surface similarity between the dominant language and the heritage language: if the heritage language allows an alternation between two or more structures, the structure which more closely matches the equivalent structure in the dominant language is more likely to be selected (Muysken, 2013).
- b. Convergence is constrained by the structure of the heritage language. If structural parallelism between the heritage language and the dominant language is lacking, convergence is not likely to occur (Silva-Corvalán, 1993, 1994, 2008).
- c. Replicating obligatory grammatical categories of the dominant language using linguistic ‘material’ of the heritage language (incipient contact-induced grammaticalization) (Backus et al., 2011).

We have seen that most of the innovative features found in heritage Ambon Malay follow the above mentioned tendencies. The next section illustrates similarities and differences among these innovative features by using hierarchical cluster analysis.

7.2.1 Cluster analysis

Cluster analysis groups objects, in this case linguistic features, into clusters based on pairwise distances. The aim of cluster analysis is to classify linguistic features in order to obtain groups (clusters) of features. Similarities or differences between features can provide additional evidence regarding their origin. The assumption is that features that cluster together may have been brought about by the same mechanisms. On the basis of the findings of Chapter 3, Chapter 4, Chapter 5 and Chapter 6, I selected the 12 linguistic features that proved to be representative of actual ‘on-going’ changes and were of statistical significance.⁸⁷ These features represent changes in the domain of word-order (preference for pre-nominal *itu* ‘D.DIST’ and *satu* ‘one’),⁸⁸ definite marking (higher frequency of =*nya* ‘DEF’), aspect marking (overuse use of *ada* ‘EXIST’, underuse of *su* ‘PRF’ and reduplication ‘ITER,

⁸⁷ For the sake of comparability, the scores were normalized to a scale from 0.0 to 1.0.

⁸⁸ The demonstrative *ini* ‘D.PROX’ was not included due to the high number of missing values. In the case of *itu* ‘D.DIST’ and *satu* ‘one’, the missing values (two in each case) were replaced with the average of the group.

INTENS’) and in specific constructions (higher frequency of DOs, PPs and APs, and lower frequency of SVCs and ‘two predicate constructions’). The list of the 12 linguistic variables is provided here below together with the reference to the chapter where the feature is discussed.

Pre-nominal *itu* ‘D.DIST’ (Chapter 3).

Pre-nominal *satu* ‘one’ (Chapter 3)

Definite marker = *nya* ‘DEF’ (Chapter 3)

Aspect marker *ada* ‘EXIST’ (Chapter 4)

Aspect marker *su* ‘PRF’ (Chapter 4)

Reduplication ‘ITER, INTENS’ (Chapter 4)

‘Two predicate constructions’ in *give*-events (Two-predicate (*give*)) (Chapter 5)

Double Object constructions (DO) (Chapter 5)

Serial verb constructions (SVC) (Chapter 6)

‘Two predicate constructions’ in resultative events (Two-predicate (*res*)) (Chapter 6)

Prepositional Phrases (PP) (Chapter 6)

Adjectival Phrases (AP) (Chapter 6)

A proximity matrix is used to calculate distances between linguistic features (see Table 7.1 on the next page). In the agglomerative hierarchical clustering procedure, each linguistic feature starts in its own cluster and pairs of clusters are merged at each step (Ward, 1963). The first step in the hierarchical clustering process is to look for the pair of features that yield the smallest internal variance when merged into one cluster. In the matrix on Table 7.1, the most similar features DO and = *nya* ‘DEF’, with dissimilarity equal to .763 (shaded gray), so merging them into one cluster incurs a smaller variance increase than merging any other pair. This process is repeated for all the features and the pairs are joined together. The graphical representation of hierarchical clustering is the dendrogram, or clustering tree, where the linguistic features are grouped together in a hierarchical fashion from the closest (most similar) to the furthest apart (the most different). The dendrogram is represented in Figure 7.1 on the next pages.

<i>ada</i> 'EXIST'	DO	SVC	PP	AP	Two- predicate (give)	Two- predicate (res.)	<i>su</i> 'PRF'	Redupli- cation	<i>itu</i> 'D.DIST'	<i>satu</i> 'one'	= <i>nya</i> 'DEF'	
<i>ada</i> 'EXIST'	.000	1012	2.251	.814	.911	2.990	2.074	2.532	2.390	3.640	2.782	.990
DO	1.012	.000	2.512	1.077	1.091	3.176	2.383	2.794	2.433	3.458	2.815	.763
SVC	2.251	2.512	.000	2.529	2.305	3.122	2.813	3.111	3.135	3.705	3.068	2.340
PP	.814	1.077	2.529	.000	.922	2.940	2.067	2.626	2.350	3.743	2.865	1.078
AP	.911	1.091	2.305	.922	.000	2.942	2.295	2.958	2.717	3.946	3.088	.958
Two-predicate (give)	2.990	3.176	3.122	2.940	2.942	.000	2.481	3.377	3.523	4.173	3.379	3.083
Two-predicate (res.)	2.074	2.383	2.813	2.067	2.295	2.481	.000	3.002	2.638	4.177	3.299	2.392
<i>su</i> 'PRF'	2.532	2.794	3.111	2.626	2.958	3.377	3.002	.000	3.237	3.234	3.268	2.746
Reduplication	2.390	2.433	3.135	2.350	2.717	3.523	2.638	3.237	.000	3.555	3.140	2.611
<i>itu</i> 'D.DIST'	3.640	3.458	3.705	3.743	3.946	4.173	4.177	3.234	3.555	.000	2.969	3.428
<i>satu</i> 'one'	2.782	2.815	3.068	2.865	3.088	3.379	3.299	3.268	3.140	2.969	.000	2.875
= <i>nya</i> 'DEF'	.990	.763	2.340	1.078	.958	3.083	2.392	2.746	2.611	3.428	2.875	.000

Table 7.1: Proximity matrix.

The dendrogram with the resulting clusters is shown in Figure 7.1.⁸⁹ Three main clusters can be identified on the basis of their similarity (the clusters are indicated by the numbers).

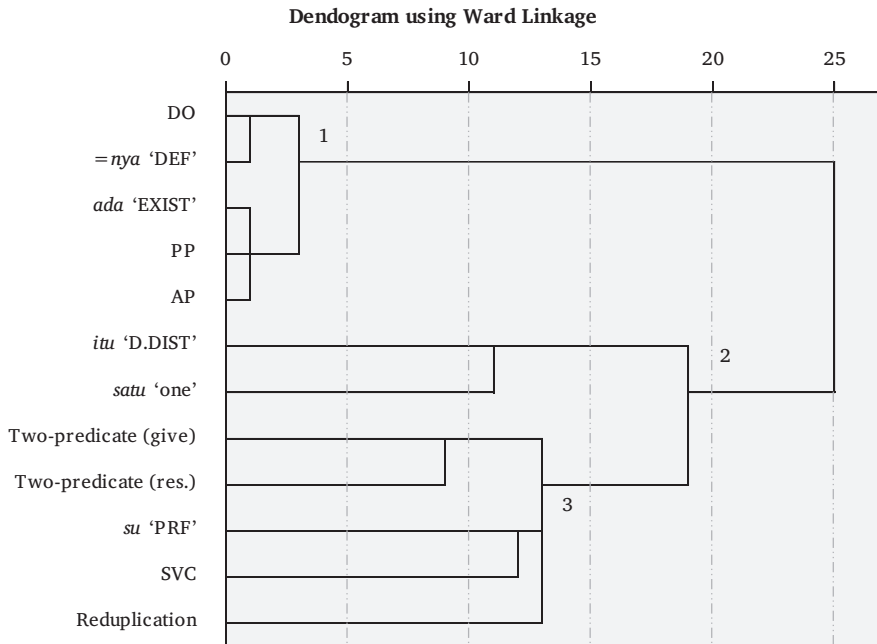


Figure 7.1: Dendrogram of the 12 linguistic variables, distinguishing three clusters

The features in Cluster 1 are fairly close to each other, thus indicating a high degree of similarity among them, whereas the features in Cluster 2 and Cluster 3 display greater internal difference. Interestingly, the features in Cluster 1 are all ‘Dutch-like’ or ‘Dutch-aligned’ features, which are used with a greater frequency by heritage speakers. Recall that, DO, PP, and AP are used increasingly more to the expenses of the Malay-like constructions (i.e., Two-predicate and SVC, respectively), and that both *ada* ‘EXIST’ and *=nya* ‘DEF’ are undergoing an incipient process of grammaticalization. In Cluster 2, we also find Dutch-like features (i.e., pre-nominal *itu* ‘D.DIST’ and prenominal *satu* ‘one’), but in this case the features are less clustered together, as indicated by the relative length of the branches. Features in Cluster 3 are all Malay-like features. The first sub-group includes the ‘two predicate

⁸⁹ A hierarchical cluster analysis (SPSS), with Ward’s method and Euclidean distances was used.

construction' (in both *give-* and resultative events), while the second sub-group includes the aspect markers, *su* 'PRF' and reduplication 'ITER, INTENS', and SVC. The fact that on the basis of heritage speakers behavior, the 'innovative' Dutch-like features cluster together *vis a vis* the 'conservative' Malay-like features supports the idea that the innovative features are brought about by the same factors, namely cross-linguistic influence, type of input and universal principles (see Section 1.3). The next section is dedicated to illustrate how the mechanisms promoting structural convergence at the individual level also have an effect at the community level.

7.2.2 Discussion

One of the most interesting aspects of heritage bilingual communities is that they allow us to observe the beginning of contact-induced change (O'Grady et al., 2011). The findings reported above and in the previous chapters suggest that contact-induced change begins to manifest itself as a change in frequency or preference between structures. Muysken (2013, p. 721) observes that "perhaps the best candidate for L1-oriented syntactic change is the relative increase in use in partial maintenance settings, under influence of another language, of a construction which is already present in a language". A similar point is made by Backus (2004, p. 179), "one way of representing the diachronic process of language change is as shifts in 'entrenchment'". In the usage-based approach, change in frequency is defined as a change in the entrenchment of a particular structure (Croft, 2000; Tomasello, 2000, 2009; Bybee, 2006; Ellis, 2006). Frequency of usage, in fact, reflects the entrenchment level of a unit: the more frequently used, the more entrenched. If frequency leads to entrenchment, entrenchment, in turns, leads to greater fluency in language use; entrenched structures are more easily activated, and more frequently selected in language production. "Language production thus provides further input for the system, which means that usage both comes from and also (re-) shapes the linguistic system itself" (Onar Valk, 2015, p. 50).

Several studies, including the present one, have shown that in bilingual communities, the frequency of a unit in one language can influence the frequency or the level of entrenchment of the corresponding unit in the other language (Johanson, 2002; Backus, 2004; Boumans, 2006; Backus et al., 2011; Onar Valk 2015). This finding is supported by many psycholinguistic studies (Hartsuiker et al., 2004; Schoonbaert et al., 2007; Hartsuiker & Pickering, 2008), which have demonstrated that bilingual speakers tend to reuse recently produced words or

syntactic structures regardless of the language. For instance, Schoonbaert et al. (2007) show that Dutch-English bilinguals are significantly more likely to use an English double-object construction after hearing the construction in Dutch than without such a prime. Psycholinguistic studies thus reveal important evidence about the role of repetition within and across languages and demonstrate that the repetition of particular syntactic structure can occur without co-activation of lexical information (Hartsuiker & Pickering, 2008).

The idea that the frequency of a structure in one language can affect the entrenchment of a somehow equivalent structure in the other language presupposes that bilingual speakers (unconsciously) establish ‘interlingual identifications’ (Weinreich, 1979; Gast & van der Auwera, 2012), ‘equivalence relations’ (Heine & Kuteva, 2005) or ‘analogy’ (Winford 2012, pp. 448-452) between forms (words or structures) and categories across the two languages. As pointed out by Alferink (2015, p. 15), “converging patterns originate from an initial similarity or equivalence, a point where both languages map onto each other to a certain extent”. One may ask how speakers establish these ‘equivalences’ between constructions in two languages, and how entrenchment works cross-linguistically. I believe that Hartsuiker’s (Hartsuiker et al., 2004; Schoonbaert et al., 2007; Hartsuiker & Pickering, 2008) psycholinguistic model of bilingual processing, originally developed to explain priming, offers a satisfactory solution, because it postulates that syntactic procedures are not language-specific, but can be shared. Only if we assume that syntax can be shared, we can theorize that a syntactic procedure of one language can trigger the use of the ‘corresponding’ procedure in the other language.

The model assumes that all information involved in linguistic processing is organized in nodes, interconnected in a network, see Figure 7.2 on the next page. Nodes can represent conceptual (at the top of the picture), lexical (the ovals) or morpho-syntactic information (the rectangles), or simply index the language which is to be activated as a whole (the flags), but essentially, they are not of a different nature as to their capacity to interact. Activation of a lexical node can lead to activation of another lexical node, just as it can co-activate a node containing a morpho-syntactic procedure or some other type of information. Thus, it is also possible for one morpho-syntactic procedure to be activated by connected nodes belonging to two languages. For instance, as shown in Figure 7.2, the concept of ‘break’ is linked to two lemmas, *pata* ‘snap, break’ in Ambon Malay and *breken* ‘to break’ in Dutch. When the Ambon Malay verb *pata* is activated, the syntactic nodes SVC and PP are also activated, because the verb *pata* can occur in a resultative Serial

Verb Construction (*He breaks a stick becomes two*) or it can occur with a resultative PP (*He breaks a stick into two*). The activation of *pata* spreads to its Dutch translation equivalent *breken*, because the two lemmas share the same conceptual node (*break*). Once *breken* is activated, its combinatorial nodes PP and VERB PARTICLE also become active. Language-specific syntactic combinations, such as SVC in Malay and VERB PARTICLE in Dutch, are only connected to the words of the language that allows for them, whereas syntactic combinations that are possible in both languages (PP) are connected (and consequently activated) by words from both languages. Since the combinatorial node PP receives its degree of activation from two languages, it is more likely to be selected (bold rectangle in the figure). The double activation and the consequent entrenchment of this structure in the repertoire account for its increased frequency in bilingual speakers *vis a vis* homeland speakers (see Chapter 6).

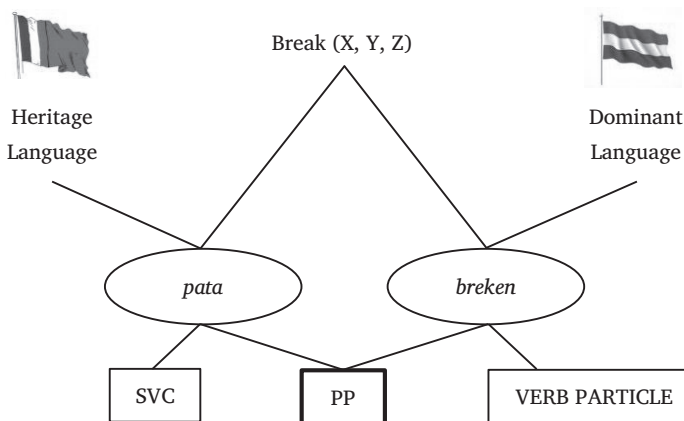


Figure 7.2: Simplified model of co-activation of syntactic procedures in Dutch-Ambon Malay bilinguals.

The continuous activation and selection of the PP node will eventually lead to a shift in entrenchment in favor of PPs and at the expenses of SVCs. This shift in entrenchment is likely to start at the (specific) lemma level (e.g., 'break' + PP), and then to extend to a schematic (less specific) level of representation (VERB + PP). One can speculate that, if Ambon Malay will be spoken in the Netherlands for enough time, the schematic representation (VERB + PP) will be extended to all verbs, and the language will undergo full restructuring and change its typological profile from a serializing language to a non-serializing language, as has happened in the case of Tetun Dili (Hajek, 2006). At this point, it is important to note that, shifts

in entrenchment caused by syntactic priming at the individual level can affect the distributional patterns at the group level as, in immigrant communities, it is mostly bilingual speakers who provide input to other bilingual speakers. I come back to this point in section 7.3, when I discuss the importance of social networks for the propagation of innovations.

Hartsuiker et al.'s processing model also (indirectly) explains why cross-linguistic influence is constrained by the structure of the languages involved. If the two languages do not share a syntactic combination, the activation of a syntactic procedure (e.g., ADJECTIVE-NOUN order) does not spread to the other language. This is illustrated in Figure 7.3: the Dutch ADJECTIVE-NOUN order does not co-activate the corresponding Ambon Malay structure, because the two languages do not share the combinatorial node. The lack of a shared syntactic procedure accounts for the fact that Ambon Malay bilinguals and monolinguals behave alike when it comes to the order of the adjective and the noun and no cross-linguistic influence from Dutch is attested in this domain (see Chapter 3).

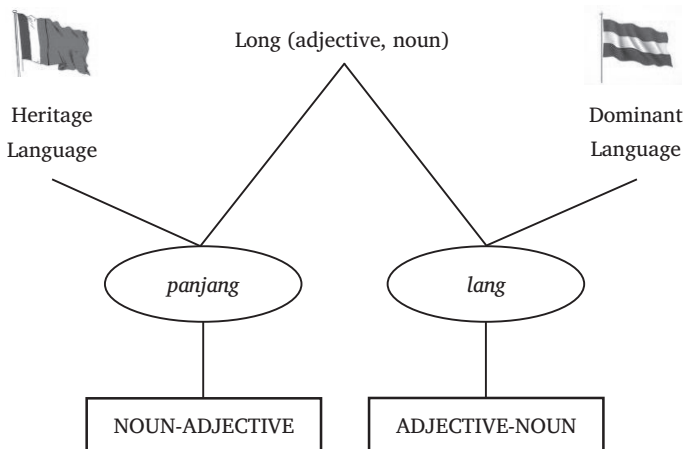


Figure 7.3: Lack of co-activation of syntactic procedures in Dutch-Ambon Malay bilinguals.

We have seen that basic principles of linguistic processing, as illustrated in Hartsuiker et al.'s model, are at the basis of language contact phenomena, such as *change in frequency*. Incipient *contact-induced grammaticalization* is more difficult to fit and to represent by means of Hartsuiker et al.'s model because grammaticalization is a diachronic process that usually extends over a long period of time (see Section 1.3.1.3). The effects of the various changes that bring about

grammaticalization can take long time, sometimes centuries, to surface. Nevertheless, grammaticalization too starts in the mind of bilingual individuals. Contact-induced grammaticalization probably begins as ‘functional convergence’ of features (Sánchez, 2004, 2006). The Functional Convergence Hypothesis, proposed by (Sánchez, 2004), states that convergence “takes place when functional features not present in one of the languages are activated by input and production in the other language” (p. 147). For instance, language A and language B share the feature Tense, but in language A Tense is associated with Aspect, while in language B it is associated with Evidentiality. Bilinguals who are dominant in language B are likely to re-associate evidentiality features to Tense when they are also speaking language A (e.g., they will use past forms to convey the reportative value). Functional convergence is, thus, favored by the frequent activation of the two languages and by partial similarity across the two languages.

The crucial question is where to represent grammatical categories (e.g., Tense and Aspect) in Hartsuiker et al.’s model. Pickering and Branigan (1998) represent Tense and Aspect at the lexical level because these features determine what form the lexical item will take. Other researchers (Flecken, 2010; Bylund & Jarvis, 2011; Jarvis & Pavlenko, 2011) locate Tense and Aspect at the conceptual level because grammatical categories are the instantiation of (language-mediated) concepts (see The Conceptual Transfer Hypothesis in Section 1.4.2). Conceptual representations are shared by the two languages and are, thus, likely to receive cross-activation. An example is given in Figure 7.4 on the next page. When the concept of ‘break’ is activated in Ambon Malay, the grammatical categories that are relevant for the expression of this concept, like Aspect (rounded rectangle), also become active. Since the conceptual node ‘break’ is shared with Dutch, the activation will spread to the Dutch grammatical categories of Tense, Aspect, and Finiteness (rounded rectangle). The activation of Aspect in Ambon Malay also activates Tense and Finiteness in Dutch. This continuous cross-activation may ultimately lead to the re-association of the Tense (and possibly also Finiteness) category to Ambon Malay morphological forms. The ‘new’ feature is re-mapped onto a form or a structure that already performs a similar function in the heritage language. According to Matras (2009, p. 27), polysemy is the key factor in this process. So, we can hypothesize that heritage speakers of Ambon Malay select the aspect marker *ada* to express the present tense (or Finiteness) because this morpheme already has a tense component in its meaning (see Section 4.3.1).

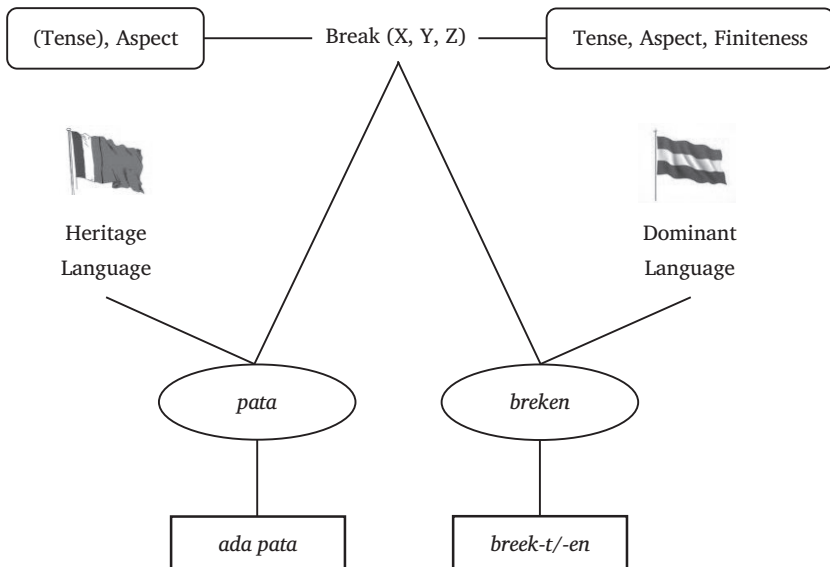


Figure 7.4: Tentative representation of grammatical categories associated with the conceptual level.

Before turning to the next sections, two important remarks should be made. First, as convergence can be seen as both a process and an outcome of language contact, the shared syntax view advocated by Hartsuiker et al. (2004) can also be seen as both a cause and an outcome. A study conducted by Bernolet, Hartsuiker and Pickering (2013) found a positive correlation between priming and L2 proficiency, with more proficient participants showing stronger priming effects and least proficient participants hardly showing any priming. On the basis of this finding, Hartsuiker (2013, p. 738) concludes that “L2 learners start with separate combinatorial nodes for their second language and only “collapse” them with the L1 representation after sufficient L2 experience. Thus, [shared syntax] is an end point in an L2 learning trajectory”. Now, one may ask why speakers ‘collapse’ representations from their L1 and L2, or more generally, what are the forces behind convergence. The most plausible answer to these questions is: ‘processing economy’. By increasing the structural overlap between two languages, converge lightens the cognitive load of having to remember and use two different linguistic systems (Silva-Corvalán, 1994). In language processing terms, convergence helps the system run more smoothly and access information with higher speed (cf. Alferink, 2015, pp. 116-117). It is plausible to assume that, if speakers collapse the syntax of their two

languages, they will try to do so also with patterns of conceptual categorization and that having one set of conceptual categories increases the economy of the system.

The second remark concerns the role of incomplete acquisition in the account of contact-induced change given so far. In Chapter 2 (Section 2.1.1), I explicitly state that the effect of incomplete acquisition or attrition on the grammar of heritage Ambon Malay cannot be tested due to the lack of L1 acquisition data. Without this type of data, it is not possible to establish whether a grammatical structure was not completely acquired by the bilingual heritage speakers (incomplete acquisition), or whether it was acquired but then lost or reduced due to lack of use (attrition). Although incomplete acquisition and attrition on the one hand, and cross-linguistic influence on the other hand, represent two different sources for heritage language divergence, they can also interact (see Section 1.3). The effects of cross-linguistic influence observed above (i.e., change in frequency of constructions and contact-induced grammaticalization of heritage morphemes) may have been exacerbated by the fact that heritage speakers have never fully acquired (or have lost) the discourse-pragmatic rules governing those features. Language acquisition is a gradient process, and even though heritage speakers have certainly acquired the morpho-syntactic rules of a particular structure (e. g. the morphological process of reduplication, attaching the enclitic =*nya* to a noun), they may have not fully acquired the frequency, contexts of usage and pragmatic functions of the structure; and these ‘incompletely acquired’ aspects are exactly what cross-linguistic influence targets. So for instance, heritage speakers have never acquired or lost the discourse-pragmatic constraints governing the use of the definite marker =*nya*, and therefore they apply the Dutch rule that predicts definite marking on all already mentioned nouns. It is possible that heritage speakers’ knowledge of rules governing definite marking has eroded under the effect of the dominant language, Dutch, where definiteness is obligatory.

To conclude this section, change in frequency and (embryonic) contact-induced grammaticalization in heritage Ambon Malay lead to divergence from the homeland variety and to convergence toward Dutch. These processes involve a shift in the entrenchment of syntactic procedures and grammatical categories, respectively. Hartsuiker et al.’s model helps us to explain how the repeated co-activation of the same, shared, syntactic node eventually leads to the entrenchment of the syntactic procedure in the repertoire of the bilinguals, who will use it independently of the language. If structural factors account for the ‘qualitative’ aspect of convergence, it is social factors that determine the ‘quantitative’ aspect and to which we now turn.

7.3 Social factors and language change

A number of social factors have been found to correlate with linguistic variation and linguistic change, in both monolingual communities (Labov, 2001; Milroy & Milroy, 1985, 1992; Trudgill 1972, 2011, Chambers, Trudgill, & Schilling-Estes, 2002, among others), and bilingual communities (Thomason & Kaufman 1988; Wei, 1994; Hulsén, 2000; Montrul, 2008, among others).

In bilingual (heritage) communities, the emergence of linguistic change is usually related to factors such as *age of onset of bilingualism* (sequential vs. simultaneous) and *parental language input* (Montrul, 2008; Unsworth, Argyri, Cornips, Hulk, Sorace, & Tsimpli, 2014). Montrul (2008) cites a number of studies showing that simultaneous bilinguals make more errors than sequential bilinguals, and (p. 115) concludes that “incomplete acquisition appears to be more dramatic in simultaneous than in sequential bilingualism, due to the reduced amount of input received in one language while the language was not yet fully developed”. This view is challenged by Kupisch (2013), who reports studies on simultaneous German-French bilinguals showing that many properties are acquired successfully also by this group. Kupisch (2013, p. 210) also raises the important issue concerning the exact determination of age of onset in both languages:

For successive [sequential] bilinguals, it is comparatively more difficult to say when exactly they have been exposed to the majority language for the first time. On the one hand, they are continuously surrounded by majority language speakers and they normally have access to the media in the majority language. On the other hand, it is rather unclear when they start to be interested in stimuli that are not directed to them personally.

The distinction between sequential vs. simultaneous bilingualism is easily determined in heritage language children, for which parents can keep language diaries indicating the amount of time spent speaking one language, or spent at daycare, school, and out of school care (depending on the child's age) and the number of hours per week spent on other activities, including reading, clubs, and sports; watching TV, etc. (see Unsworth et al., 2014). In the case of adult heritage speakers, we lack this information, and it is therefore hard to establish with certainty the exact onset of the dominant language.

Parental language is also an important predictor of linguistic innovations. For instance, Raschka, Wei and Lee (2002) report that Cantonese heritage speakers whose parents consistently and exclusively use Cantonese have good levels of

language ability, whereas heritage speakers whose parents use Cantonese and English interchangeably have a lower level of language proficiency. Similarly, Irizarri van Suchtelen (2014) shows that, although all Spanish heritage speakers in the Netherlands diverge considerably from their monolingual peers, heritage speakers with two Spanish speaking parents perform more native-like than heritage speakers who grew up with only one Spanish speaking parent and did not productively use Spanish in childhood.

Generation or age is yet another factor that correlates with language maintenance and language proficiency in heritage language communities (Silva-Corvalán, 1994; Wei, 1994; Hulsen, 2000). The heritage language has been found to decline with each generation, following the pattern $G1 > G2 > G3$. Generally speaking, first generation speakers are dominant in the heritage language, second generation speakers have relatively strong skills in both the heritage and the majority languages, whereas third generation speakers are dominant in the majority language; beyond the third generation, few heritage speakers retain a functional command of their language (Benmamoun et al., 2010, p. 79; Carreira & Kagan, 2011, p. 42). An example is provided by Hulsen (2000), who looks at lexical access in production and comprehension in Dutch heritage speakers in Australia, using a picture-naming and a picture-matching task. She reports that second generation Dutch were able to perform both tasks, but they differed significantly in both speed and accuracy of lexical retrieval from both first generation speakers and a control group of Dutch speakers in the Netherlands. In contrast, third generation speakers were only able to perform the picture matching task (comprehension) but were unable to perform the picture-naming task (production). The language of first generation speakers is usually the most similar to the homeland variety, whereas the language of second- and especially third-generation speakers feature many innovations and is characterized by incomplete acquisition and loss of linguistic features (Carreira & Kagan, 2011, p. 79).

Social network structure has also proved to be a factor influencing the degree of language preservation and language ability in immigrant communities. Most of the findings demonstrate that a close-knit territorially bounded network with strong ties acts as a conservative force helping to resist outside innovations and promoting language maintenance (Wei, 1994; Wei, Milroy, & Ching, 2000; Chau, 2011). Territorially based networks, in fact, enhance the frequency and the intensity of contact among the community members, which, in turn, favor language preservation. For instance, Ambon Malay speakers living in Moluccan municipalities

in the Netherlands retain a higher rate of language maintenance than their peers living outside the wards (Veenman, 1994). Similar findings have been reported for different Chinese immigrant communities. A comparison between two groups of families in Tyneside, England, - one with strong ties and affiliated to the True Jesus church, the other with weak ties and lacking a centralized institution- shows that the families of the first group have an higher level of language maintenance and language proficiency compared to the families in the second group (Wei et al., 2000). Another example is that provided by Chau (2011), who compares Cantonese heritage speakers in Amsterdam (an urban area), where a large community of Cantonese is present, to Cantonese heritage speakers in Venlo (a rural area), who are rather isolated. Heritage speakers in Amsterdam use Cantonese to a larger extent (they watch Cantonese soaps and listen to Canto-pop) than the Cantonese heritage speakers in Venlo, because the former can share their experiences with their Cantonese speaking peers who live in the neighborhood, whereas the latter do not have many Chinese peers and therefore prefer Dutch soaps and non-Chinese music. According to Chau (2011), the low exposure to Cantonese of Venlo heritage speakers accounts for the many deviations found in their speech (i.e., wrong use of nominal classifiers, inverted double object construction). Finally, other social factors can have an effect on language maintenance and proficiency, such as ethnic identity orientation, attitude, learning goals and motivation (Kondo-Brown, 2003; Chinen & Tucker, 2005; Carreira & Kagan, 2011).

We have seen that many studies report a positive interaction between language maintenance and proficiency and social factors, such as generation and social network, which heavily influence the amount of (heritage) language use. The interaction between language use and language preservation, however, is more complex than it first appears, because frequent heritage language use can entail a higher rate of maintenance but it can also entail a higher rate of innovations (or accelerated change). Two studies report a non-linear interaction between language use and native-like language proficiency. The first study investigates first language attrition among German immigrants in Canada and the Netherlands (Schmid, 2007, 2011). The results of the study show that the ‘attriters’ differ from the control group in Germany with respect to lexical access, lexical diversity and fluency measures, but this difference does not depend on the amount of L1 use in daily life. Interestingly, length of residence in the foreign country seems to have an effect on individuals with the lowest and the highest rate of L1 use. Schmid (2011, p. 171) concludes that:

These findings suggest that change and deterioration of the L1 which may be witnessed among migrant populations may be determined by two opposite poles: speakers who do not use their L1 at all may experience some degree of ‘atrophy’, while those who live in a bilingual migrant community where L1 and L2 are used frequently alongside each other and mixed to some degree may find themselves sharing in a language with accelerated signs of contact induced change.

The second study focuses on null subjects and Voice Onset Time (VOT) in three generations of heritage language speakers (Cantonese, Italian, Russian, and Ukrainian) in Toronto (Nagy, 2015). The results of the study show that the three generations of Italians, Cantonese and Russians exhibit the same rate of null subjects and there is no evidence for higher rate of null subjects among speakers with a higher level of exposure. Similarly, VOT patterns are the same among the three generations of Italians (although a difference is present among the Russians and the Ukrainians), but in any case the difference does not seem to depend on social variables. Nagy (2015, p. 325), thus, concludes that “correlation is not found between linguistic performance and measures of reported language exposure, use or preference”.

To conclude, although there is an increasing body of evidence showing that heritage speakers with a history of low exposure to or use of the heritage language exhibit less homeland-like patterns, this is not always the case, as both infrequent and frequent exposure can accelerate on-going language change. So, if on the one hand close-knit network structures can help to maintain the ‘integrity’ of the heritage language, on the other hand, once an innovation has been introduced, they can boost the innovation process and promote language change within the community.

7.3.1 Effect of social factors

In some of the previous chapters (Chapter 5 and Chapter 6), I reported qualitative observations on the relation between the linguistic innovations observed in heritage Ambon Malay and the language history of the speakers (i.e., age of acquisition of Dutch, the main language spoken with parents). In this section, I investigate this relation in a more systematic and quantitative way by testing possible correlations between the 12 linguistic features listed in Section 7.2.1 and a number of social variables.

On the basis of previous research (see Section 7.3), the following, variables were chosen: *age of onset of bilingualism* (sequential, simultaneous), *main language spoken in the family* (Ambon Malay, Dutch, Ambon Malay & Dutch), *generation* (second generation, generation 2.5⁹⁰, third generation), *place where the speaker grew up* (Moluccan camp, Moluccan ward, outside Moluccan ward), and *place where the speaker lives* (Moluccan ward, outside Moluccan ward). Some of these variables, however, are correlated: *age of onset of bilingualism* strongly correlates with *main language spoken in the family* ($r(32) = .866, p < .001$), and also with *place where the speaker grew up* ($r(32) = .673, p < .001$). In fact, sequential bilinguals are mostly individuals who grew up in households in Moluccan camps, where Malay was the main language spoken. Due to this correlation, the two variables (*main language spoken in the family* and *place where the speaker grew up*) were not included in the model. A General linear model (in SPSS) was used to test the possible effect of the social variables (predictors = *age of onset of bilingualism*, *generation*, *place where the speaker lives*) on the 12 linguistic variables (responses). The results are summarized in Table 7.2, on the next page.

⁹⁰ Speakers belonging to this generation have one parent from the first generation and the other parent from the second generation (see Section 2.1.2.1).

	AoA		LIVE		GEN		AoA*LIVE		AoA*GEN		LIVE*GEN		R
	p	PES	p	PES	p	PES	p	PES	p	PES	p	PES	
DO	.031	.179	.000	.546	.997	.000	.000	.455	.897	.001	.057	.142	.515
PP	.909	.001	.489	.020	.896	.009	.180	.180	.540	.016	.736	.005	-.058
AP	.125	.095	.655	.008	.992	.001	.496	.020	.697	.006	.851	.002	-.055
=nya 'DEF'	.000	.505	.000	.593	.628	.038	.000	.619	.277	.049	.005	.280	.657
ada 'EXIST'	.538	.016	.253	.054	.938	.005	.746	.004	.074	.127	.819	.002	.002
itu 'D.DIST'	.737	.005	.001	.376	.067	.202	.080	.123	.924	.000	.026	.191	.294
satu 'one'	.039	.165	.601	.012	.053	.217	.480	.021	.699	.006	.579	.013	.086
SVC	.989	.000	.044	.158	.258	.107	.340	.038	.240	.057	.401	.030	.224
Two-predicate (give)	.733	.005	.260	.053	.469	.061	.047	.154	.243	.056	.117	.056	.120
Two-predicate (Res.)	.776	.003	.217	.063	.857	.013	.812	.002	.449	.024	.405	.029	-.118
su 'PRF'	.938	.000	.922	.000	.575	.045	.819	.002	.705	.006	.070	.130	-.053
Reduplication	.511	.018	.852	.001	.750	.024	.409	.029	.705	.006	.610	.011	-.148

Table 7.2: AoA = age of acquisition of Dutch, LIVE = where the speaker lives, GEN = generation, p = p value, PES = significant effect sizes, R = adjusted R squared values, (give) = give-events, (res.) = resultative events. Shaded cells indicate significant results.

Given the small sample size (only 32 speakers), we can only expect to find statistical significance for the strongest effects, and where we do not find a significant effect it may either be that there is no such effect or that it is too weak to surface in the current sample. We find significant effects for six variables, namely in the use of DO, =*nya* 'DEF', pre-nominal *itu* 'D.DIST', prenominal *satu* 'one', and SVC, and 'two predicate constructions' in *give*-events. The *place where the speaker lives* seems to be the best predictor of linguistic innovations: speakers living outside a Moluccan ward have an higher rate of Dutch-like features (i.e., DO, =*nya* 'DEF' and pre-nominal *itu* 'D.DIST') than speakers living in a Moluccan ward, who in turn, have an higher rate of SVC (a Malay-like feature). In the case of another Malay-like feature, the 'two predicate construction' (in *give*-events), there is an interaction with *age of onset of bilingualism*, such that sequential bilinguals living in a Moluccan ward display the highest rate. *Age of onset of bilingualism* is somehow more difficult to interpret: with respect to pre-nominal *satu* 'one', it gives the expected result, namely that simultaneous bilinguals use the pre-nominal position (Dutch-like feature) more often than sequential; however, with respect to =*nya* 'DEF', the result is the opposite, sequential bilinguals use =*nya* 'DEF' more often than simultaneous bilinguals. A similar observation holds for DO. The interaction between *age of onset of bilingualism* and *place where the speaker lives* shows that sequential bilinguals living outside a ward have the highest rate of DO and =*nya* 'DEF'. *Generation* does not play a role for any of the features. The interaction with *place where the speaker lives* shows that speakers of the second generation and of the generation 2.5 living outside a Moluccan ward are the most innovative (higher rate of =*nya* 'DEF' and pre-nominal *itu* 'D.DIST').

Notably, *place where the speaker lives* is the best predictor of innovations, speakers living outside Moluccan wards are the most innovative, although in some cases, it is the sequential bilinguals living outside a Moluccan ward who have the highest rate of Dutch-like features. If we test these findings using the clusters of features displayed in Figure 7.1 (Section 7.2.1), we obtain similar results, see Table 7.3 on the next page. Living outside a Moluccan ward is by far the most important factor for the linguistic variables in Cluster 1 and Cluster 2 (The Dutch-like features), while living in a Moluccan ward has a positive effect on the linguistic variables in Cluster 3 (the Malay-like features). *Age of onset of bilingualism* has an effect on the variables in Cluster 1, such that sequential bilinguals living outside Moluccan wards display the highest rate.

	AoA		LIVE		GEN		AoA*LIVE		AoA*GEN		LIVE*GEN		R
	p	PES	p	PES	p	PES	p	PES	p	PES	p	PES	
CLUSTER 1	.002	.332	.000	.516	.860	.013	.000	.583	.924	.000	.036	.171	.583
CLUSTER 2	.110	.103	.012	.234	.028	.258	.516	.018	.752	.004	.077	.125	.249
CLUSTER 3	.805	.003	.036	.710	.797	.019	.062	.138	.386	.031	.782	.003	.098

Table 7.3: AoA = age of acquisition of Dutch, LIVE = where the speaker lives, GEN = generation, p = p value, PES = significant effect sizes, R = adjusted R squared values, (give) = give-events, (res.) = resultative events. Shaded cells indicate significant results.

CLUSTER 1 = DUTCH-LIKE FEATURES (DO, = *nya* 'DEF', *ada* 'EXIST', PP, AP)

CLUSTER 2 = DUTCH-LIKE FEATURES (pre-nominal *itu* 'D.DIST', pre-nominal *satu* 'one')

CLUSTER 3 = MALAY-LIKE FEATURES (*su* 'PRF', reduplication, SVC, 'two predicate construction')

7.3.2 Discussion

The previous section has shown that the effects of the social variables – *age of onset of bilingualism*, *place where the speaker lives* and *generation* - on the 12 linguistic variables. *Age of onset of bilingualism* does not have any influence on most of the linguistic variables, and when it does, it interacts with *place where the speaker lives*. The first issue that this section addresses is, then, why does *age of onset of bilingualism* not have a straightforward effect in the shaping of heritage Ambon Malay grammar.

We have seen that, in some cases, sequential bilinguals show a higher rate of innovation than simultaneous bilinguals, and that second generation speakers seem more innovative than third generation ones (only for the variables in cluster 2). This apparently counterintuitive finding is actually in line with the observation of Schmid (2007, 2011) reported in the previous section, namely that language change may be found among speakers who frequently use their L1 and L2 alongside each other. Now, it is plausible to assume that, having received considerable Ambon Malay input in childhood, sequential bilinguals and second generation speakers are more confident about their linguistic skills in the heritage language and, thus use the language more frequently than the other speakers. The frequent use of the dominant and the heritage languages alongside each other creates the conditions for syntactic priming and the subsequent shift in entrenchment that derives from it. Thus, if high exposure in childhood translates into relatively frequent use in adulthood, then the language of sequential bilinguals can indeed show ‘accelerated signs of contact induced change’.

The second point related to *age of onset of bilingualism* is the nature of the linguistic variables that are correlated with it. The innovative features described in this dissertation are rather different from the features described by Montrul (2008) and Unsworth et al. (2014). The phenomena investigated by these authors are errors in tense-aspect-mood inflection morphology or in the assignment of grammatical gender. For this type of phenomena, *age of onset of bilingualism* has proved to be an important factor predicting the incomplete acquisition of morphological marking (but cf. Kupisch, 2013). The phenomena described in this dissertation are of a somehow different nature: we do not deal with errors and inflectional morphology, but rather with changes in frequency. Acquiring the form-meaning mapping of a particular construction, together with its frequency, pragmatics, and contexts of usage may be a process that continues well beyond the age of five (the threshold for

sequential vs. simultaneous bilingualism), and that needs continuous language input through the lifetime. After adolescence, sequential and simultaneous heritage speakers become virtually identical, as Dutch becomes the main functional language for them. Hence, *age of onset of bilingualism* may not make a difference because the innovative features of Ambon Malay are not acquired once and for all in childhood but are part of a gradient process of language acquisition that is continuous through the lifetime and is sustained by other intervening factors, such as the amount and the type of Ambon Malay speaking contacts an individual has. The observation that innovations seem to depend more on continuous lifetime exposure rather than on *age of onset of bilingualism* supports the importance of social network, to which we now turn.

Social network, represented here by the social variable *place where the speaker lives*, is by far the most important factor in the heritage Ambon Malay community. Heritage speakers living outside a Moluccan ward show the highest rate of Dutch-like features, whereas speakers living in a Moluccan ward retain Malay-like features more firmly. This finding coincides with previous studies on the immigrant Chinese community (Wei, 1994; Chau, 2011) in the U.K. and other studies on the Ambon Malay community (Huwaë, 1992; Veenman, 1994; Tahitu & Lasomer, 2001), which found a correlation between social network and language proficiency and language maintenance. Now, one may ask why does living in or outside a Moluccan ward have an effect on the heritage language. The first, and the most intuitive, possibility is that speakers living outside a Moluccan ward have less chances to speak Ambon Malay compared to their peers living in a ward, so their heritage language shows signs of ‘atrophy’. Since they are fully immersed in a Dutch-speaking environment, there are higher chances that, when they speak Ambon Malay, they will rely on Dutch-like structures.

However, we have seen that, among the speakers living in a city, the most innovative are sequential bilinguals, who supposedly use their heritage language frequently. This, somehow counterintuitive, result is compatible with Schmid’s idea (2011) that in order to be an innovator, a speaker needs to use his/her (heritage) language. As mentioned above, sequential bilinguals probably feel more comfortable speaking the heritage language than (some) simultaneous bilinguals, and this leads them to use Ambon Malay and Dutch alongside each other frequently. The continuous and frequent switch from one language to the other is likely to increase the chances of syntactic priming, eventually leading to a shift in the entrenchment of Malay-like structures in favor of Dutch-like structures. Now, the question remains,

why are (sequential) heritage speakers in the city (outside a Moluccan ward) the most innovative?

According to the ‘weak ties’ model of language change, people who are not part of a close-knit, territorially based, network tend to be the innovators in linguistic changes (Milroy & Milroy, 1985, 1992). Thanks to their ‘mobility’ and the many acquaintances they have (weak ties), these individuals are more likely to be exposed to new (linguistic) information than the members of a close-knit network. Milroy & Milroy (1985, p. 366) point out that, new (linguistic) ideas pass from one group to another via weak-network links, while “information relayed through strong ties tends not to be innovatory”. In other words, weak ties function as bridges through which new information and/or linguistic innovation spreads from one group of individuals to another group (Milroy & Milroy, 1985, 1992). The process of propagation of linguistic innovations has been described by Enfield (2003, p. 366) in the following way:

Some individual or individuals begin to habitually perform a new linguistic act, exposing those in their personal network to the idea, with the result that those who are exposed then replicate this performance (given sufficient motivation to do so), and in turn expose more people in their own social networks (as well as those who began the process in the first place, revalidating and encouraging the usage, and leading it to take further hold).

Thus, it is likely that Ambon Malay heritage speakers living in a city (outside a Moluccan ward) participate in several networks and interact with (at least some) individuals in Ambon Malay; when they move from one network to another, they may (more or less voluntarily) diffuse the innovations that they have picked up in these former interactions.

7.4 Conclusions

The previous chapters of this dissertation have illustrated and discussed a number of on-going changes (innovations) that are attested among heritage speakers of the Moluccan community in the Netherlands. Heritage Ambon Malay has been found to diverge from the homeland variety spoken on Ambon, Indonesia, and to converge toward Dutch. *Convergence* is mainly instantiated by changes in frequency or preference and contact-induced grammaticalization. The cognitive motivation

behind convergence is to increase the similarity between the two linguistic systems in order to achieve greater 'processing economy'.

Both structural and social factors constraint the shape and the amount of language change that takes place in heritage Ambon Malay. The typological profile of Ambon Malay and Dutch determines which structures will be affected by cross-linguistic influence and which will not be affected. The amount of exposure to Ambon Malay received in childhood and the social network of the speaker determine the direction and the incidence of the change. Sequential bilingualism does not always entail less cross-linguistic influence. In some cases, sequential bilinguals are subject to more cross-linguistic influence (higher rate of Dutch-like feature) because they frequently use the two languages alongside each other. Finally, the *place where the speaker lives* strongly predicts the amount of innovation in the heritage language, where individuals living outside Moluccan wards are the groundbreakers in most of the 'on-going' changes.

CHAPTER 8

Conclusions and future research

The aim of this dissertation was to compare heritage Ambon Malay, as spoken in the Netherlands, to homeland Ambon Malay, as spoken in the Moluccas, in order to detect signs of divergence between these two varieties, and signs of convergence between the heritage variety and Dutch. Since the 1950s, the two varieties have developed in two ways: homeland Ambon Malay in the Moluccas has been influenced by Standard Indonesian, while heritage Ambon Malay in the Netherlands has been influenced by Tangsi Malay, and subsequently by Dutch, the dominant language of bilingual heritage speakers. The chapters of this dissertation have focused on some specific areas of heritage Ambon Malay grammar and have provided quantitative analysis of the patterns. The data and the analysis carried out in each chapter allow us to answer the questions posed in the introduction:

- Does heritage Ambon Malay diverge from its homeland variety?
- Is heritage Ambon Malay changing under the influence of Dutch? How does this ‘on-going’ change manifest itself?
- What are the factors driving this ‘on-going’ change?

The answer to the first question (Does heritage Ambon Malay diverge from its homeland variety?) is undoubtedly yes. Heritage Ambon Malay has been found to diverge from homeland Ambon Malay in all the four grammatical areas investigated in this dissertation, namely nominal modification, tense-aspect, *give*-constructions and resultative-constructions. The differences between the two varieties are both qualitative and, most importantly, quantitative, thus indicating that the changes are pervasive in a single speaker and within the speech community. The findings of Chapter 3, Chapter 4, Chapter 5, and Chapter 6, together with the findings of other studies (Tahitu, 1988, Huwaë; 1992; Aalberse & Moro, 2014) demonstrate that heritage Ambon Malay has changed and is still changing rapidly due to unbalanced bilingualism and due to contact with Dutch. This leads us to the second question.

The answer to the second question (Is heritage Ambon Malay changing under the influence of Dutch? How does this ‘on-going’ change manifest itself?) is yes. Dutch is the dominant language of heritage Ambon Malay speakers and, as such, is perhaps

the most important impetus of change. Bilingual speakers, in fact, have been found to collapse the syntax of their two languages (as much as possible) in order to lighten the cognitive load of having to remember and use two language systems. As a result, the grammar of heritage Ambon Malay is becoming increasingly similar to the grammar of Dutch. Convergence toward Dutch also accounts for divergence from homeland Ambon Malay: the more heritage Ambon Malay converges toward Dutch, the more it diverges from the homeland variety. But, how does the influence from Dutch manifest itself? It mainly manifests itself in two ways: change in frequency and contact-induced grammaticalization. If Ambon Malay has two (or more) equally possible options, heritage speakers will prefer the option also present in the dominant language, Dutch, therefore increasing its frequency; or they will turn a pragmatically marked option into a pragmatically unmarked one to match the frequency of that option in Dutch. If a grammatical contrast is obligatorily marked in Dutch, heritage speakers will tend to overtly express it in Ambon Malay by recruiting existing structures and grammaticalizing them in the heritage language.

The answer to the third question (What are the factors driving this ‘on-going’ change?) is: cross-linguistic influence from Dutch, the qualitatively and quantitatively different input that heritage speakers were exposed to, and universal principles of language development in the context of language disuse, such as simplification, conceptual naturalness, preference for certain iconic constructions. As mentioned in the previous paragraph, cross-linguistic influence is perhaps the most relevant factor driving patterns of divergence and convergence in heritage Ambon Malay. Being the dominant language of bilingual heritage speakers, Dutch is slowly but surely influencing all aspects of the heritage grammar. Another factor that accounts for divergence is the different type of input heritage speakers were exposed to, namely the language spoken by first generation speakers in the Netherlands. This language variety differs from the homeland variety in two ways: it was heavily influenced by Tangsi Malay and is probably subject to attrition. The Tangsi Malay elements present in the language of first generation speakers have been passed to the language of heritage speakers of the second and third generation, and this has enhanced the divergence of the heritage variety from the homeland variety. Furthermore, heritage speakers received input in the heritage language from first generation speakers, whose language may have been already attrited, and this contributes to further divergence from the homeland language. Lastly, universal principles of language development in the context of language disuse also play a role in shaping heritage Ambon Malay grammar. Simplification, conceptual naturalness,

preference for certain iconic constructions are yet other causes of divergence between the heritage grammar and the homeland grammar. These factors are likely to interact with each other (so that, for instance, contact with Dutch can accelerate a change driven by conceptual naturalness) and, in many cases, the multicausal explanation is perhaps to be preferred.

The above mentioned factors, namely cross-linguistic influence, type of input, and universal principles, account for the changes attested in heritage Ambon Malay. Social factors, such as age on onset of bilingualism and the place where the speaker lives determine the extent of the change within and across different individuals. The place where the speaker lives is the most important factor in the heritage Ambon Malay community, followed by age of onset of bilingualism. Heritage speakers living outside a Moluccan ward show the highest rate of innovative Dutch-like features, whereas speakers living in a Moluccan ward retain conservative Malay-like features more firmly. The role of age of onset of bilingualism is more difficult to interpret. In some cases simultaneous bilingualism correlates with a high rate of innovative Dutch-like features, but in other cases it does not. Moreover, sequential bilinguals living outside a Moluccan ward are found to be the most innovative. Sequential bilinguals probably feel more comfortable speaking the heritage language than (some) simultaneous bilinguals, and this leads them to frequently use Ambon Malay and Dutch alongside each other, thus creating a chance for cross-linguistic influence to take place.

This dissertation has investigated patterns of divergence and convergence in four grammatical areas, namely nominal modification, tense-aspect, *give*-constructions and resultative-constructions. Many other areas of heritage Ambon Malay grammar need to be explored or investigated further in future research. Among these, there is modality, a domain that seems to be subject to interesting contact phenomena. A preliminary study that I conducted shows that the Ambon Malay necessity modal *musti* 'must' is extending its semantic range to resemble the Dutch equivalent *moeten* 'must' (see Moro, 2015). The Dutch modal *moeten* 'must' includes four modal meanings, namely deontic, participant external, participant-internal and epistemic. The Ambon Malay modal *musti* 'must' only includes the deontic and participant external meaning. So, while in Dutch it is perfectly fine to say that someone *moet niezen* 'must/has to sneeze' (a participant-internal necessity), the expression *musti bersin* 'must/has to sneeze' in Ambon Malay is infelicitous. The heritage Ambon Malay modal *musti* 'must' has undergone semantic extension under the influence of Dutch and has acquired the participant-internal meaning previously absent in the

language, so that the expression *musti bersin* ‘must/has to sneeze’ is commonly used among heritage speakers. It is likely that contact with Dutch has accelerated an internal change driven by conceptual naturalness. The extension from participant-external to participant-internal modality is very common cross-linguistically and, thus, it can be expected to take place independently of contact. Beside modality, the reader can find suggestions for future research in the chapters of this dissertation and in the work of Tahitu (1988, 1989) Huwaë (1992), Lekawael (2011) and Aalberse & Moro (2014).

Another important area for future research is the language of first generation speakers, which cannot be overlooked, if we wish to correctly disentangle the sources of divergence in heritage speakers. Further data from this group of speakers will help us to establish which innovations are proper of the language of heritage speakers and which are passed on to them from the attrited language of their parents. Data from Ambon Malay L1 children are also necessary, if we wish to establish with certainty whether heritage speakers had the chance to fully acquire a certain linguistic feature or not, and how age of acquisition interacts with transfer. Finally, additional data from the heritage community, especially spontaneous speech or conversation data, can contribute to our understanding of language change by adding another piece to the puzzle and hopefully provide converging evidence for the claims I proposed in this dissertation.

References

- Aalberse, S. P., & Moro, F. R. (2014). Stability in Chinese and Malay heritage languages as a source of divergence. In K. Braunmüller, S. Höder & K. Kühn (Eds.), *Stability and divergence in language contact: Factors and mechanisms* (pp. 141–162). Amsterdam and Philadelphia: John Benjamins.
- Aalberse, S. P., & Muysken, P. C. (2013). *Perspective on heritage languages*. Position paper for the Structural Changes in Heritage Languages Workshop, Noordwijkerhout, The Netherlands. Retrieved July 18, 2014 from http://www.ru.nl/linc/news_and_events/events/workshop-'structural/.
- Abdurachman, P. R. (1972). *Some Portuguese loanwords in the vocabulary of speakers of Ambonese Malay in Christian villages of Central Moluccas*. (Seri Lembaran Khusus 1). Jakarta: Lembaga Research Kebudayaan Nasional.
- Adelaar, K. A., & Prentice, D. J. (1996). Malay: its history, role and spread. In S. A. Wurm, P. Mühlhäusler & D. T. Tyron (Eds.), *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas* (pp. 673-693). Berlin: Mouton de Gruyter.
- Adelaar, K. A. (2005). Structural diversity in the Malayic subgroup. In K. A. Adelaar & N. Himmelmann (Eds.), *The Austronesian languages of Asia and Madagascar* (pp. 202-226). London and New York: Routledge.
- Aikhenvald, A. Y., & Dixon, R. M. (2006). Serial verb constructions in typological perspective. In A. Y. Aikhenvald & R. M. Dixon (Eds.), *Serial verb constructions: A cross-linguistic typology* (pp. 1-68). Oxford: Oxford University Press.
- Albirini, A., Benmamoun, E., & Saadah, E. (2011). Grammatical features of Egyptian and Palestinian Arabic heritage speakers' oral production. *Studies in Second Language Acquisition*, 33(2), 273-303
- Alferink, I. (2015). *Dimensions of convergence in bilingual speech and gesture*. (LOT Dissertation Series 391). Utrecht: LOT.

- Ameka, F., de Witte, C., & Wilkins, D. (1999). Picture series for positional verbs: Eliciting the verbal component in locative descriptions. In D. Wilkins (Ed.), *Manual for the 1999 Field Season* (pp. 48-54). Nijmegen: Max Planck Institute for Psycholinguistics.
- Amirrachman, R. A. (2012). *Peace education in the Moluccas, Indonesia: between global models and local interests*. Unpublished doctoral dissertation, Universiteit van Amsterdam. Retrieved July 31, 2015 from <http://hdl.handle.net/11245/2.145859>.
- Andersen, R. W., & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition*, 16(2), 133-156.
- Arka, I. W. (2011). On modality and finiteness in Indonesian: complexities of =*nya* nominalisation. In *Proceedings of the International Workshop on Tense, Aspect, Mood and Evidentiality in Indonesian Languages*, Tokyo University of Foreign Studies (pp. 73-89). Retrieved August 4, 2015 from <http://lingdy.aacore.jp/en/activity/indonesiatam/timetable.html>.
- 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.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Backus, A. (1996). *Two in one. Bilingual speech of Turkish immigrants in the Netherlands*. Tilburg: Tilburg University Press.
- Backus, A. (2004). Convergence as a mechanism of language change. *Bilingualism Language and Cognition*, 7(2), 179-181.
- Backus, A., Doğruöz, A., & Heine, B. (2011). Salient stages in contact-induced grammatical change: Evidence from synchronic vs. diachronic contact situations. *Language Sciences*, 33(5), 738-752.

- Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: form, meaning, and use*. Malden, MA: Blackwell.
- Bartels, D. (1986). Can the train ever be stopped again? Developments in the Moluccan community in the Netherlands before and after the hijackings. *Indonesia*, 41, 23-45.
- Baxter, A. N. (1990). Some observations on verb serialization in Malacca Creole Portuguese. *Boletim de Filologia*, 31, 161-184.
- Benmamoun, E., Montrul, S., & Polinsky, M. (2010). White paper: Prolegomena to heritage linguistics. Harvard University. Retrieved, August 4, 2015 from <http://www.international.ucla.edu/media/files/hl-whitepaper.pdf>.
- Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2013). From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition*, 127(3), 287-306.
- Bisang, W. (2011). Grammaticalization and typology. In H. Narrog & B. Heine (Eds.), *The Oxford handbook of grammaticalization* (pp. 105-117). Oxford: Oxford University Press.
- Boas, H. C. (2003). *A constructional approach to resultatives*. Stanford, CA: CSLI Publications.
- Bohnenmeyer, J., Bowerman, M., & Brown, P. (2001). Cut and break clips. In S. C. Levinson & N. J. Enfield (Eds.), *Manual for the field season 2001* (pp. 90-96). Nijmegen, Netherlands: Max Planck Institute for Psycholinguistics.
- Boogaart, R. (1999). *Aspect and temporal ordering: A contrastive analysis of Dutch and English*. Unpublished doctoral dissertation, Vrije Universiteit, Amsterdam.
- Boumans, L. (2006). The attributive possessive in Moroccan Arabic spoken by young bilinguals in the Netherlands and their peers in Morocco. *Bilingualism: Language and Cognition*, 9(3), 213-231.

- Bowerman, M. (2011). Linguistic typology and first language acquisition. In J. J. Song (Ed.), *The Oxford handbook of linguistic typology* (pp. 591-617). Oxford: Oxford University Press.
- Bowerman, M., Gullberg, M., Majid, A., & Narasimhan, B. (2004). Put project: the cross-linguistic encoding of placement events. In A. Majid (Ed.), *Field Manual Volume 9* (pp. 10-24). Nijmegen: Max Planck Institute for Psycholinguistics.
- Bowern, C. (2008). *Linguistic fieldwork: a practical guide*. Basingstoke, England: Palgrave Macmillan.
- Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In G. Bouma, I. Krämer, & J. Zwarts (Eds.), *Cognitive foundations of interpretation* (pp. 69-94). Amsterdam: KNAW.
- Broekhuis, H., Corver, N., & Vos, R. (2015). *Syntax of Dutch. Verbs and Verb Phrases: Vol. 1*. Amsterdam: Amsterdam University Press.
- Brown, J. R., & Putnam, M. T. (2015). Functional convergence and extension in contact: Syntactic and semantic attributes of the progressive aspect in Pennsylvania Dutch. In J. B. Salmons & J. C. Johannessen (Eds.), *Germanic heritage languages in North America: Acquisition and change* (pp. 135-160). Amsterdam and Philadelphia: John Benjamins.
- Bruyn, A., Muysken, P. C., & Verrips, M. (1999). Double-object constructions in the creole languages: development and acquisition. In M. DeGraff (Ed.), *Language creation and language change* (pp. 329-37). Cambridge, MA: MIT Press.
- Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 82(4), 711-733.
- Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: The University of Chicago Press.

- Bylund, E., & Jarvis, S. (2011). L2 effects on L1 event conceptualization. *Bilingualism: Language and Cognition*, 14(1), 47-59.
- Carreira, M. (2004). Seeking explanatory adequacy: A dual approach to understanding the term “heritage language learner”. *Heritage Language Journal*, 2(1), 1-25.
- Carreira, M., & Kagan, O. (2011). The results of the National Heritage Language Survey: Implications for teaching, curriculum design, and professional development. *Foreign Language Annals*, 44(1), 40-64.
- Carroll, M., & von Stutterheim, C. (2003). Typology and information organisation: perspective taking and language-specific effects in the construal of events. In A. Giacalone Ramat (Ed.), *Typology and second language acquisition* (pp. 365-403). Berlin: Mouton de Gruyter.
- Chambers, J. K., Trudgill, P., & Schilling-Estes, N. (Eds.) (2002). *The handbook of language variation and change*. Malden, MA: Blackwell.
- Chau, W. C. (2011). The Influence of Dutch on Cantonese in the Netherlands. Unpublished Thesis, Universiteit Leiden.
- Chevalier, J. F. (2004). Heritage language literacy: Theory and practice. *Heritage Language Journal*, 2(1), 1-19.
- Chinen, K., & Tucker, G. R. (2005). Heritage language development: Understanding the roles of ethnic identity and Saturday school participation. *Heritage Language Journal*, 3(1), 27-59.
- Cleary-Kemp, J. (2007). Universal uses of demonstratives: evidence from four Malayo-Polynesian languages. *Oceanic Linguistics*, 46(2), 325-347.
- Colleman, T. (2006). *De Nederlandse datiefalternatie. Een constructioneel en corpus-gebaseerd onderzoek*. Unpublished doctoral dissertation, Universiteit Gent.

- Colleman, T. (2009). Verb disposition in argument structure alternations: a corpus study of the dative alternation in Dutch. *Language Sciences*, 31, 593-611.
- Colleman, T., & Bernolet, S. (2012). Alternation biases in corpora vs. picture description experiments: DO-biased and PD-biased verbs in the Dutch dative alternation. In D. Divjak & S. T. Gries (Eds.), *Frequency effects in language representation* (pp. 87-125). Berlin and Boston: Walter de Gruyter.
- Collins, J.T. (1974). Catatan ringkas tentang bahasa Ambon. *Dewan Bahasa*, 18(4), 151-162.
- Collins, J. T. (1980a). *Ambonese Malay and creolization theory*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Collins, J.T. (1980b). Laha, a language of the Central Moluccas. *Indonesia Circle*, 23, 3-19.
- Collins, J. T. (1981). Pertembungan linguistik di Indonesia Timur: Bahasa Melayu dan bahasa Asilulu di pulau Ambon. *Dewan Bahasa*, 25(7), 30-55.
- Collins, J. T. (1983). Syntactic change in Ambonese Malay: The possessive construction. In J. T. Collins (Ed.), *Studies in Malay dialects: Part 2, NUSA*, 17, (pp. 28-41). Jakarta: Badan Penyelenggara Seri Nusa.
- Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.
- Croft, W. (2000). *Explaining language change: An evolutionary approach*. Harlow, England: Pearson Education.
- Croft, W., Barðdal, J., Hollmann, W., Sotirova, V., & Taoka, C. (2010). Revising Talmy's typological classification of complex event constructions. In H. C. Boas (Ed.), *Contrastive studies in construction grammar* (pp. 201-235). Amsterdam and Philadelphia: John Benjamins.

- Cuyckens, H. (1993). The Dutch spatial preposition “in”: A cognitive semantic analysis. In C. Zelinsky-Wibbelt (Ed.), *The semantics of prepositions: From mental processing to natural language processing* (pp. 27-71). Berlin: Mouton de Gruyter.
- De Hoop, H., Haverkort, M., & Van den Noort, M. (2004). Variation in form versus variation in meaning. *Lingua*, 114(9), 1071-1089.
- de Prada Pérez, A. (2010). Subject position in Spanish in contact with Catalan: Language similarity vs. interface vulnerability. In M. Iverson, I. Ivanov, T. Judy, J. Rothman, R. Slabakova, & M. Tryzna (Eds.), *Proceedings of the 2009 Mind/Context Divide Workshop* (pp. 104-115). Somerville, MA: Cascadilla Proceedings Project.
- de Prada Pérez, A. (2015). Spanish-Catalan bilingual subject pronoun production and the Vulnerability Hypothesis. Paper presented at the Bilingualism in the Hispanic and Lusophone World Conference, Leiden University, The Netherlands. Retrieved on August 4, 2015 from <http://media.leidenuniv.nl/legacy/booklet-final-14-01-2015.pdf>.
- de Prada Pérez, A., & y Cabo, D. P. (2011). Invariable *gusta* in the Spanish of heritage Speakers in the U.S. In J. Herschensohn & D. Tanner (Eds.), *Proceedings of the Generative Approaches to Second Language Acquisition* (pp. 110-120). Somerville, MA: Cascadilla Proceedings Project.
- den Ouden, D. B., Fix, S., Parrish, T. B., & Thompson, C. K. (2009). Argument structure effects in action verb naming in static and dynamic conditions. *Journal of Neurolinguistics*, 22(2), 196-215.
- Donaldson, B. (2008). *Dutch: A comprehensive grammar* (2nd ed). London and New York: Routledge.
- Donohue, M. (2011). Papuan Malay of New Guinea. In C. Lefebvre (Ed.), *Creoles, their substrates, and language typology* (pp. 413-436). Amsterdam and Philadelphia: John Benjamins.

- Dryer, M. S. (2007). Clause types. In T. Shopen (Ed.), *Language typology and syntactic description: Clause structure* (pp. 224–275). Cambridge: Cambridge University Press.
- Ellis, N. (2006). Language acquisition as rational contingency learning. *Applied Linguistics*, 27(1), 1-24.
- Enfield, N. J. (2003). *Linguistic epidemiology: Semantics and grammar of language contact in Mainland Southeast Asia*. London: RoutledgeCurzon.
- Englebretson, R. (2003). *Searching for structure: The problem of complementation in colloquial Indonesian conversation*. Amsterdam and Philadelphia: John Benjamins.
- Evans, N., Levinson, S. C., Enfield, N. J., Gaby A., & A. Majid. (2004). Reciprocal constructions and situation type. In A. Majid (Ed.), *Field Manual Volume 9* (pp. 25-30). Nijmegen: Max Planck Institute for Psycholinguistics.
- Ewing, M. C. (2005). Colloquial Indonesian. In K. A. Adelaar & N. Himmelmann (Eds.), *The Austronesian languages of Asia and Madagascar* (pp. 227-254). London and New York: Routledge.
- Field, A. (2005). *Discovering statistics with SPSS* (2nd edition). London: SAGE Publications.
- Filip, H. (2011). Aspectual class and aktionsart. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning*. Berlin: De Gruyter Mouton
- Fishman, J. A. (Ed.) (2001). *Can threatened languages be saved? Reversing language shift, revisited: A 21st century perspective*. Clevedon, England: Multilingual Matters.
- Flecken, M. (2010). *Event conceptualization in language production of early bilinguals*. (LOT Dissertation Series 256). Utrecht: LOT.

- Florey, M., & van Engelenhoven, A. (2001). Language documentation and maintenance programs for Moluccan languages in the Netherlands. *International Journal of the Sociology of Language*, 151, 195-219.
- Gast, V., & van der Auwera, J. (2012). What is 'contact-induced grammaticalization?' Evidence from Mayan and Mixe-Zoquean languages. In B. Wiemer, B. Wälchli, & B. Hansen (Eds.), *Grammatical replication and borrowability in language contact* (pp. 381-426). Berlin and Boston: Walter de Gruyter.
- Gil, D. (2001). Creoles, complexity and Riau Indonesian. *Linguistic typology*, 5, 325-371.
- Goldberg, A. E., & Jackendoff, R. (2004). The English resultative as a family of constructions. *Language*, 80(3), 532-568.
- Gordon, M., Ghushchyan, E., McDonnell, B., Rosenblum, D., & Shaw, P. A. (2012). Sonority and central vowels: A cross-linguistic phonetic study. In S. Parker (Ed.), *The sonority controversy* (pp. 219-256). Berlin and Boston: Walter de Gruyter.
- Grangé, P. (2011). Aspect in Indonesian: free marker versus affixed or clitic markers. In *Proceedings of the International Workshop on Tense, Aspect, Mood and Evidentiality in Indonesian Languages*, Tokyo University of Foreign Studies (pp. 43-63). Retrieved August 4, 2015 from <http://lingdy.aacore.jp/en/activity/indonesiatam/timetable.html>.
- Grimes, B. D. (1991). The development and use of Ambonese Malay. *Pacific Linguistics*, A-81, 83-123.
- Haeseryn, W., Romijn, K., Geerts, G., de Rooij, J., & van den Toorn, M.C. (1997). *Algemene Nederlandse Spraakkunst*. Tweede, geheel herziene druk. Groningen: Martinus Nijhoff; Deurne: Wolters Plantyn. Consulted in the online electronic version E-ANS, online URL <http://ans.ruhosting.nl/e-ans/> [Last accessed 27 Januray 2016].

Haig, G., & Schnell, S. (2010). *Annotations using GRAID (Grammatical Relations and Animacy in Discourse)*. Pdf Manual Version 5.4. Retrieved January 4, 2016 from <http://www.linguistik.uni-kiel.de/GRAIDmanual5.4.pdf>.

Hajek, J. (2006). Serial verbs in Tetun Dili. In A. Y. Aikhenvald & R. M. Dixon (Eds.), *Serial verb constructions: A cross-linguistic typology* (pp. 239-253). Oxford: Oxford University Press.

Hartsuiker, R. J. (2013). Bilingual strategies from the perspective of a processing model. *Bilingualism: Language and Cognition*, 16(4), 737-739.

Hartsuiker, R. J., & Pickering, M. J. (2008). Language integration in bilingual sentence production. *Acta Psychologica*, 128(3), 479-489.

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(6), 409-414.

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, 75(2), B27-B39.

Haspelmath, M. (2011). On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology*, 15(3), 535-567.

Haspelmath, M., Michaelis, S. M., & the APiCS Consortium. (2013). Ditransitive constructions with 'give'. In S. M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (Eds.) *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved July 18, 2014 from <http://apics-online.info/parameters/60>.

Heine, B. (2008). Contact-induced word order change without word order change. In P. Siemund & N. Kintana (Eds.), *Language contact and contact languages* (pp. 33-60). Amsterdam and Philadelphia: John Benjamins.

Heine, B., & Kuteva, T. (2005). *Language contact and grammatical change*. Cambridge: Cambridge University Press.

- Heine, B., & Kuteva, T. (2007). *The genesis of grammar: A reconstruction*. Oxford: Oxford University Press.
- Heine, B., & Kuteva, T. (2008). Constraints on contact-induced linguistic change. *Journal of Language Contact*, 2(1), 57-90.
- Hellwig, B., & Lüpke, F. (2001). Caused positions. In S. C. Levinson & N.J. Enfield (Eds.), *Manual for the field season 2001*, (pp. 126-128). Nijmegen: Max Planck Institute for Psycholinguistics.
- Hengeveld, K. (2011). The grammaticalization of tense and aspect. In H. Narrog & B. Heine (Eds.), *The Oxford handbook of grammaticalization*, (pp. 580-594). Oxford: Oxford University Press.
- Himmelmann, N. (2005). The Austronesian languages of Asia and Madagascar: Typological characteristics. In K. A. Adelaar and N. Himmelmann (Eds.), *The Austronesian languages of Asia and Madagascar* (pp. 110–181). London and New York: Routledge.
- Hulsen, M. E. H. (2000). *Language loss and language processing: Three generations of Dutch migrants in New Zealand*. Unpublished doctoral dissertation, Radboud Universiteit Nijmegen. Retrieved August 4, 2015 from <http://repository.ubn.ru.nl/handle/2066/18901>.
- Hulsen, M., de Bot, K., & Weltens, B. (2002). “Between two worlds.” Social networks, language shift, and language processing in three generations of Dutch migrants in New Zealand. *International Journal of the Sociology of Language*, 153, 27-52.
- Huwaë, R. (1992). *Tweetaligheid in Wierden: het taalgebruik van jongeren uit een Molukse gemeenschap*. Unpublished doctoral dissertation, Universiteit van Amsterdam.
- Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition*, 12(1), 3-69.

- Irizarri van Suchtelen, P. (2014). Maintained and acquired heritage Spanish in the Netherlands: The case of dative constructions. *Applied Linguistics Review*, 5(2), 375-400.
- Jacob, J. A., & Grimes, C. E. (2011). Aspect and directionality in Kupang Malay serial verb constructions: Calquing on the grammars of substrate languages. In C. Lefebvre (Ed.), *Creoles, their substrates, and language typology* (pp. 337-366). Amsterdam and Philadelphia: John Benjamins.
- Jansen, B., Lalleman, J., & Muysken, P. C. (1981). The alternation hypothesis: acquisition of Dutch word order by Turkish and Moroccan foreign workers. *Language Learning*, 31(2), 315-336.
- Jarvis, S., & Odlin, T. (2000). Morphological type, spatial reference, and language transfer. *Studies in Second Language Acquisition*, 22(4), 535-556.
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge
- Jäger, F., Norcliffe, E., Housel, K., & Bohnemeyer J. (2008). *Follow-up experiments on sentence production in Yucatec*. Retrieved January 5, 2016 from <https://hlplab.wordpress.com/2008/07/26/follow-up-experiments-on-sentence-production-in-yucatec/>.
- Johanson, L. (2002). Contact-induced change in a code-copying framework. In M. Jones & E. Esch (Eds.), *Language change: The interplay of internal, external and extra-linguistic factors* (pp. 285-314). Berlin: Mouton de Gruyter.
- Johanson, L. (2013). *Structural factors in Turkic language contacts*. New York: Routledge.
- Kaufmann, I., & Wunderlich, D. (1998). *Cross-linguistic patterns of resultatives*. Manuscript. Heinrich-Heine Universität, Düsseldorf. Retrieved on August 4, 2015 from http://www.researchgate.net/profile/Dieter_Wunderlich/publication/.

- Kirsner, R. S. (1979). Deixis in discourse: An exploratory quantitative study of the modern Dutch demonstrative adjectives in discourse and syntax. In T. Givón (Ed.), *Discourse and Syntax* (pp. 355-375). New York: Academic Press.
- Kita, S. (1995). Recommendations for data collection for gesture studies. In D. Wilkins (Ed.), *Extensions of space and beyond: Manual for field elicitation for the 1995 field season* (pp. 35-42). Nijmegen: Max Planck Institute for Psycholinguistics.
- Klein, F. (1980). A quantitative study of syntactic and pragmatic indications of change in the Spanish of bilinguals in the U.S. In W. Labov (Ed.), *Locating language in time and space* (pp. 69-82). New York: Academic Press.
- Kluge, A. (2014). *A grammar of Papuan Malay*. (LOT Dissertation Series 361). Utrecht: LOT.
- Kondo-Brown, K. (2003). Heritage language instruction for post-secondary students from immigrant backgrounds. *Heritage Language Journal*, 1(1), 1-25.
- Koontz-Garboden, A. (2004). Language contact and Spanish aspectual expression: a formal analysis. *Lingua*, 114(9), 1291-1330.
- 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.
- Kusters, W. (2003). *Linguistic Complexity: The Influence of Social Change on Verbal Inflection* (LOT Dissertation Series 77). Utrecht: LOT.
- Labov, W. (2011). *Principles of linguistic change: Social factors*. Oxford: Blackwell.
- Laleko, O. (2010). *The syntax-pragmatics interface in language loss: Covert restructuring of aspect in heritage Russian*. Unpublished doctoral dissertation, University of Minnesota. Retrieved, July 31, 2015 from <http://www.tc.umn.edu/~olaleko/files/Dissertation.pdf>.

Laleko, O., & Polinsky, M. (2013). Marking topic or marking case: A comparative investigation of heritage Japanese and heritage Korean. *Heritage Language Journal*, 10(2), 40-62.

Lee, N., Ping, L., & Nomoto, H. (2009). Colloquial Singapore English got: Functions and substratal influences. *World Englishes*, 28(3), 293-318.

Lehmann, C. (1992). Word order change by grammaticalization. In M. Gerritsen & D. Stein (Eds.), *Internal and external factors in syntactic change* (pp. 395-416). Den Haag and Berlin: Mouton de Gruyter.

Leirissa, R. Z., Pattykaihatu, J. A., Luhukay, H., Talib, U., & Maelissa, S. (2004). *Ambonku*. Ambon: Pemerintah Kota Ambon

Lekawael, R. (2011). *The development of Moluccan Malay as a heritage language in the Netherlands*. Unpublished Thesis, Radboud Universiteit Nijmegen.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago press.

Levin, B., & Hovav, M. R. (2005). *Argument realization*. Cambridge: Cambridge University Press.

Lewis, P. M., Simons, G. F., & Fennig, C. D. (Eds.) (2015). *Ethnologue: Languages of the World* (18th edition). Dallas, TX: SIL International. Online URL <http://www.ethnologue.com>. [Last accessed 31 July 2015].

Litamahuputty, B. (1994). The use of *biking* and *kasi* in Ambonese Malay. *Cakalele*, 5, 11-31.

Litamahuputty, B. (2012). *Ternate Malay: Grammar and texts* (LOT Dissertation Series 307). Utrecht: LOT.

Ma, L. (2006). *Acquisition of the perfective aspect marker le of Mandarin Chinese in discourse by American college learners*. Unpublished doctoral dissertation, University of Iowa. Retrieved July 31, 2015 from <http://ir.uiowa.edu/etd/68/>.

- Mailoa, J. P. (2006). *Kamus bahasa harian dialek orang Ambon*. Jakarta: Kulibia Printing.
- Malchukov, A., Haspelmath, M., & Comrie, B. (2010). Ditransitive constructions: a typological overview. In A. Malchukov, M. Haspelmath & B. Comrie (Eds.), *Studies in ditransitive constructions: A comparative handbook* (pp. 1-64). Berlin and New York: Mouton de Gruyter.
- Maskikit-Essed, R., & Gussenhoven, C. (forthcoming). No stress, no pitch accent, no prosodic focus: The case of Ambonese Malay. *Phonology*.
- Matras, Y. (2009). *Language contact*. Cambridge: Cambridge University Press.
- Mazurkewich, I., & White, L. (1984). The acquisition of the dative alternation: Unlearning overgeneralizations. *Cognition*, 16, 261-283.
- McWhorter, J. (2001a). The world's simplest grammars are creole grammars. *Linguistic Typology*, 5(2/3), 125-166.
- McWhorter, J. (2001b). What people ask David Gil and why: Rejoinder to the replies. *Linguistic Typology*, 5(2-3), 388-412.
- Michaelis, S. M., & Haspelmath, M. (2003). Ditransitive constructions: Creole languages in a cross-linguistic perspective. *Creolica*, Retrieved July 18, 2014 from <http://www.creolica.net/>.
- Michaelis, S. M., Maurer, P., Haspelmath, M., & Huber, M. (Eds.). (2015) *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online URL <http://apics-online.info/> [Last accessed 31 July 2015].
- Milroy, J., & Milroy, L. (1985). Linguistic change, social network and speaker innovation. *Journal of Linguistics*, 21(2), 339-384.
- Milroy, L., & Milroy, J. (1992). Social network and social class: Toward an integrated sociolinguistic model. *Language in Society*, 21(1), 1-26.

- Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition*, 7(2), 125-142.
- Montrul, S. (2005). Second language acquisition and first language loss in adult early bilinguals: Exploring some differences and similarities. *Second Language Research*, 21(3), 199-249.
- Montrul, S. (2008). *Incomplete acquisition in bilingualism. Re-examining the age factor*. Amsterdam and Philadelphia: John Benjamins.
- Montrul, S. (2009). Knowledge of tense-aspect and mood in Spanish heritage speakers. *International Journal of Bilingualism*, 13(2), 239-269.
- Montrul, S. (2010). Dominant language transfer in adult second language learners and heritage speakers. *Second Language Research*, 26(3), 293-327.
- Montrul, S. (2011). Morphological errors in Spanish second language learners and heritage speakers. *Studies in Second Language Acquisition*, 33(2), 163-192.
- Montrul, S. (2012). Is the heritage language like a second language?. *Eurosla Yearbook*, 12(1), 1-29.
- 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.
- Montrul, S., & Ionin, T. (2010). Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism: Language and Cognition*, 13(4), 449-473.
- Montrul, S., & Sánchez-Walker, N. (2013). Differential object marking in child and adult Spanish heritage speakers. *Language Acquisition*, 20(2), 109-132.
- Moro, F. R. (2014). Resultative constructions in heritage Ambon Malay in the Netherlands. *Linguistics in the Netherlands*, 31(1), 78-92.

- Moro, F. R. (2015). Modal categories in contact: the case of heritage Ambon Malay in the Netherlands. *Heritage Language Journal*, 12(3), 271-291.
- Moro, F. R, & Klamer, M. (2015). Give-constructions in heritage Ambon Malay in the Netherlands. *Journal of Language Contact*, 8(2), 263-298.
- Moro, F. R. (in press). Aspectual distinctions in Dutch-Ambon Malay bilingual heritage speakers. *International Journal of Bilingualism*.
- Moro, F. R. & Irizarri van Suchtelen, P. (forthcoming). Dominant language transfer in heritage languages: Redefining the 'structural', and the 'transfer' in 'structural transfer'. In H. Peukert, T. Kupisch, K. Bührig, & I. Gogolin (Eds.), *Dynamics of Linguistic Diversity*. Amsterdam and Philadelphia: John Benjamins.
- Muysken, P. C. (2005). Possessive constructions in language contact settings. *Monash University Linguistics Papers*, 4(1), 11-19.
- Muysken, P. C. (2010a). Ethnolects as a multidimensional phenomenon. In M. Norde, B. de Jonge & C. Hasselblatt (Eds.), *Language contact: New perspectives* (pp. 7-26). Amsterdam and Philadelphia: John Benjamins.
- Muysken, P. C. (2010b). Scenarios for Language Contact. In R. Hickey (Ed.), *The handbook of language contact* (pp. 265-281). Malden, MA: Wiley-Blackwell.
- Muysken, P. C. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition*, 16(4), 709-730.
- Nagy, N. (2015). A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua*, 164, 309-327.
- Nivens, R. (1994). Written Ambonese Malay, 1895-1992. *Cakalele*, 5, 33 - 75.
- Nordhoff, S. (2009). *A grammar of upcountry Sri Lanka Malay*. (LOT Dissertation Series 226). Utrecht: LOT.

- O'Grady, W. (2012). Designing acquisition experiments for heritage language populations. Paper presented at The Sixth Heritage Language Research Institute, University of California, Los Angeles. Retrieved August 4, 2015 from <http://www.nhlrc.ucla.edu/nhlrc/events/institute/2012/schedule>.
- O'Grady, W., Lee, O. S., & Lee, J. H. (2011). Practical and theoretical issues in the study of heritage language acquisition. *Heritage Language Journal*, 8(3), 23-40.
- O'Grady, W., Kwak, H., Lee, O., & Lee, M. (2011). An emergentist perspective on heritage language acquisition. *Studies in Second Language Acquisition*, 33(2), 223-245.
- Olsson, B. (2013). *Iamitives: Perfects in Southeast Asia and beyond*. Unpublished Thesis, University of Stockholm.
- Onar Valk, P., & Backus, A. (2013). Syntactic change in an immigrant language: from non-finite to finite subordinate clauses in Turkish. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 4(2), 7-29.
- Onar Valk, P. (2015). *Transformation in Dutch Turkish Subordination? Converging evidence of change regarding finiteness and word order in complex clauses*. (LOT Dissertation Series 394). Utrecht: LOT.
- Paauw, S. (2008). *The Malay contact varieties of Eastern Indonesia: a typological comparison*. Unpublished doctoral dissertation, State University of New York at Buffalo. Retrieved July 31, 2015 from <http://wings.buffalo.edu/linguistics/people/faculty/dryer/dryer/PaauwMalayIndonesia.pdf>.
- Payne, T. E. (1997). *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39(4), 633-651.

- Philippa, M., Debrabandere, F., Quak, A., Schoonheim, T., & van der Sijs, N. (2011). *Etymologisch Woordenboek van het Nederlands*. Online URL <http://etymologie.nl/> [Last accessed 4 August 2015].
- Pires, A., & Rothman, J. (2009). Disentangling sources of incomplete acquisition: An explanation for competence divergence across heritage grammars. *International Journal of Bilingualism*, 13(2), 211-238.
- Polinsky, M. (1998). A non-syntactic account of some asymmetries in the double object construction. In J. Koenig (Ed.), *Discourse and cognition: Bridging the gap* (pp. 403-422). Stanford, CA: CSLI Publications.
- Polinsky, M. (2006). Incomplete acquisition: American Russian. *Journal of Slavic linguistics*, 14(2), 191-262.
- Polinsky, M. (2008a). Gender under incomplete acquisition. *Heritage Language Journal*, 6(1), 40-71.
- Polinsky, M. (2008b). Without aspect. In G. Corbett & M. Noonan (Eds.), *Case and grammatical relations* (pp. 263-282). Amsterdam and Philadelphia: John Benjamins.
- Polinsky, M. (2008c). Heritage language narratives. In D. Brinton, O. Kagan & S. Bauckus (Eds.), *Heritage language education: A new field emerging*. New York: Routledge.
- Polinsky, M. (2011). Reanalysis in adult heritage language: A case for attrition. *Studies in Second Language Acquisition*, 33(2), 305-328.
- Polinsky, M., & Kagan, O. (2007). Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass*, 1(5), 368-395.
- Raschka, C., Wei, L., & Lee, S. (2002). Bilingual development and social networks of British-born Chinese children. *International Journal of the Sociology of Language*, 153, 9-26.

- Rice, S., & Kabata, K. (2007). Crosslinguistic grammaticalization patterns of the allative. *Linguistic Typology*, 11(3), 451-514.
- Ross, M. (2007). Calquing and metatypy. *Journal of Language Contact*, 1(1), 116-143.
- Şahin, H. (2015). Cross-linguistic influences: Dutch in contact with Papiamentu and Turkish. (LOT Dissertation Series 405). Utrecht: LOT.
- Sakel, J. (2007). Types of loan: Matter and pattern. In Y. Matras & J. Sakel (Eds.), *Grammatical borrowing in cross-linguistic perspective* (pp. 15-29). Berlin and New York: Mouton de Gruyter.
- Sánchez, L. (2004). Functional convergence in the tense, evidentiality and aspectual systems of Quechua Spanish bilinguals. *Bilingualism: Language and Cognition*, 7(02), 147-162.
- Sánchez, L. (2006). Bilingual grammars and creoles: Similarities between functional convergence and morphological elaboration. In C. Lefebvre, L. White & C. Jourdan (Eds.), *L2 Acquisition and Creole Genesis* (pp. 277-294). Amsterdam and Philadelphia: John Benjamins.
- Schmid, M. S. (2007). The role of L1 use for L1 attrition. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language attrition: Theoretical perspectives* (pp. 135-153). Amsterdam and Philadelphia: John Benjamins.
- Schmid, M. S. (2011). Contact x time: External factors and variability in L1 attrition. In M. S. Schmid & W. Lowie (Eds.), *Modeling bilingualism: From structure to chaos. In honor of Kees de Bot* (pp. 155-176). Amsterdam and Philadelphia: John Benjamins.
- Schmiedtová, B., von Stutterheim, C., & Carroll, M. (2011). Language-specific patterns in event construal of advanced second language speakers. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (pp. 66-107). Bristol: Multilingual Matters.

- Schoenmakers-Klein Gunnewiek, M. (1997). Dutch language loss in Brazil and the Conceptual Hypothesis. In H. Klatter & S. Kroon (Eds.), *Dutch overseas: Studies in maintenance and loss of Dutch as an immigrant language* (Studies in Multilingualism). (pp. 115-140). Tilburg: Tilburg University Press.
- 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(2), 153-171.
- Seliger, H. W., & Vago, R. M. (1991). *First language attrition*. Cambridge: Cambridge University Press.
- Senft, G. (Ed.). (2008). *Serial verb constructions in Austronesian and Papuan languages*. Canberra: Pacific Linguistics.
- Shi, M. (2011) *Incomplete knowledge of aspect in heritage speakers of Mandarin Chinese in the Netherlands*. Unpublished Thesis, Radboud Universiteit Nijmegen.
- Silva-Corvalán, C. (1993). On the permeability of grammars: evidence from Spanish and English contact. In W. J. Ashby, M. Mithun & G. Perissinotto (Eds.), *Linguistic perspectives on romance languages* (pp. 19-43). Amsterdam and Philadelphia: John Benjamins.
- Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.
- Silva-Corvalán, C. (2008). The limits of convergence in language contact. *Journal of Language Contact*, 2(1), 213-224.
- Slobin, D. (1991). Learning to think for speaking: Homeland language, cognition, and rhetorical style. *Pragmatics*, 1, 7-26.
- Slomanson, P. (2006). Sri Lankan Malay morphosyntax. In A. Deumert & S. Durrleman-Tame (Eds.), *Structure and variation in language contact* (pp. 135-158). Amsterdam and Philadelphia: John Benjamins.

- Slomanson, P. (2011). Dravidian features in the Sri Lanka Malay verb. In C. Lefebvre (Ed.), *Creoles, their substrates, and language typology* (pp. 383-411). Amsterdam and Philadelphia: John Benjamins.
- Smeets, H., & Veenman, J. (2000). More and more at home: Three generations of Moluccans in the Netherlands. In H. Vermeulen & R. Penninx (Eds.), *Immigrant integration: The Dutch case* (pp. 36-63) Amsterdam: Het Spinhuis.
- Sneddon, J. N. (2003a). Diglossia in Indonesian. *Bijdragen tot de Taal-, Land-en Volkenkunde*, 159(4), 519-549.
- Sneddon, J. N. (2003b). *The Indonesian language: Its history and role in modern society*. Sydney: UNSW Press.
- Soolsma, R. (2013). *De progressief in het Nederlands en het Moluks Maleis*. Unpublished Thesis, Radboud Universiteit Nijmegen.
- Sorace, A. (2005). Selective optionality in language development. In L. Cornips & K. P. Corrigan (Eds.), *Syntax and variation: Reconciling the biological and the social* (pp. 55-80). Amsterdam and Philadelphia: John Benjamins.
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1-33.
- Tagliamonte, S. A. (2006). *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.
- Tahitu, B. (1988). The sound system of Melaju Sini: Malay as spoken by younger Moluccans in the Netherlands. *Bijdragen tot de Taal-, Land-en Volkenkunde*, 144(2/3), 276-296.
- Tahitu, B. (1989). *Melaju Sini. Het Maleis van Molukse jongeren in Nederland*. Unpublished doctoral dissertation, Universiteit Leiden.
- Tahitu, B. (2000). *Daftar kata-kata Moluks Maleis-Nederlands*. Utrecht: Landelijk Steunpunt Educatie Molukkers.

- Tahitu, B., & Lasomer, X. (2001). Moluks-Maleis. In G. Extra & J. J. de Ruiter (Eds.), *Babylon aan de Noordzee: Nieuwe talen in Nederland* (pp. 153-174). Amsterdam: Bulaaq.
- Takaria, D., Pieter, C., & Muis, M. (1998). *Kamus bahasa Melayu Ambon-Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Talmy, L. (2000). *Toward a cognitive semantics: Typology and process in concept structuring*. Cambridge, MA: MIT Press.
- Theijssen, D. (2012). *Making choices: Modelling the English dative alternation*. Unpublished doctoral dissertation, Radboud Universiteit Nijmegen.
- Thomason, S. G. (2001). *Language contact*. Edinburgh: Edinburgh University Press.
- Thomason, S., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Thurgood, E. (2001). The development of articles in Baba Malay. *Anthropological Linguistics*, 43(4), 471-490.
- Tjia, J. (1992). Partikel-partikel Dalam Klausa, Kalimat dan Wacana Melayu Ambon: Catatan Pendahuluan. *Cakalele*, 3, 43-61.
- Tjia, J. (1997). *Verb serialization in Ambonese Malay*. Unpublished Thesis, University of Oregon.
- Tjia, J. (2004). The development of possessive and passive-like constructions in Ambonese language. *Studies in Malay Dialects, NUSA*, 53, 51-62.
- Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive linguistics*, 11(1), 61-82.
- Tomasello, M. (2009). *Constructing a language: A usage-based theory of language acquisition*. Harvard: Harvard University Press.

- Torres Cacoullos, R. (2000). *Grammaticization, synchronic variation, and language contact: A study of Spanish progressive-ndo constructions*. Amsterdam and Philadelphia: John Benjamins.
- TNS Opinion & Social (2012). *Europeans and their languages. Special Eurobarometer 386*. Retrieved on June 30 2015, from ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf.
- Trudgill, P. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society*, 1(2), 179-195.
- Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Unsworth, S., Argyri, F., Cornips, L., Hulk, A., Sorace, A., & Tsimpli, I. (2014). The role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics*, 35(4), 765-805.
- Valdés, G. (2000). Spanish for native speakers: Vol 1. *AATSP Professional development series handbook for teachers K-16*. New York: Harcourt College Publishers.
- van Deusen-Scholl, N. (2003). Toward a definition of heritage language: Sociopolitical and pedagogical considerations. *Journal of Language, Identity, And Education*, 2(3), 211-230.
- van Engelenhoven, A. (2008). *Ini apel ni nya* 'This here apple now'. Deictics in the Malay speech of Southwest Malukan migrants in the Netherlands. *Wacana, Jurnal Ilmu Pengetahuan Budaya*, 10(1), 90-116.
- van Hout, A., de Swart, H., & Verkuyl, H. J. (2005). Introducing perspectives on aspect. In H. J. Verkuyl, H. De Swart & A. van Hout (Eds.), *Perspectives on aspect* (pp. 1-17). Dordrecht, Netherlands: Springer.
- van Kemenade, A., & Los, B. (2003). Particles and prefixes in Dutch and English. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology* (pp. 79-117). New York: Kluwer Academic.

- van Minde, D. (1997). *Malayu Ambong*. Unpublished doctoral dissertation, Research School CNWS, Universiteit Leiden.
- van Minde, D. (2002). European loan words in Ambonese Malay. In K. A. Adelaar & R. A. Blust (Eds.), *Between worlds: linguistic paper in memory of David John Prentice* (pp. 193-216). Canberra: Pacific Linguistics.
- van Minde, D., & Tjia, J. (2002). Between perfect and perfective. The meaning and function of Ambonese Malay *su* and *suda*. *Bijdragen tot de taal-, land- en volkenkunde*, 158(2), 283-303.
- van Rijswijk, R., Muntendam A., & Dijkstra, T. (in preparation). *Heritage speakers, their L1 and their L2: Towards a new definition*.
- van Staden, M., & Reesink, G. (2008). Serial verb constructions in a linguistic area. In G. Senft (Ed.), *Serial verb constructions in Austronesian and Papuan languages* (pp. 17-54). Canberra: Pacific Linguistics.
- van Wagtenonk, J. (2008). *The first modern diaspora community in the Netherlands: Lessons from the Moluccan experience*. (Independent Study Project Collection). Paper 47. Retrieved on July 31, 2015, from http://digitalcollections.sit.edu/isp_collection/47/.
- Veenman, J. (1994). *The social integration of Moluccans*. Rotterdam: Koninklijke Vermande/ ISEO.
- Veenstra, T. (1996). *Serial verbs in Saramaccan. Predication and creole genesis*. Den Haag: Holland Academic Graphics.
- Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.
- Verkuyl, H. J. (1989). Aspectual classes and aspectual composition. *Linguistics and Philosophy*, 12(1), 39-94.
- Vermeulen, H., & Penninx, R., (Eds.) (2000). *Immigrant integration: The Dutch case*. Amsterdam: Het Spinhuis.

- Voigt, H. (1994). *Code-wisseling, taalverschuiving en taalverandering in het Melaju Sini*. Unpublished doctoral dissertation, University of Tilburg.
- von Stutterheim, C., & Nüse, R. (2003). Processes of conceptualisation in language production: Language-specific perspectives and event construal. *Linguistics*, 41(5; ISSU 387), 851-881.
- Vos, K. (2008). *Codeswitching onder Molukkers in Nederland*. Unpublished thesis, Rijksuniversiteit Groningen.
- Ward, J. H., Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58, 236-244.
- Wasow, T. (2002). *Postverbal behavior*. Stanford, CA: CSLI Publications.
- Wei, L. (1994). *Three generations, two languages, one family: Language choice and language shift in a Chinese community in Britain*. Clevedon, England: Multilingual Matters.
- Wei, L., Milroy, L., & Ching, P. S. (2000). A two-step sociolinguistic analysis of codeswitching and language choice: the example of a bilingual Chinese community in Britain. In L. Wei (Ed.), *The bilingualism reader* (pp. 188-209). London & New York: Routledge.
- Weinreich, U. (1979). *Languages in contact: Findings and problems*. Den Haag: Walter de Gruyter.
- Whong-Barr, M., & Schwartz, B. D. (2002). Morphological and syntactic transfer in child L2 acquisition of the English dative alternation. *Studies in Second Language Acquisition*, 24, 579-616.
- Wible, D., & Huang, P. Y. (2003). Using learner corpora to examine L2 acquisition of tense-aspect markings. In Proceedings of the Corpus Linguistics 2003 Conference (pp. 889-898). Lancaster University: University Centre for Computer Corpus Research on Language. Retrieved August 4, 2015 from

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.498.5658&rep=rep1&type=pdf>.

- Winford, D. (2003). *An introduction to contact linguistics*. Malden, MA: Blackwell.
- Winford, D. (2012). *Creole Languages*. In R. I. Binnick (Ed.), *The Oxford Handbook of Tense and Aspect*. (pp. 428-457). Oxford: Oxford University Press.
- Yip, V., & Matthews, S. (2007). *The bilingual child: Early development and language contact*. Cambridge: Cambridge University Press.
- Zdorenko, T., & Paradis, J. (2008). The acquisition of articles in child second language English: fluctuation, transfer or both?. *Second Language Research*, 24(2), 227-250.
- Zwarts, J. (1995). Lexical and functional direction in Dutch. *Linguistics in the Netherlands*, 12(1), 227-238.

Appendix 1:

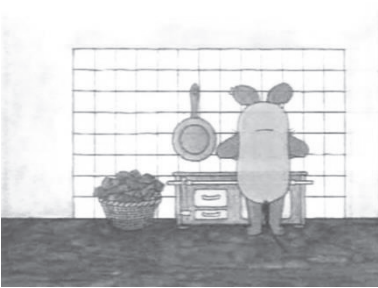
Stimuli for the simultaneous video description task (Chapter 3 and Chapter 4)



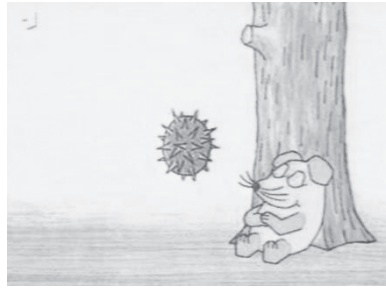
The mouse wants to eat a banana⁹¹



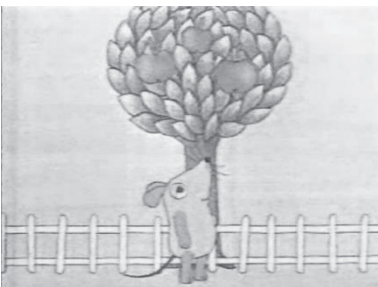
The mouse plays guitar⁹¹



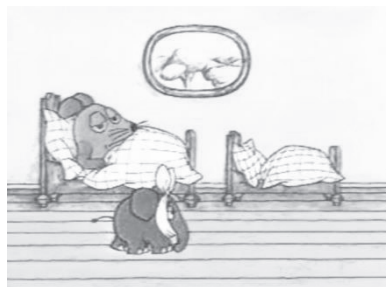
The mouse tries to cook a pancake⁹¹



A fruit falls next to the mouse⁹¹

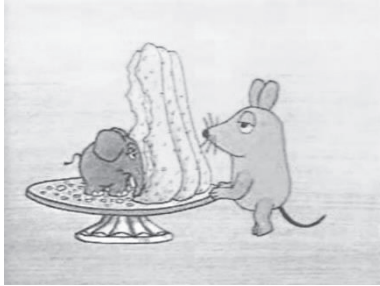


The mouse tries to pick an apple from a tree⁹¹

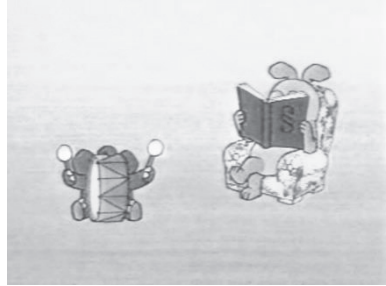


The mouse tries to sleep but is disturbed by the elephant⁹¹

⁹¹ SOURCE: The German children's series *Die Sendung mit der Maus*.



The mouse wants to eat a cake but the elephant, who is hiding behind the cake, has already eaten half of it⁹¹



The mouse tries to read, the elephant annoys him by playing drum⁹¹



A man washes his hands⁹²



A man washes an apple and eats it⁹²



A man lights a match⁹²



A man plays with a ball, two children arrive and steal his ball⁹³

⁹² SOURCE: Haig, G., & Schnell, S. (2010). *Annotations using GRAID (Grammatical Relations and Animacy in Discourse)*. Pdf Manual Version 5.4.

⁹³ SOURCE: S. Kita. (1995). Recommendations for data collection for gesture studies. In D. Wilkins (Ed.), *Extensions of space and beyond: Manual for field elicitation for the 1995 field season* (pp. 35-42). Nijmegen: Max Planck Institute for Psycholinguistics.



Two boys try to grab a t-shirt stuck on a tree⁹³



A man puts three balls in a box but a boy steals the balls⁹³

Appendix 2:

Stimuli for the video-clip retelling task

2.1 Video-clips used in Chapter 4

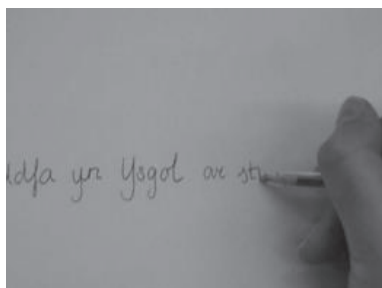
Activity video-clips



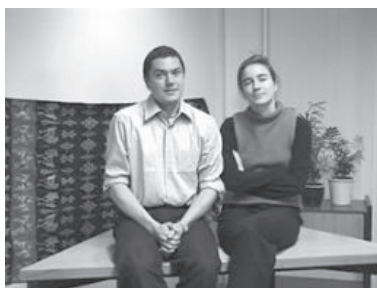
Sew⁹⁴



Swim⁹⁴



Write⁹⁴



Sit on table⁹⁵

⁹⁴ **SOURCE:** M. Starren and the team members of the NWO Project “Grammaticised forms underlying information structure: Hurdles for advanced learners in achieving native-like competence” (2005-2012), online URL <http://www.nwo.nl/en/research-and-results/research-projects/i/68/968.html> [Last accessed 14 January 2016].

⁹⁵ **SOURCE:** N. Evans, S. C. Levinson, N. J. Enfield, A. Gaby & A. Majid. (2004). Reciprocal constructions and situation type. In A. Majid (Ed.), *Field Manual Volume 9* (pp. 25-30). Nijmegen: Max Planck Institute for Psycholinguistics.



Sleep⁹⁶

Accomplishment video-clips



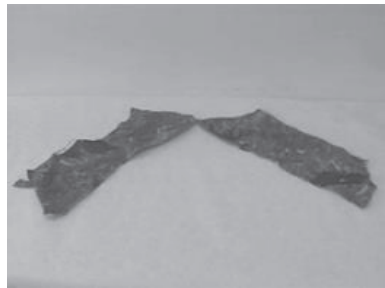
Cut off branch⁹⁷



Cut fish⁹⁷



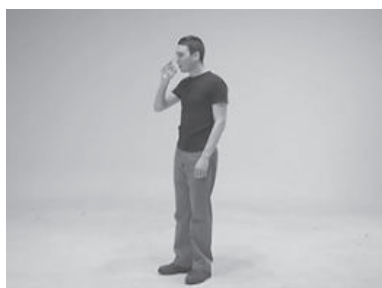
Cut hair⁹⁷



Tear magically⁹⁷

⁹⁶ SOURCE: D. B. den Ouden and colleagues, Northwestern University, IL, USA. See D. B. den Ouden, S. Fix, T. B. Parrish, & C. K. Thompson (2009). Argument structure effects in action verb naming in static and dynamic conditions. *Journal of Neurolinguistics*, 22(2), 196-215.

⁹⁷ SOURCE: J. Bohnemeyer, M. Bowerman & P. Brown. (2001). Cut and break clips. In S. C. Levinson & N. J. Enfield (Eds.), *Manual for the field season 2001* (pp. 90-96). Nijmegen: Max Planck Institute for Psycholinguistics.

Tear cloth⁹⁷Cut carrot⁹⁷Descend stairs⁹⁸Lift bucket⁹⁹Drink water⁹⁹Handshake¹⁰⁰

⁹⁸ **SOURCE:** M. Starren and the team members of the NWO Project “Grammaticised forms underlying information structure: Hurdles for advanced learners in achieving native-like competence” (2005-2012), online URL <http://www.nwo.nl/en/research-and-results/research-projects/i/68/968.html> [Last accessed 14 January 2016].

⁹⁹ **SOURCE:** D. B. den Ouden and colleagues, Northwestern University, IL, USA. See D. B. den Ouden, S. Fix, T. B. Parrish, & C. K. Thompson (2009). Argument structure effects in action verb naming in static and dynamic conditions. *Journal of Neurolinguistics*, 22(2), 196-215.

¹⁰⁰ **SOURCE:** N. Evans, S. C. Levinson, N. J. Enfield, A. Gaby & A. Majid. (2004). Reciprocal constructions and situation type. In A. Majid (Ed.), *Field Manual Volume 9* (pp. 25-30). Nijmegen: Max Planck Institute for Psycholinguistics.

Achievement video-clips



Break rope¹⁰¹



Break magically¹⁰¹



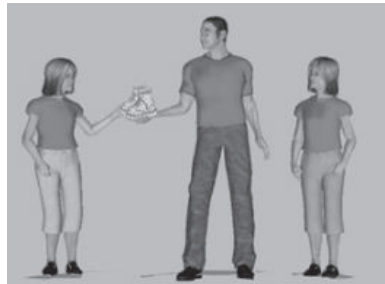
Break pot¹⁰¹



Break stick¹⁰¹



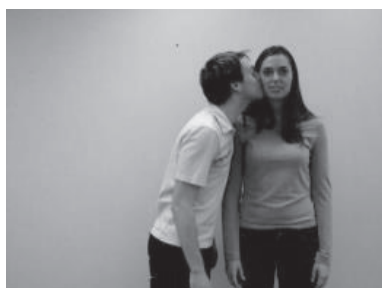
Throw ball¹⁰²



Give shoes¹⁰²

¹⁰¹ **SOURCE:** J. Bohnemeyer, M. Bowerman & P. Brown. (2001). Cut and break clips. In S. C. Levinson & N. J. Enfield (Eds.), *Manual for the field season 2001* (pp. 90-96). Nijmegen: Max Planck Institute for Psycholinguistics.

¹⁰² **SOURCE:** F. Jäger, E. Norcliffe, K. Housel, J. Bohnemeyer and colleagues, University of Rochester, NY, USA, online URL <https://hlplab.wordpress.com/2008/07/26/follow-up-experiments-on-sentence-production-in-yucatec/> [Last accessed 14 January 2016].

Give Backpack¹⁰³Kick ball¹⁰³Hit ball¹⁰³Smell flower¹⁰³Kiss woman¹⁰³Push man¹⁰³

¹⁰³ SOURCE: D. B. den Ouden and colleagues, Northwestern University, IL, USA. See D. B. den Ouden, S. Fix, T. B. Parrish, & C. K. Thompson (2009). Argument structure effects in action verb naming in static and dynamic conditions. *Journal of Neurolinguistics*, 22(2), 196-215.



Put apple¹⁰⁴



Put head¹⁰⁴

2.2 Video-clips used in Chapter 5



Show jacket¹⁰⁵



Give 1 of 2 backpacks¹⁰⁵



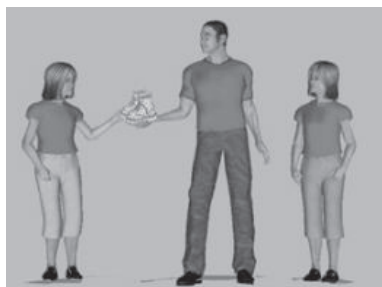
Show book¹⁰⁵



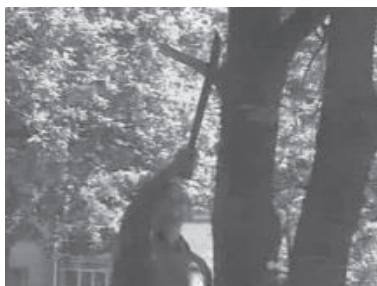
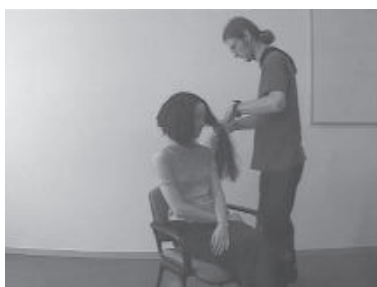
Offer cereal¹⁰⁵

¹⁰⁴ **SOURCE:** M. Bowerman, M. Gullberg, A. Majid & B. Narasimhan (2004). Put project: the cross-linguistic encoding of placement events. In A. Majid (Ed.), *Field Manual Volume 9* (pp. 10-24). Nijmegen: Max Planck Institute for Psycholinguistics.

¹⁰⁵ **SOURCE:** F. Jäger, E. Norcliffe, K. Housel, J. Bohnemeyer and colleagues, University of Rochester, NY, USA, online URL <https://hlplab.wordpress.com/2008/07/26/follow-up-experiments-on-sentence-production-in-yucatec/> [Last accessed 14 January 2016].

Give shoes¹⁰⁵Give backpacks¹⁰⁵

2.3 Video-clips used in Chapter 6

Break rope¹⁰⁶Cut off branch¹⁰⁶Cut fish¹⁰⁶Cut hair¹⁰⁶

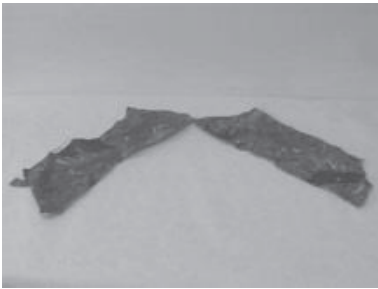
¹⁰⁶ SOURCE: J. Bohemeyer, M. Bowerman & P. Brown. (2001). Cut and break clips. In S. C. Levinson & N. J. Enfield (Eds.), *Manual for the field season 2001*, (pp. 90-96). Nijmegen: Max Planck Institute for Psycholinguistics.



Break magically¹⁰⁶



Break pot¹⁰⁶



Tear magically¹⁰⁶



Break stick¹⁰⁶

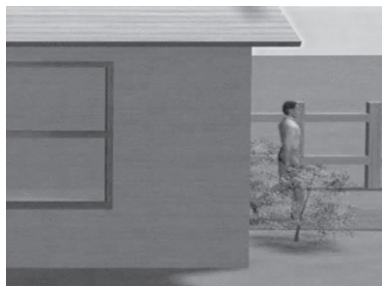


Tear cloth¹⁰⁶



Cut carrot¹⁰⁶

2.4 Video-clips used as fillers



Exit house¹⁰⁷



Enter house¹⁰⁷



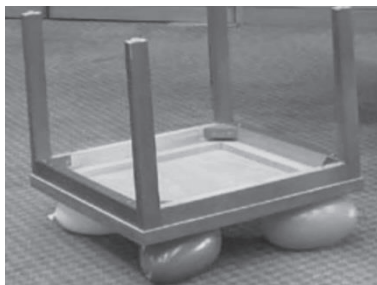
Under table¹⁰⁸



Before trunk¹⁰⁸



Squeeze ball¹⁰⁸



Squeezed under table¹⁰⁸

¹⁰⁷ SOURCE: S. Kita. (1995). Recommendations for data collection for gesture studies. In D. Wilkins (Ed.), *Extensions of space and beyond: Manual for field elicitation for the 1995 field season* (pp. 35-42). Nijmegen: Max Planck Institute for Psycholinguistics.

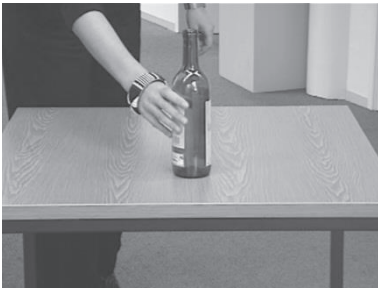
¹⁰⁸ SOURCE: F. Ameka, C. de Witte & D. Wilkins. (1999). Picture series for positional verbs: Eliciting the verbal component in locative descriptions. In D. Wilkins (Ed.), *Manual for the 1999 Field Season* (pp. 48-54). Nijmegen: Max Planck Institute for Psycholinguistics.



In basket¹⁰⁸



Ball tree¹⁰⁹



Put bottle¹⁰⁹



Stick ground¹⁰⁹



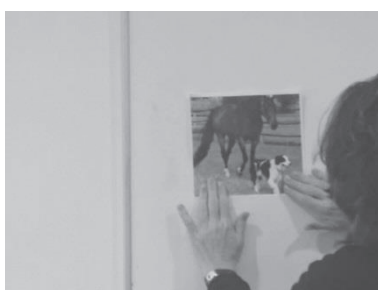
Put ladder¹⁰⁹



Books¹¹⁰

¹⁰⁹ **SOURCE:** B. Hellwig & F. Lüpke. (2001). Caused positions. In S. C. Levinson & N.J. Enfield (Eds.), *Manual for the field season 2001*, (pp. 126-128). Nijmegen: Max Planck Institute for Psycholinguistics.

¹¹⁰ **SOURCE:** N. Evans, S. C. Levinson, N. J. Enfield, A. Gaby & A. Majid. (2004). Reciprocal constructions and situation type. In A. Majid (Ed.), *Field Manual Volume 9* (pp. 25-30). Nijmegen: Max Planck Institute for Psycholinguistics.

Remove head¹¹¹Remove picture¹¹¹Take can¹¹¹Put picture¹¹¹Smile woman¹¹²Write board¹¹²

¹¹¹ **SOURCE:** M. Bowerman, M. Gullberg, A. Majid & B. Narasimhan (2004). Put project: the cross-linguistic encoding of placement events. In A. Majid (Ed.), *Field Manual Volume 9* (pp. 10-24). Nijmegen: Max Planck Institute for Psycholinguistics.

¹¹² **SOURCE:** D. B. den Ouden and colleagues, Northwestern University, IL, USA. See D. B. den Ouden, S. Fix, T. B. Parrish, & C. K. Thompson (2009). Argument structure effects in action verb naming in static and dynamic conditions. *Journal of Neurolinguistics*, 22(2), 196-215.



Lick letter¹¹²



Follow man¹¹²



Sit down¹¹²



Hug woman¹¹²



Sneeze¹¹²



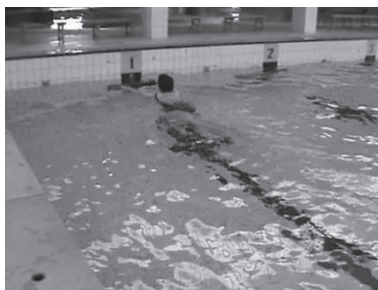
Applaud woman¹¹²



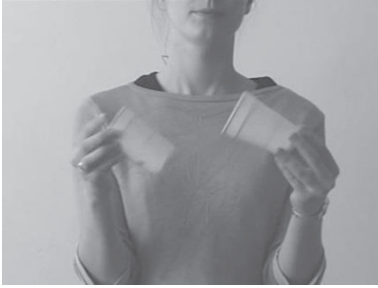
Grab arm¹¹²



Wash clothes¹¹²

Sail¹¹³Write letter¹¹³Sail to shore¹¹³Sew pants¹¹³Roll off¹¹³Swim to shore¹¹³

¹¹³ **SOURCE:** M. Starren and the team members of the NWO Project “Grammaticised forms underlying information structure: Hurdles for advanced learners in achieving native-like competence” (2005-2012), online URL <http://www.nwo.nl/en/research-and-results/research-projects/i/68/968.html> [Last accessed 14 January 2016].



Open cups¹¹⁴



Cut finger¹¹⁴



Take icecream¹¹⁵

¹¹⁴ **SOURCE:** J. Bohnemeyer, M. Bowerman & P. Brown. (2001). Cut and break clips. In S. C. Levinson & N. J. Enfield (Eds.), *Manual for the field season 2001* (pp. 90-96). Nijmegen: Max Planck Institute for Psycholinguistics.

¹¹⁵ **SOURCE:** F. Jäger, E. Norcliffe, K. Housel, J. Bohnemeyer and colleagues, University of Rochester, NY, USA, online URL <https://hlplab.wordpress.com/2008/07/26/follow-up-experiments-on-sentence-production-in-yucatec/> [Last accessed 14 January 2016].

Appendix 3:

The sociolinguistic questionnaire

Date:

Location:

Topic: *Background information*

1. What is your full name?
2. How old are you?
3. Where do you live?
4. With whom do you live?

Topic: *Partner and children*

5. Do you have a partner/spouse?
6. What is the origin of your partner/spouse?
7. Since how many years are you together/married?
8. Do you have children? How many?
9. How old are your children?

Topic: *Growing up (the participant is encouraged to talk about his/her childhood)*

10. When were you born?
11. Where were you born?
12. Where did you grow up?
13. How did you grow up?
14. With whom did you grow up?
15. Who lived at home?
16. How do/did you usually spend the holidays?

Topic: *Education and work (the participant is encouraged to talk about his/her present life)*

17. What is your education?
18. What kind of schools did you attend?
19. What is your current job?
20. What was your previous job?
21. Where have you lived?

Topic: *Parents or care givers*

22. Which is the country/island/village of origin of your mother/father?
23. What is the education of your mother/father?
24. What is/was the job of your mother/father?

Topic: *Moving to the Netherlands*¹¹⁶

25. When did you/or your parents move to the Netherlands?
26. How old were you when you moved to the Netherlands?/How old were your parents when they moved to the Netherlands?
27. With whom did you/they come?

Topic: *Link to the homeland*¹¹⁷

28. How often do you visit the Moluccas/Ambon?
29. When was the last time you have been there, and for how long did you stay?
30. How much contact do you have with friends/family there?
31. How do you communicate with them?
32. What kind of media do you watch/read/listen to? (in what languages)
33. How do you imagine your life in ten years? Do you plan to go (back) to the Moluccas/Ambon?

Topic: *Link to the Netherlands*¹¹⁸

34. Do you have relatives living in the Netherlands?
35. Which language do you speak with them?
36. What do you think about their Ambon Malay?

Topic: *Languages and language use*

37. Which languages do you speak? And how well?
38. Which language do you consider your mother tongue? That is, the language you learned at home and which you are most comfortable with.
39. When did you learn Ambon Malay? And how?
40. When did you learn Dutch? And how?

¹¹⁶ These questions were only asked to Ambon Malay speakers in the Netherlands, namely first generation speakers and heritage speakers.

¹¹⁷ See previous footnote.

¹¹⁸ These questions were only asked to Ambon Malay speakers in Ambon, Indonesia.

41. When did you learn (any additional language that the informant says to speak well)? And how?
42. Which language did you speak with your mother/father/siblings/friends when you were a child? And now?
43. Which language did you parents spoke to each other?
44. Did/do you have grandparents? Which language did/do you speak with them?
45. In which situations do you use Ambon Malay?
46. In which situations do you use Dutch?
47. In which situations do you use (any additional language that the informant says to speak well)?
48. Which language(s) do you use most?
49. Which language to you speak with your partner/spouse?
50. Do you raise your children in Ambon Malay?
51. Do your children speak Ambon Malay?

Topic: *The future of Ambon Malay in the Netherlands*¹¹⁹

52. How much is it used at present?
 53. How well is it spoken at present?
 54. What is the difference between the language spoken by the first generation and the language spoken by the other generations? Can you give an example?
 55. What do you say when someone asks about your identity? What do you feel yourself?
-
56. Can you tell me the legend *Batu Badaong*?
 57. Can you tell me about the conflict in Ambon in 1999?

¹¹⁹ These questions were only asked to Ambon Malay speakers in the Netherlands, namely first generation speakers and heritage speakers.

Nederlandse Samenvatting

In dit proefschrift wordt een vergelijking gemaakt tussen het Ambonees en Maleis als erftaal, zoals gesproken in Nederland, en de variëteit zoals gesproken op de plek van herkomst, Ambon, in Indonesië. In deze studie wordt specifiek gekeken naar de divergentie ten opzichte van de taal in het land van herkomst en convergentie naar de dominante taal, het Nederlands, waarbij het onderzoek zich richt op enkele specifieke gebieden van de grammatica van het Ambonees Maleis in Nederland en daarbij de geobserveerde patronen kwantitatief analyseert. De volgende algemene vragen komen aan de orde in dit proefschrift: Verschilt het Ambonees Maleis in Nederland van dat in het land van herkomst? Is het Ambonees Maleis in Nederland aan het veranderen onder invloed van het Nederlands? Hoe manifesteert deze verandering zich? Wat zijn de factoren die achter deze 'voortgaande' verandering zitten? In elk hoofdstuk van dit proefschrift wordt een systematische vergelijking gemaakt tussen het Ambonees Maleis in Nederland, het Ambonees Maleis zoals gesproken in het land van herkomst en het Nederlands. Als diagnostiek voor verandering worden de volgende vier linguïstische kenmerken gebruikt: nominale modificatie (volgorde van nominale modificatoren), aspectuele onderscheidingen (frequentie en distributie van aspectmarkeerders), bezitoverdracht en resultatieve constructies. Deze specifieke kenmerken zijn gekozen omdat ze een belangrijk onderdeel van de grammatica van het Ambonees Maleis in Nederland bestrijken, van woordvolgorde tot grammaticale constructies, en zodoende inzicht geven in de mate waarin de erftaal divergeert van de taal in het land van herkomst en convergeert naar het Nederlands.

In Hoofdstuk 1 wordt een overzicht gegeven van de meest relevante informatie omtrent erftalen en erftaalsprekers, de verschillende oorsprongen van divergentie en convergentie in erftalen, de werkhypotheses van dit onderzoek, de geschiedenis van de Ambonees Maleis sprekende gemeenschap in Ambonees en Nederland, en de grammatica van het Ambonees Maleis. Een substantieel deel van Hoofdstuk 1 wordt besteed aan het bespreken van de factoren waarvan is bewezen dat ze een rol spelen in het verklaren van taalverandering door taalcontact in tweetalige erftaalgemeenschappen. De voornaamste factoren die verantwoordelijk zijn voor divergentie tussen de erftaalvariëteit en de herkomstlandvariëteit zijn invloed vanuit de dominante taal, het verschil in type Maleis waaraan erftaalsprekers blootgesteld werden, en universele principes betreffende taalverwerving in contactsituaties. Erftaalsprekers zijn tweetalig in het Nederlands en het Ambonees Maleis, waarbij het

Nederlands hun functioneel dominante taal is. Het mag daarom worden verondersteld dat het Nederlands hun gebruik van het Ambonees Maleis in verschillende mate beïnvloedt in verschillende aspecten van de taal. Daarnaast worden erftaalsprekers blootgesteld aan een type input dat zowel kwantitatief als kwalitatief verschilt van die van de sprekers in het land van herkomst, de Centrale Molukken. De meerderheid van de erftaalsprekers heeft het Tangsi Maleis leren spreken, een afwijkende variëteit van het Maleis dat in 1951 met ex-KNIL soldaten en hun families naar Nederland is gekomen. Bovendien gebruiken zij de erftaal slechts in een beperkt aantal domeinen. Ten slotte spelen universele principes betreffende taalontwikkeling een belangrijke rol in de vorming van de grammatica van erftalen. Sommige van de geobserveerde veranderingen in erftalen kunnen niet verklaard worden aan de hand van de structuur van de betreffende talen, en lijken gemotiveerd te zijn door processen als simplificatie, regularisatie van paradigma's, en voorkeur voor bepaalde typen (ongemarkeerde) structuren.

In Hoofdstuk 2 wordt de methodologie uiteengezet die in het huidige onderzoek gebruikt is voor het selecteren van sprekers, het verzamelen van data en het analyseren van de resultaten. In dit hoofdstuk wordt ten eerste beschreven welke typen sprekers zijn opgenomen in de steekproef, hoe ze geselecteerd zijn, en waar en wanneer het veldwerk uitgevoerd is. Vervolgens worden in het hoofdstuk de problemen besproken omtrent het vaststellen van de juiste Ambonees Maleise referentiegroep en kwesties betreffende de dataverzameling in de eerste generatiegroep. Het tweede deel van dit hoofdstuk presenteert het elicitatiemateriaal dat in dit onderzoek gebruikt is, en de procedures betreffende transcriptie, codering en analyse van de data.

In Hoofdstuk 3 worden woordvolgordeveranderingen in het domein van nominale modificatie behandeld. In het Ambonees Maleis worden nominale modificatoren (aanwijzend voornaamwoorden, telwoorden, bijvoeglijk naamwoorden, en markeerders van definietheid) normaal gesproken achter het zelfstandig naamwoord geplaatst (de postnominale positie), maar aanwijzend voornaamwoorden en telwoorden kunnen ook vóór het zelfstandig naamwoord geplaatst worden (prenominale positie). In het Nederland worden nominale modificatoren altijd in prenominale positie geplaatst. Dit hoofdstuk laat zien dat deze gedeeltelijke overlap een concreet effect heeft op woordvolgorde, in de vorm van een toegenomen frequentie van woordvolgordepatronen die gedeeld worden met het Nederlands (d.w.z. prenominale aanwijzend voornaamwoorden en het prenominale telwoord *satu* 'één'). In dit hoofdstuk wordt ook getest of de

verandering in de lineaire volgorde van deze morfemen correleert met een grammaticale herinterpretatie op basis van het model van de Nederlandse lidwoorden. De resultaten laten zien dat er tot dusverre nog geen grammaticale herinterpretatie plaatsgevonden lijkt te hebben. Een geval waar een beginnend proces van grammaticalisatie door taalcontact mogelijk wel aan de hand zou kunnen zijn betreft de markeerder van definietheid = *nya*, die statistisch gezien frequenter gebruikt wordt door de erftaalsprekers dan door de sprekers in het land van herkomst. In dit hoofdstuk wordt beargumenteerd dat de hogere frequentie van = *nya* bij erftaalsprekers mogelijk gedeeltelijk veroorzaakt wordt door grammaticalisatie van de categorie 'definietheid' als gevolg van invloed vanuit het Nederlands, en gedeeltelijk door de afwijkende taalvariëteit waaraan erftaalsprekers worden blootgesteld, te weten het Tangsi Maleis.

Hoofdstuk 4 onderzoekt het systeem van aspectmarkering in het Ambonees Maleis in Nederland. Ten eerste worden in dit hoofdstuk de tijds- en aspectonderscheidingen van het Ambonees Maleis van het land van herkomst en het Nederlands behandeld, en vervolgens worden mogelijke gevallen van divergentie nader bekeken door de gebruiksfrequentie van aspectmarkeerders in de erftaalgroep te vergelijken met die van drie referentiegroepen, namelijk sprekers van het Ambonees Maleis in het herkomstland, eerste generatie sprekers van het Ambonees Maleis in Nederland, en eentalige sprekers van het Nederlands. De bevindingen laten zien dat de erftaalvariëteit op twee punten verschilt van de variëteit in het land van herkomst: de markeerder *ada* 'EXIST' wordt vaker en in nieuwe contexten gebruikt, terwijl de markeerder *su* 'PRF' en het proces van verbale reduplicatie juist significant minder vaak worden gebruikt. De frequentie van de markeerder *mau* 'want' lijkt vrij stabiel te zijn. Aan de hand van grammaticalisatietheorieën en bevindingen uit andere onderzoeken wordt in dit hoofdstuk beargumenteerd dat *ada*, naast zijn functie van markeerder van progressief aspect, ook de functie van markeerder van tegenwoordige tijd/finietheid heeft aangenomen, een innovatie die kan worden verklaard door invloed van het Nederlands. In dit hoofdstuk wordt ook beargumenteerd dat taal-interne factoren, zoals vorm-betekenis correspondentie, frequentie en akoestische opvallendheid een belangrijke rol spelen in het behoud dan wel verlies van aspectuele vormen.

Hoofdstuk 5 concentreert zich op de uitdrukking van semantische gebeurtenissen van bezitoverdracht in het Ambonees Maleis in Nederland. In dit hoofdstuk worden de verschillende strategieën om bezitoverdracht uit te drukken in het Ambonees Maleis en het Nederlands besproken. In beide talen wordt zowel de 'Dubbele Object

(DO) constructie' (*Jan geeft Marie een boek*) als de 'Prepositionele Object (PO) constructie' (*Jan geeft een boek aan Marie*) gebruikt, echter met verschillende frequenties. In het Ambonees Maleis heeft het gebruik van PO altijd de voorkeur, terwijl in het Nederlands DO de voorkeur heeft in gesproken taal, en PO in elicities in specifieke contexten. Bovendien kunnen bezitsoverdracht in het Ambonees Maleis ook worden uitgedrukt door het gebruik van twee predicaten in één enkele zin (de 'twee predicaten' constructie). Een kwantitatieve analyse van de data laat zien dat erftaalsprekers de DO-constructie significant vaker gebruiken dan sprekers in het land van herkomst, terwijl ze de 'twee predicaten' constructie juist significant minder vaak gebruiken. In het hoofdstuk wordt beargumenteerd dat de oorzaak van deze frequentieveranderingen ligt in de invloed vanuit het Nederlands en in universele principes betreffende taalontwikkeling in het geval van gereduceerde input. Het hoofdstuk laat zien dat er ook kwalitatieve verschillen tussen de erftaalsprekers en sprekers in het land van herkomst van het Ambonees Maleis zijn: de verschillende voorzetsels die beide groepen gebruiken in de PO-constructie zijn een gevolg van de verschillen in hun sociale geschiedenis en in het type input waaraan de erftaalsprekers zijn blootgesteld (Tangsi Maleis).

Hoofdstuk 6 behandelt resultatieve constructies. In dit hoofdstuk worden de verschillende strategieën om resultatieve evenementen in het Ambonees Maleis en het Nederlands uit te drukken besproken. Hoewel in het Ambonees Maleis een voorkeur geldt voor seriële werkwoordsconstructies (SVCs) (*Zij breekt een stok wordt twee*), en in het Nederlands voor werkwoordspartikels (*Zij breekt een tak af*), komen in beide talen ook voorzetselconstituenten (*Zij breekt een stok in tweeën*) en bijvoeglijk naamwoordsconstituenten (*Zij slaat een vaas kapot*) voor. Erftaalsprekers gebruiken SVC's significant minder vaak dan sprekers in het land van herkomst, terwijl ze voorzetsels en adjectieven juist significant vaker gebruiken. Het feit dat de erftaalsprekers zich verwijderen van de voorkeuren van de sprekers in het land van herkomst en zich ontwikkelen in de richting van die van sprekers van het Nederlands toont duidelijk aan dat invloed vanuit het Nederlands de voornaamste oorsprong is van divergentie. In het hoofdstuk worden ook de veranderingen besproken in de keuze van specifieke voorzetsels die gebruikt worden. Erftaalsprekers van het Ambonees Maleis gebruiken voornamelijk de voorzetsels *ka* 'naar' en *dalang* '(binnen)in' (*ka dua* 'in tweeën'). De semantische uitbreiding in het gebruik van deze voorzetsels is een intern proces dat gemotiveerd wordt door universele principes, en versneld door contact met het Nederlands.

In Hoofdstuk 7 worden de bevindingen van de voorgaande hoofdstukken bijeengebracht om te kijken welke structurele en sociale factoren verantwoordelijk zijn voor de divergentie- en convergentiepatronen in de verschillende grammaticale domeinen van het Ambonees Maleis in Nederland. Eerst wordt gekeken naar structurele factoren. Door gebruik te maken van hiërarchische clusteranalyses wordt in dit hoofdstuk aangetoond dat de linguïstische kenmerken van het Ambonees Maleis in Nederland te verdelen zijn in twee clusters: de innovatieve ‘Nederlands-achtige’ kenmerken en de meer conservatieve ‘Maleis-achtige’ kenmerken. De Nederlands-achtige kenmerken zijn alle kenmerken die door erftaalsprekers gebruikt worden om de compatibiliteit tussen het Ambonees Maleis en het Nederlands te maximaliseren. Aan de hand van psycholinguïstische modellen van tweetalige taalverwerking wordt in dit hoofdstuk geïllustreerd hoe sprekers frequentiepatronen vanuit de ene taal naar de andere kopiëren met als doel om de twee systemen zo gelijk mogelijk te maken. In het tweede deel van dit hoofdstuk, gericht op sociale factoren, blijkt dat de factor *plaats waar de spreker woont* de beste voorspeller van linguïstische innovaties is. Sprekers die buiten een Molukse wijk wonen hebben een hoger aandeel Nederlands-achtige kenmerken dan sprekers die daarbinnen wonen. Bovendien wordt in dit hoofdstuk aangetoond dat de interactie tussen *leeftijd van aanvang van tweetaligheid* en *plaats waar de spreker woont* grote effecten kan hebben op het taalgebruik van erftaalsprekers.

Hoofdstuk 8 sluit dit proefschrift af door de onderzoeksvragen zoals gesteld in Hoofdstuk 1 te beantwoorden. Het antwoord op de eerste vraag (Verschilt het Ambonees Maleis in Nederland van de herkomstlandvariëteit?) is beslist ja. Het Ambonees Maleis in Nederland blijkt in alle vier de in dit proefschrift onderzochte grammaticale domeinen te verschillen van het Ambonees Maleis zoals gesproken in het land van herkomst. De verschillen tussen de twee variëteiten zijn zowel kwalitatief als kwantitatief, waarmee wordt aangetoond dat de veranderingen zijn doorgedrongen tot zowel individuele sprekers als tot de gehele gemeenschap van sprekers. Het antwoord op de tweede vraag (Is het Ambonees Maleis in Nederland aan het veranderen onder de invloed van het Nederlands?) is ook ja. Het Nederlands is de dominante taal van erftaalsprekers van het Ambonees Maleis en zodoende misschien wel de belangrijkste oorzaak van verandering. Deze verandering manifesteert zich op twee manieren: verandering in frequentie en grammaticalisatie door taalcontact. In het geval dat het Ambonees Maleis twee (of meer) even goede opties biedt, zullen erftaalsprekers de voorkeur geven aan die optie die ook bestaat in hun dominante taal, het Nederlands, waarmee ze de frequentie ervan verhogen.

Als een grammaticaal contrast verplicht uitgedrukt moet worden in het Nederlands, zullen erftaalsprekers de neiging hebben om dit ook openlijk uit te drukken in het Ambonees Maleis, door bestaande structuren te gebruiken en deze verder te grammaticaliseren in de erftaal. Het antwoord op de derde vraag (Wat zijn de factoren die achter deze ‘voortgaande’ verandering zitten?) is als volgt: invloed vanuit het Nederlands, kwalitatieve en kwantitatieve verschillen in input waaraan erftaalsprekers worden blootgesteld, en universele principes betreffende taalontwikkeling van een taal die in onbruik raakt, zoals simplificatie, conceptuele natuurlijkheid, en voorkeur voor bepaalde iconische constructies. Het is waarschijnlijk dat deze factoren met elkaar in wisselwerking zijn en een cumulatief effect hebben.

De hierboven genoemde factoren – wederzijdse beïnvloeding tussen talen, het type input, en universele principes – kunnen de veranderingen die zijn aangetoond in het Ambonees Maleis in Nederland verklaren. Sociale factoren, zoals de plaats waar de sprekers wonen en de leeftijd van aanvang van tweetaligheid, bepalen de mate waarin de verandering zichtbaar wordt binnen en tussen verschillende individuen. De plaats waar de spreker woont is de belangrijkste factor in het Ambonees Maleis in de Nederlandse gemeenschap, gevolgd door leeftijd van aanvang van tweetaligheid. Erftaalsprekers die buiten een Molukse wijk wonen hebben een groter aandeel innovatieve Nederlands-achtige kenmerken, terwijl sprekers die binnen een Molukse wijk wonen meer vasthouden aan conservatieve Maleis-achtige kenmerken.

Curriculum vitae

Francesca Romana Moro was born on September 7, 1985, in Rome, Italy. In 2008, she obtained her B.A. degree in Indonesian language and literature from the University of Naples “L’Orientale”. With the support of the Darmasiswa Indonesian Scholarship Program, she spent a year at the Udayana University in Denpasar, Bali. Afterwards, she continued in the Research Master program in Linguistics at Leiden University, where she graduated in February 2011. From June to October 2011, she worked as a Student Assistant at the Radboud University Nijmegen in the ERC Project “Traces of Contact” headed by prof. P. C. Muysken. In November 2011, she began her PhD research in the Centre for Language Studies (CLS) at the Radboud University Nijmegen. This dissertation is the result of her research.

Currently, she is employed as a postdoctoral fellow at the Leiden University Centre for Linguistics (LUCL). Her research is part of the VICI Project “Reconstructing the past through languages of the present: the Lesser Sunda Islands” funded by The Netherlands Organization for Scientific Research and headed by prof. M. A. F. Klamer.