UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Bilingualism and language contact

Aalberse, S.; Muysken, P.

DOI
10.1017/9781316831922.027
Publication date
2019
Document Version
Final published version
Published in
The Cambridge handbook of bilingualism

License
Article 25fa Dutch Copyright Act (https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment)
Link to publication

Citation for published version (APA):
Aalberse, S., & Muysken, P. (2019). Bilingualism and language contact. In A. De Houwer, & L. Ortega (Eds.), *The Cambridge handbook of bilingualism* (pp. 524-543). Cambridge University Press. https://doi.org/10.1017/9781316831922.027

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:26 Mar 2025

26

Bilingualism and Language Contact

Suzanne Aalberse and Pieter Muysken

1 Introduction

This chapter discusses bilingualism from the perspective of language contact studies. We could consider language contact the umbrella term for all situations where languages come into contact, either through interaction between groups of speakers who have diverse backgrounds or because elements of more than one language come together in an individual speaker. That latter case would refer to bilinguals. Although in principle all parts of bilingualism fall under language contact studies, we focus on work that is more explicitly placed in the domain of language contact, that is, the study of the consequences of bilingualism for language use and structure, rather than on individual bilingualism itself.

The term language contact is thus used here to refer to the effects that speakers' bilingual practices have on the languages they use. These effects may involve any part of the language system: phonetics, intonation, and phonology; the words themselves; morphology; syntax; semantics, discourse, and pragmatics. The effects may be temporary or more permanent, entrenched in the language system as it were. Thus, language contact is a rather vague term for a vast area, particularly since bilingual practices are extremely diverse and all aspects of language are involved.

While there has been a steady stream of more incidental observations about language contact in earlier periods, starting at least in the 19th century, Weinreich (1953) and Haugen (1953) gave the field a major impulse both conceptually and empirically. Weinreich had multilingualism in Switzerland (French, Italian, Swiss, German) as his empirical base; he discussed several types of bilinguals and different types of results of what he called interference (i.e., mutual influence of the languages) in their language production. Haugen focused on Norwegian as a heritage language in the Norwegian immigrant community in the United States and brought in historical considerations as well. He can be credited as the founder of the field of heritage language studies. He is also known for his

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

focus on the ecology of languages, the basic tenet of which is that the environment in which a language is spoken greatly influences its use and subsequent development.

It is impossible to even begin to summarize all the work here in the broad field of language contact studies, which has been approached from very different angles. These include psycholinguistics and neurolinguistics, sociolinguistics and anthropological linguistics, as well as historical and comparative linguistics. Winford (2003) offers a broad but detailed introduction to the field, with chapters providing both summaries and detailed commentaries on lexical and structural borrowing, social and structural aspects of code-switching, bilingual mixed languages, second language learning, pidgins, and creoles.

In this chapter, we limit ourselves rather drastically in two ways. First, we discuss just a few types of bilingual practices. They are four: code-switching and lexical borrowing; contact-induced language change; mixed languages, pidgins, and creoles; and heritage languages. Second, we limit ourselves to the lexicon and to syntax. The four fields we cover (and then only in a very superficial way) have received considerable scholarly attention, and the study of the lexicon and syntax has been very important for them.

We focus on the more system-oriented domains of bilingual practice, and on convergence in bilingual speech, that is, the process by which two languages become structurally more alike, while disregarding important fields such as bilingual acquisition or second language acquisition. Although the focus of the subdisciplines we discuss has been primarily scholarly, there are important societal implications of the research reported on here. One concerns issues of language purity as linked to ethnic identity: in many communities, the influence of other languages, perceived or real, is a matter of concern both to politicians and to ordinary citizens, and the study of language contact can help explain what is actually happening (see Fuller, Chapter 6, this volume, and Lo Bianco, Chapter 8, this volume). Another one concerns the language portrayed in educational materials for migrant communities in schools (see Bigelow & Collins, Chapter 2, this volume; and García & Tupas, Chapter 20, this volume) or adult education (see Simpson, Chapter 13, this volume) or groups of traditional heritage speakers. Is it the traditional varieties of language that are used here, or the evolved varieties bilingual speakers use in their daily lives?

2 Code-Switching and Lexical Borrowing

Since Poplack's (1980) study, code-switching (defined as the use of different languages in the same conversation) and lexical borrowing (that is, the bringing of a word from one language into another one) have been the focus of considerable attention across a wide range of linguistic subdisciplines. In her study, Poplack examined code-switching between Spanish

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

and English in the Puerto Rican community in New York using audio recordings of spontaneous conversations. The patterns of switching she found were relatively symmetrical, that is to say, speakers switched back and forth evenly between the two languages. The principal constraints on the switching that Poplack discovered were that switches generally occurred where the word order of the two languages was the same and did not occur within a word. An additional finding was that only highly proficient bilinguals switched frequently within a clause. There is an ongoing discussion on the status of asymmetrical code-switching in Poplack's model, specifically as regards switching in which there is a clear base language into which elements are inserted. Poplack argues that these cases are all instances of borrowing or sometimes nonce borrowing (when a word from another language is only introduced once), while other researchers assume there to be insertional code-switching (i.e., putting a fragment from another language into a sentence while code-switching).

Myers-Scotton (1993) based her Matrix Language Frame (MLF) model primarily on earlier work in which she recorded and analyzed Swahili–English code-switching in Kenya. Most of her data show a clear asymmetrical pattern in which one language, often Swahili, functions as the base or matrix language, and elements from another language, generally English, are inserted into this matrix language. The word order, grammatical structure, and functional elements such as articles and verb endings (which Myers-Scotton labels system morphemes) come from the matrix language. The grammatical model that Myers-Scotton developed is part of a broader pragmatic model in which a particular language functions as the unmarked (i.e., expected) code in specific conversations and the inserted elements come from a marked (i.e., unexpected) language. There has been extensive discussion regarding the possibility of capturing all cases of code-switching in asymmetrical models, as the strict version of Myers-Scotton's MLF model would require. Many researchers question the possibility of assigning a single matrix language to all cases in which code-switching occurs.

The work by Poplack and Myers-Scotton has been the basis for a series of intense debates on the nature of, and possible grammatical constraints on, code-switching within the clause. The term code-switching refers to the juxtaposition of elements that can be seen to belong to different languages, but sometimes code-mixing may seem to be the better term since, when code-switching occurs within a clause, the result appears to be a new mixed language, subject to specific constraints on mixing.

The controversy regarding models and conceptualizations involves several theoretical, grammatical frameworks, from Universal Grammar–based minimalism to cognitive grammar. Muysken (2000) attempted to summarize some of the findings in the debate in order to arrive at some sort of consensus on the types of code-switching

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

encountered. The basic distinction he introduced is between asymmetrical insertion, in which there is a clear basic or matrix language, and more symmetrical alternation, in which there is a shift or juxtaposition from one basic language to the other. A third type of code-switching he distinguished is congruent lexicalization, where the languages involved are so similar in structure or lexicon or both that there is neither insertion nor alternation, but rather a rapid back-and-forth between the two languages. While some researchers have recognized the validity of this three-way distinction, it has proven difficult to unambiguously classify all instances of code-switching in terms of these three types.

Several studies have tried to put the process of code-switching in a more dynamic perspective, focusing on changes across different generations. Based on Li Wei's fieldwork on cross-generational change in language choice in families in the Chinese Tyneside (Newcastle, United Kingdom) community, Milroy and Li Wei (1995) proposed a dynamic network model to explain cross-generational shifts in code-switching strategies. In a similar vein, Backus (1996) explored patterns of code-switching in the Turkish-Dutch immigrant community in the Netherlands. Backus shows that first-generation speakers who arrived as adults produce simple insertions of basic culture-specific content words, while the insertions are more complex in the group that arrived from Turkey as a child; the second and third generations show more alternation patterns with a clearer separation of the two languages. A typical insertional pattern from Backus's work (1998, p. 9) can be seen in example (1) (for explanation of the abbreviations in the glosses, see the Appendix):

(1) iki gün once işte *bioscoop*-a, *vragen* yap-tıydı-m
two day before interj cinema-DAT ask do-PLUPF-1SG
"but two days before I had asked her along to the movies"
(Dutch elements are in italics, Turkish elements are not)

All work on code-switching mentioned so far has taken a sociolinguistic perspective involving recordings in specific bilingual communities and taking variation and speaker characteristics into account. A few efforts have been made to interpret code-switching patterns from a more formal perspective, such as minimalism in generative grammar. A typical example is MacSwan's (1999) approach, which was originally based on Nahuatl–Spanish code-switching data from central Mexico but takes a much broader perspective. In the minimalist approach, the goal is not to posit grammatical constraints specific to code-switching (such as "do not switch between a pronoun and a verb") but rather to let all constraints on the process of code-switching inside the clause follow from universal properties of our grammatical competence. Undoubtedly, this is a valuable research strategy. The main problem with it so far is that it postulates a priori that the underlying categorical makeup of different languages is identical (e.g., gender features in one language correspond to gender

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

features in another language). This assumption is put in doubt in much typological research.

In recent years, there has been an increase in psycholinguistic and computational modeling work on code-switching, and it is likely this will increase as new methods and more ecologically valid experimental designs to elicit code-switching become available.

3 Contact Languages

The field of contact linguistics looks at contact languages, that is, languages that emerge out of the meeting of different languages in a bilingual population. The field is too vast to cover here in its entirety. The main topics under discussion in contact linguistics are the role of first and second language acquisition for language contact, the role of language typology and language universals, processes of language simplification, complexification, and grammaticalization (i.e., the emergence of new grammatical categories such as "tense" and "number"), and language change. It is difficult to point to a single debate in this area. An overview of the whole field is the introduction by Velupillai (2015).

Contact linguists distinguish between new languages that are spoken as second languages for communication between speakers of different language groups, known as pidgins, languages spoken as community languages that have emerged out of earlier contact, known as creoles, and new languages that contain words or morphemes that originate from different language sources, known as mixed languages. There is much debate about the precise definitions and delineation of these terms, however. One point of disagreement, for instance, is whether it is feasible to define creole languages in terms of their typological features, or whether we need to define them solely in terms of the circumstances of their genesis. In the former approach, one would say, for example, that creole languages usually have basic subject-verb-object word-order, have a consonant-vowel syllable structure, etcetera. By contrast, an approach focusing on origins makes no such claims but simply defines creoles as contact languages that became stabilized in their grammar due to acquisition by children and their codification as community languages (see Chevrot & Ghimenton, Chapter 25, this volume, for related discussion of the difficulties involved in distinguishing between "language" and "dialect").

A related question is whether a well-defined class of creole languages exists, or whether the very notion of creole fractures into several dimensions that may or may not be jointly present in individual languages. Whatever the answer, diverse types of language-mixing phenomena can be found, some of which are briefly illustrated here.

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

Muysken (1981) introduced the notion of relexification, defined as the whole replacement of the lexical shapes of words by those of another language (e.g., Spanish *salt-ar* "jump" becoming *jump-ar*, incorporating now an English word form). He based it on rather unusual data from the mixed language of a community in Ecuador often referred to as Media Lengua. In this language, the traditional Amerindian language Quichua (Ecuadorian Quechua) is spoken alongside Spanish. However, for several generations the in-group vernacular was a form of Quichua in which the phonological shape of all roots was replaced by a Spanish shape while most of the Quichua lexical semantics, grammar, morphology, and phonetics was maintained. All word roots are thus Spanish (although pronounced as in Quichua), but the suffixes are almost all Quichua. A typical example is (2) from Muysken's unpublished fieldwork data, where the underlined elements are of Spanish lexical origin – *no* "no," *undi* (< *(d)onde*) "where," *i-* (< *ir*) "go," and *kiri-* (< *querer*) "want":

(2) [[*no* [*undi-mun* *i-na*] *kiri-k*]-*kuna-lla*]
neg where-dir go-nmlz want-ag-pl-del
"those that do not want to go anywhere"

Notice that the sentence in (2) is a far from simple structure, containing a double embedding, following Quichua grammar. Mixed languages are not necessarily simple languages.

Bakker (1997) wrote a detailed description of another mixed language, Michif, which involves the Algonquian language Cree as well as French. Michif probably emerged in the 19th century during the fur trade era in eastern Canada, but is currently spoken in some communities further to the west in Canada and in adjacent parts of the United States. In contrast to Media Lengua, where mixing mostly involves the phonological part of the lexical entry, in Michif, the basic split is between the verbal and the clausal systems, which are Cree in origin, and the nominal system, which is French. As shown in example (3) from Bakker (1997, p. 6), sentences are thus split between clearly French (nominal) parts (in italics here) and clearly Cree parts:

(3) êkwa pâstin-am *sa* *bouche* ôhi *le* *loup* ê-wî-otin-ât
and open-he.it his.F mouth this-OBV the.M wolf COMP-want-take-he.him
"and he opened his mouth and the wolf wanted to take him"

Every word in Michif is either Cree or French, and the full noun phrase is ordinarily French in origin. The Michif pattern is reminiscent of situations of intra-sentential insertional code-switching in which entire French noun phrases may be inserted, but the origin of Michif as the result of bilingual practices and insertional code-switching is controversial.

Thomason's (1997) book describes a number of the most important contact languages, including Hiri Motu from Papua New Guinea, Pidgin Delaware from the East Coast of the United States, Ndyuka-Trio pidgin

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

from Surinam, pidgin and creole languages with Arabic lexicons, Sango from the Central African Republic, forms of Swahili that have undergone restructuring as they were learned by populations of second language learners, Michif, Media Lengua, Callahuaya from Bolivia, Medniy Aleut from Siberia, and Ma'a (Mbugu) from Tanzania. Thomason presents many different typological patterns found among these languages, as well as differences in the processes that led to their genesis.

Lefebvre (1998) took up the notion of relexification introduced for Media Lengua and tried to account for the genesis of Haitian Creole. She postulated that the primary mechanism involved in its emergence was the relexification of Fongbe grammatical structures with French word forms. Speakers of the West African Fongbe language used relexification as a type of special early second language learning strategy in unusual circumstances. Since there is no inflectional morphology in Fongbe, relexification in this case would lead to structures that only involve French word material. However, the underlying grammatical organization would be Fongbe. An example is (4), from Lefebvre (1998, p. 197):

(4)	Mari	te	*prepare*	pat	Haitian Creole
	Mari	kò	Dà	wÒ	Fongbe
	Mary	ant	prepare	dough	

"Mary prepared dough"
(in the Haitian Creole sentence, the French word is italicized)

Although many grammatical patterns in Haitian Creole are similar to Fongbe patterns, not all researchers agree that the primary mechanism leading to the genesis of Haitian Creole and related languages was indeed relexification.

The collective volume edited by Meakins and O'Shannessy (2016) contains a synthesis of the exciting work that has been done on the way speakers of Aboriginal communities in Australia have reacted to colonization and the spread of English in the continent. Contact between Aboriginal languages and the colonial language of English has led to some very interesting contact languages in Australia, the best known of which are Gurindji Kriol and Light Warlpiri. These languages share the main division that Bakker (1997) encountered for Michif. The nominal frame (including case markers and derivational morphology) is from one language, in this case from one of the traditional languages, Gurindji or Warlpiri (in contrast with Michif, where the nominal frame was from the "new" language, French) and the verbal tense-mood-aspect system from another language, in this case English or Kriol. A third mixed language from Australia is New Tiwi, with again the main division between nominal and verbal frames, but in this case New Tiwi sides with Michif. However, these mixed languages are only part of the story of language contact in Australia, which involves mutual influences between the Aboriginal languages themselves, the emergence of contact languages such as Kriol,

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

and the interaction between English and Kriol and various Aboriginal languages.

Speaking broadly, the overviews of mixed languages in the literature show a considerable range of mixing phenomena in different language pairs, often involving opposition or hierarchy. Different oppositions include those between (1) nominal versus verbal frames, as in the case of Michif, where nominal material is from French, and (2) lexical versus grammatical frames, as in the case of Media Lengua, where lexical material is from Spanish. Mixed languages may emerge as languages of ethnically mixed communities, in communities undergoing rapid cultural changes, or sometimes as languages of interethnic communication. Several factors having to do with power and prestige of the languages involved, the degrees of speakers' proficiency in the different languages, and structural properties of the languages involved determine which of these oppositions comes into play in a specific mixed language.

4 Contact-Induced Language Change

The role of language contact in language change has always been controversial and continues to be so. One of the main reasons is that different time frames produce different results: long-term historical developments often seem to lead to more profound changes than short-term contacts in contemporary bilingual communities. Contact-induced language change is by no means omnipresent, it seems, at least on the structural level. Lexical change, as in borrowing, does seem to be virtually always there.

The controversy about language change involves three partly related questions asked and discussed in the sections that follow.

4.1 Can Language Change Be the Result of Contact with Another Language?

For this question, the consensus among most researchers is yes. There are well-documented cases of contact-induced language change. However, for many individual cases, this remains a controversial issue. While in earlier periods, resistance to the idea that contact-induced change existed came mostly from structuralist and generative quarters, surprisingly, current criticisms of the notion of contact-induced change come mostly from within the variationist paradigm, which posits that variation within spoken language is central. Numerous examples can be given, such as Poplack and Tagliamonte (2001). The reasoning in this paradigm is generally the following. In the variety that has undergone the change, the variable constraints (i.e., the probabilities for a particular change to occur in a specific grammatical context) on the phenomenon in question resemble those in an ancestral variety. Hence, attributing them to an external factor

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

such as a different language is not necessary and possibly counterintuitive. A typical example was Labov's original demonstration (1969) that the variable constraints on the presence of the copula (*she strong* versus *she is strong*) in African American Vernacular English (AAVE) resembled those on the contraction of the copula in Standard English (*she's strong* versus *she is strong*). This resemblance made postulating the influence of an earlier creole in the genesis of AAVE less attractive as an explanation. This type of argumentation is particularly strong if the variable constraints involved do not follow from universal properties of human processing (an example would be that the sound *s* is often deleted after a *t* or *k*, since pronouncing the sequence of these consonants is not easy), in which case the resemblance could be accidental.

4.2 What Conditions the Processes of Contact-Induced Language Change?

The fuller and more specific version of this question is: are the processes of contact-induced language change, where present, mostly conditioned by structural features of the languages involved, or by sociolinguistic conditions? Thomason and Kaufman (1988) in particular have addressed this issue. They argue that different types of language contact processes or scenarios lead to different outcomes, and they introduced the principal distinction between maintenance and shift contexts. In maintenance contexts, a socially subordinate (minority) language is maintained, but undergoes (predominantly lexical) influence from a dominant (majority) language (for discussion of power relations between languages on a societal level, see Fuller, Chapter 6, this volume, and Lo Bianco, Chapter 8, this volume). In shift contexts, languages are abandoned in favor of the dominant language and leave primarily non-lexical traces in the resulting variety (e.g., Italian intonation patterns in Argentinian Spanish). This type of influence is labeled substrate influence. The different contexts are subject to different hierarchies of what elements are taken over from the other language. Thus, in maintenance contexts, there is a hierarchy of which elements are most easily borrowed, starting with nouns, and ending with elements like verb endings.

Van Coetsem (1988) introduced the fundamental distinction between recipient language agentivity (i.e., cases where the language undergoing change is strongest, as with borrowing) and source language agentivity (i.e., where the source language is strongest, which van Coetsem labels interference). This roughly corresponds to Thomason and Kaufman's (1988) distinction formulated in sociolinguistic terms between maintenance-related and shift-related processes. The central idea in van Coetsem's work is that in borrowing processes, the principles operating in the recipient language determine the shape and structure of a borrowed item (e.g., a word), while in interference, the principles operating in the

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

source language determine these properties. An example of interference is the case of substrate influence in the Andes, where semantic and phonetic patterns of the substrate language, Quechua, are transferred to the language to which the population of speakers has shifted, viz., to their version of Andean Spanish. When Quechua speakers adopt a word from Spanish (borrowing), the word is adapted to Quechua phonology and morphology. In van Coetsem's (1988) model, this is a case of recipient language agentivity. Notice that the notion of agency here is couched in terms of the languages involved, rather than the speakers of those languages, but it could be formulated in terms of speaker agency as well.

4.3 Are There Any Restrictions on Which Parts of a Language Can Be Influenced by Another?

The question here, stated more precisely, is: if there can be contact-induced change, are there perhaps certain parts or components of language that cannot be influenced by another language? Several scholars have addressed this issue. Most approaches to language contact make a general distinction between lexical and structural transfer. This distinction is based on the type of linguistic material that is transferred from one language into another (Matras, 2009; Thomason & Kaufman, 1988; Winford, 2003). Matras (2009) distinguishes matter replication, which involves the transfer of a word's phonological form, from pattern replication, which involves "the organization, distribution and mapping of the grammatical or the semantic meaning while the form itself is not borrowed" (p. 146). Matter replication is often lexical, such as the borrowing of the English word *awkward* into Dutch, but it can also include the borrowing of stock phrases such as the use of sentences like *forget it* from English into another language. An example of pattern replication is the use of the combination of words as in *bij de weg* ("by the way"), which is literally translated from English into Dutch. Although all the words are Dutch, the combination of the words to form the meaning "incidentally" is inspired by English (the literal Dutch meaning here would normally be "near the road").

Johanson (2002) studied developments in the Turkic languages and their interactions with surrounding languages, particularly with those belonging to the group of Indo-Iranian languages. His distinction between transfer of form and transfer of structure resembles the distinction put forward by Matras (2009), but rather than speaking of transfer or replication, Johanson speaks of copying and copies. One characteristic of copies is that they are like the original, but almost never exactly the same. The main distinction Johanson makes is the one between global and selective copies. Global copies are code-switches where all components from one language are copied into an utterance in the other language (thus involving both form and meaning). Selective copies import

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

parts of one language into the other, but not the complete form-meaning package. Johanson then distinguishes four types of selective copies, namely (1) material, (2) semantic, (3) combinatorial, and (4) frequential copies.

A material copy means that the external shape of a language element, the word form, is copied, but not necessarily with the same meaning or combinatorial properties as in the original language. For example, the word *beamer* is borrowed from English into Dutch and German, but with a very different meaning than its use in English: it is only used to refer to video projectors, whereas in English, the word is mainly used to refer to a kind of car. Semantic copies refer to the use of words in a set of meanings from another language. For example, in Russian, *xorošo* means "well," as in *he did well on his exam.* It cannot be used in contexts like *this was . . . well . . . embarrassing* in Russian, but this use is possible in American Russian as the result of a semantic copy of the English word *well*. The meanings of *well* have been copied into American Russian. Combinatorial copies refer to the incorporation of collocations in one language into the other. The Spanish word *rascacielos*, meaning "skyscraper," is an example of a combinatorial copy, where the combination of the verb *rasca*, "scrape(s)," and the noun *cielos*, "skies," is copied into Spanish, modeled on the English form *skyscraper*. Last, a frequential copy refers to the process of a change in frequency of a certain structure as the result of language contact. In Section 5, we discuss the decrease of the use of subjunctive in heritage Spanish, making the language more like English. We also discuss the increase of demonstratives in heritage Chinese in the Netherlands. Both changes in frequency are examples of frequential copies, which make the number of uses more like the dominant language.

The Australian linguist Ross, who worked on many languages of the Pacific, both in Australia and in New Guinea, coined the concept of metatypy (Ross, 1996). This refers to a change in grammatical typological profile under the influence of a dominant language, hence the term metatypy, typological metamorphosis. A language that undergoes metatypy is usually a smaller community language used primarily within the community itself, while many speakers are also speakers of a language of wider use, which often has more prestige. Some members of the smaller community use the second language more frequently than their community language, and this leads to changes both in the semantic and pragmatic organization of their first language and in its grammar, at the sentence, clause, phrase, and word levels. The case that triggered Ross's interest was the Austronesian language Takia in New Guinea, which had become more Papuan-like under the influence of the more widely spoken Papuan language Waskia, adopting some of the word order features and semantic patterns of that language. The influence of the other language must be substantial for the outcome to be labeled metatypy, but no exact criteria

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

for calling this type of change metatypy have been proposed (and may indeed not be possible).

Thus, historical linguists like Ross are willing to assume profound contact-induced language change, often involving long-term processes of adaptation. In contrast, sociolinguists have not found this type of change in communities with a shorter history of contact. Torres Cacoullos and Travis (2015) worked on New Mexican Spanish, including code-switched speech in the New Mexican Spanish–English bilingual community. They worked on processes of grammaticalization and the linguistic outcomes of language contact in a variationist framework. In contrast with the type of results Ross reported for New Guinea, they have argued that it is very difficult to find convincing evidence for structural influence of English on New Mexican Spanish, even though English is also the dominant language in this case.

5 Heritage Languages from a Contact Perspective

This section discusses heritage languages from a contact perspective. We use the term heritage language as defined by the editors, that is, to denote a language that someone learns as a child in the family and within the additional context of a community of speakers (e.g., neighborhood, places of worship), often in competition with the dominant language of the majority society, in which the heritage language speaker may end up being more comfortable and/or proficient. Note that all heritage speakers we discuss are proficient (to some extent) both in the home/community language as well as in the official language(s) of the country and thus are always bilinguals. Heritage languages have already been discussed from an acquisition perspective in other chapters in this volume (see in particular Singleton & Pfenninger, Chapter 4, and Keijzer & de Bot, Chapter 14). This chapter focuses on heritage languages from a contact perspective.

How would one recognize a language contact framework as opposed to, for example, an acquisition framework? Although the boundaries between the two approaches are not fixed, we identify some factors that play a role in differentiating the two discussed in what follows. We illustrate our points with examples from two edited volumes on Germanic heritage languages in North America that are closely connected to the contact perspective (Johannessen & Salmons, 2015, and Page & Putnam, 2015).

One factor that is characteristic of the contact approach to heritage languages is time depth. Language contact studies can involve very long periods of language contact, even as long as a few thousand years. Heritage language studies from a contact perspective could (but do not have to) involve the study of language use in multiple generations of speakers across time. The volume edited by Page and Putnam (2015), for example,

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

includes a study on Volga German in Ellis County, Kansas, in the United States. Volga German is the kind of German spoken by German Catholics who settled along the Volga River in Russia in 1760 and who bought land from the Kansas Pacific Railroad in 1875 to resettle in Ellis and neighboring counties in west central Kansas. During their periods both in Russia and in Kansas, the community kept speaking German. Another study included in the same volume compares heritage Norwegian at various points in time, namely in 1940, 1990, and 2010. Effects of language contact over such longer periods of time and comparisons of the contact languages in different periods are typical of the contact perspective.

The time depth issue has an effect on the type of linguistic domains that are studied. Whereas acquisition studies tend to focus more on what is not acquired or lost, language contact studies could also look at what additions language contact has brought to the new variant. For example, Volga German as described earlier has developed a new prepositional case. The volume edited by Johannessen and Salmons (2015) includes a study on the use of English modals that connects this use with hierarchies that can be borrowed from contact languages, showing that transfer of modals is common in contact situations. The same volume includes a study showing that heritage German has lost rounded vowels and that connects loss of rounding to more general patterns of change in historical linguistics. Yet another study shows that Pennsylvania Dutch, a variant of German in the United States, has extended the use of the progressive. All these forms of variation are described from a language change perspective, connecting them to other forms of language change, rather than from the perspective of the lack of acquisition.

The time depth issue also influences the methods applied. Information on the language use of eight generations ago, for example, cannot be based on experiments, since the speakers are no longer alive. Data might come from diaries, grammatical descriptions, recordings of speech from earlier time periods, or letters. Speakers who are still alive and know how to speak the language are usually quite old. The implications of working with older speakers are also described in Keijzer and de Bot (Chapter 14, this volume).

Another possible difference between the contact perspective and the acquisition perspective is attention to variation and stylization. Rather than focusing on learnability issues, authors with a contact perspective look at the strategic use of different languages. One could look at the role of identity marking in the selection of a particular language and one could study why and in what context the other language is used or why languages are mixed (see De Houwer, Chapter 17, this volume, for discussion of some of these issues). In Section 4, we discussed the use of literal translations from one language into the other. These are known as calques. Speakers can use calques subconsciously or more consciously as a creative form of language play that can mark group identity and creativity. For example, Dutch writer and comedian Arjen Lubach uses the Dutch

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

expression *wees daar of wees vierkant* ("be there or be square"), which is a literal translation from English, but the word *square* is translated into just one of its English meanings (not the one intended in "be there or be square"). The Dutch translation does not rhyme and sounds odd (comparable to "be there or be rectangular"), but is funny to English–Dutch bilingual speakers who realize it is a direct translation from English. Purposely deviating from the norm to create an ideal self is also discussed by Singleton and Pfenninger (Chapter 4, this volume) in connection with research by Nance (2015) showing how "new speakers" of Scottish Gaelic do not make the traditional velarized/palatalized rhotic distinction, because the preferred ideal self is oriented toward a new speaker mode and the traditional native speaker target of older speakers is considered inauthentic.

The time depth issue also affects how linguistic variation is regarded. Studies from a contact perspective will consider variations in forms as the result of language change and as natural reactions to the changing linguistic ecology, rather than consider them errors. Relevant to the linguistic ecology is the fact that in many heritage communities, speakers with different dialect backgrounds come together (see Chevrot & Ghimenton, Chapter 25, this volume, for discussion of dialectal variation). This yields the formation of koines, that is, new variants of a language that arise when mutually intelligible variants of one language come together. Characteristic of koine formation is that features (in the lexicon, morphology, etc.) that are specific to one dialect are lost or replaced. The volume edited by Johannessen and Salmons (2015) includes a study on the effects of koineization in heritage Swedish and argues that the use of some English words in American Swedish is motivated by the need to avoid dialect-specific lexical forms.

Language contact studies focus on contact. This implies that both languages under consideration are studied. Whereas second and heritage language acquisition studies may just focus on what parts of language Y are difficult for speakers independently of their other language(s), language contact studies always look at both languages. Looking at both languages is not unique to the contact perspective, however: many acquisition studies do it too, but a contact perspective must involve both languages, whereas this is not obligatory for acquisition studies (although see Treffers-Daller, Chapter 15, this volume, and Hammer & Edmonds, Chapter 19, this volume, for arguments that bilingual abilities must always be measured and tested bilingually). Because language contact can imply the birth of a new contact language, the community perspective is also more popular. How do multiple speakers with the same bilingual background make use of elements of different languages?

Having discussed relevant characteristics of the contact perspective to heritage languages, we now move to two more detailed case studies of

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

contact-induced change in heritage speakers. A pioneer in this domain is Silva-Corvalán (e.g., 1994), who has published extensively on heritage languages from a contact perspective, basing herself on Spanish in Los Angeles as a form of Spanish–English language contact. An important notion she expressed in her work is the notion of indirect change or change in frequency. The fact that she stressed the notion of change is typical of a contact perspective. The idea behind the notion of indirect change is that language contact does not bring about totally new structures that were alien to the language before, but that it can cause an increase or a decrease in the use of a certain structure, just as discussed in Section 4 with regard to the notion of frequential copying.

Silva-Corvalán (1994, p. 133) argued that the existence of superficially parallel structures in the languages in contact makes those structures of the language permeable to change. Here we discuss two types of change in frequency from her work that are due to superficial overlap between structures. The first type concerns the extension of the discourse-pragmatic functions of an element to parallel the functions of the translational equivalent in the other language. For example, English can use the proximal demonstratives *this* and *these* to introduce indefinite specific referents in order to foreground that referent, as in *yesterday I met this guy* . . . Non-heritage Spanish does not use deictics to refer to new entities in the discourse position in subject position, but example (5) from Silva-Corvalán (1994, p. 136) shows how some heritage speakers do make use of deictics this way:

(5) pero esta mula tenía la la idea que si no comía
but this mule holds the the idea that if not eat
algo en la casa . . .
something in the house
"but this mule had the idea that if he did not eat something at the house . . . "

The use of *esta* ("this") in (5) is grammatical in non-heritage Spanish too, but in non-heritage Spanish it would only be used for already mentioned referents, while in heritage Spanish the possibilities of its use are extended to refer to new referents in the discourse.

A second type of indirect change that Silva-Corvalán (1994) mentions is the increased use of a structure that parallels a structure in the other language. For example, both non-heritage Spanish and English allow subordinate clauses with an overt complementizer (such as in *the professor thinks that she* . . .) as well as with null complementizers (such as *the professor thinks Ø she* . . .). Null complementizers are much more frequent in English than in Spanish. Silva-Corvalán shows that the earlier Spanish speakers come into contact with English, the more likely they are to use null complementizers in a conversation. She compared the use of null complementizers in her corpus of audio-recorded conversations across three

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

groups of speakers. The first group, speakers who arrived in the United States after the age of 12, showed 4 participants out of 13 using null complementizers. The second group, sequential bilinguals, had 6 out of 11 speakers using null complementizers, and the third group of speakers, simultaneous bilinguals, had 9 out of 11 speakers who used null complementizers. While the use of null complementizers increased per group, no speaker used null complementizers in relative clauses like *thank you for the present that you gave me*. Although null complementizers are frequently used in this context in English, its use is ungrammatical in Spanish. The absence of null complementizers in relative clauses illustrates the role of superficial parallel structure in permeability to change. Since the use of null complementizers is ungrammatical in Spanish, this context is not permeable to change.

Now that we have looked at two forms of indirect change in Silva-Corvalán's work, we move to a very special kind of shift in frequency, namely a shift in frequency that can imply the incipient birth of a new grammatical category. The contact-induced rise of new grammatical categories is closely connected to the study of linguistic areas. Linguistic areas are geographical areas where genetically unrelated or distantly related languages are spoken that have become structurally alike due to language contact. Note that in linguistic areas languages become structurally more similar, without speakers thinking of these variants as mixed variants, whereas in mixed languages the role of both languages is clear, because matter of both languages is used (e.g., words and morphology from both languages). Convergence is connected only to pattern replication (see Section 4.3), which is more subconscious and therefore less noticeable to speakers. Convergence is the process whereby languages become structurally more alike. The process of convergence can lead to the emergence of a linguistic area.

A form of convergence that has received quite some attention in the study of linguistic areas is the contact-induced rise of definite markers such as articles. Linguistic area studies suggest that a language without articles such as *the* can develop them under influence of a language with articles, thereby making one of the two languages structurally more similar. Often articles derive historically from demonstrative pronouns. For example, the English article *the* is derived from the distal demonstrative pronoun *that*. While demonstrative pronouns indicate deictic and definiteness information, articles indicate definiteness only. The rise of articles tends to go hand in hand with encoding definiteness obligatorily. For example, languages without articles can say *I read book* but languages with articles have to encode information about definiteness, that is, they can state *I read a book* or *I read the book*, whereas bare nouns are only allowed in very restricted contexts.

Aalberse, Zou, and Andringa (2017) explored whether in a Mandarin–Dutch contact situation there was any evidence of increased

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

demonstrative use in Mandarin, possibly as a result of the contact with Dutch, a language that, unlike Mandarin, has articles (demonstratives are part of Mandarin). In order to test this, they compared different speakers of Mandarin with each other: first- and second-generation Chinese immigrants in the Netherlands, where Dutch is the dominant language, and nonimmigrant Mandarin speakers in China. The immigrants were all in contact with Dutch, and Mandarin was a heritage language for them. The second-generation immigrants used Dutch more frequently than Mandarin. The first-generation immigrants varied considerably in their use of Dutch.

Based on a narrative elicitation task in Mandarin Chinese, Aalberse et al. (2017) found that both first- and second-generation speakers in the Netherlands used more demonstratives than Mandarin speakers in China when they were referring to persons or objects that had already been mentioned. The use of bare nouns, which is grammatical in Mandarin but ungrammatical in Dutch, decreased. One could interpret the decrease in the use of bare nouns as a form of convergence with Dutch.

In addition, Aalberse et al. (2017) found that the second-generation immigrants showed a preference for the Mandarin distal demonstrative *nage* ("that") when referring to already mentioned referents or objects. The distal demonstrative in Dutch is most closely connected to the definite article in Dutch both in meaning and contexts of use. Aalberse et al. therefore suggested that second-generation speakers make an extra step in the convergence process by not just avoiding constructions that are ungrammatical in Dutch (bare nouns) but in addition by selecting a translational equivalent *nage* ("that") which is closest to the Dutch context.

In addition to quantitative differences Aalberse et al. (2017) analyzed qualitative uses of demonstratives in immigrant speakers. They described two contexts where immigrants used demonstratives and nonimmigrant speakers did not use them. One was the use of demonstratives for multiple arguments in one utterance as shown in example (6), from Aalberse et al. (p. 41), and the other was the use of demonstratives in reference maintenance contexts as shown in example (7), from Aalberse et al.'s unpublished data:

(6)

他	就	把	那	个	石	头放	在	那个	凳子上,
he	then	BA	DEM	CL	rock	put	on	DEM CL	chair,
另外	那个		朋友	就	站上去。				
other	DEM CL		friend	then	stand on				

"he then puts the rock on the chair, and the other kid stands on top"
(This refers to a scene with two children where a child is trying to reach something in a tree. To make himself taller, he stands on a rock that another child had placed on a chair.)

(7)

他们	想	用	那个	球	扔	到
they	want to	use	DEM-CL	ball	throw	at

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

那	个	衣服	上,	希望	那	个	衣服	会	掉下来。
DEM	CL	clothes	up,	hope	DEM	CL	clothes	will	fall down

"they want to throw the ball at the clothes, hoping that the clothes will fall off"

In contexts like (7), the referent ("clothes") is a maintained argument in the second half of the sentence. Nonimmigrant speakers tend to use bare forms in this case, whereas Mandarin speakers in the Netherlands use demonstratives. Note that the use of demonstratives is grammatical, but feels redundant to nonimmigrant speakers.

A first step in the potential use of articles in Mandarin in the Netherlands could be that heritage speakers use more demonstrative markers than non-heritage speakers. Such an increase in frequency relates to the notions of frequential copying (Johanson, 2002) and indirect change (Silva-Corvalán, 1994). The characteristic that sets this type of frequency change apart from changes in frequency discussed in relation to the increased use of null complementizers in Spanish is that in the long run this change in frequency might end up causing a change in the grammatical function of the element: the change in frequency might be an indication that the demonstrative is becoming an article. The observation that second-generation heritage speakers in the Aalberse et al. (2017) study preferred the distal demonstrative makes the structural difference between Dutch and Mandarin smaller and can thus be seen as the next step in the convergence process.

6 Conclusion

In this chapter we have sketched concepts and results from language contact studies in four different domains: code-switching and lexical borrowing; contact-induced language change; mixed languages, pidgins, and creoles; and heritage languages. However, we have mostly limited ourselves to the lexicon and syntax.

The systematic study of phonetics, intonation, and phonology is just beginning in contact linguistics, as is the study of semantics, discourse, and pragmatics. Useful starting points for these domains are studies by Queen (2012), Nance (2015), and Yao and Chang (2016) for phonology, and by Babel (2014) and Andersen, Furiassi, and Mišić Ilić (2017) for semantics and pragmatics. The volume on Germanic heritage languages in North America (Johannessen & Salmons, 2015) also includes chapters on these less researched topics and thus forms a good starting point for these domains as well.

One crucial point in our overview was that the outcome of language contact varies per situation. This notion of variable outcomes is captured in Haugen's (1953) notion of language ecology, which assumes that the

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

setting in which languages are spoken affects the type of changes we observe in them. The same holds for language development in bilingual situations. Whether and how people code-switch, whether creoles show influence of speakers' L1 or of universal principles, whether and in which way languages are mixed and how heritage languages change – the answers here will differ given the specific context and given the specific set of languages involved.

A crucial distinction introduced in Section 4 is the opposition between what Matras (2009) has called matter versus pattern replication. Matter replication, the use of word forms from one language into the other such as in the case of borrowing words like *email* or complete phrases like *the American dream* from English into another language is a very common phenomenon. The specific identity of the donor language is usually also clear. The status of pattern replication is more debatable since the issue to what extent languages may undergo structural influence from each other through bilingual usage is unresolved. A first question that needs to be asked when one observes change in language contact situations is whether a change could have also taken place without language contact. If language contact does seem a probable trigger for change, the next question is whether that change is a general effect of bilingualism or an effect of crosslinguistic influence.

Although some issues remain unresolved, we would like to underline that a necessary condition for structural (as opposed to lexical) influence is sustained bilingual interaction in the daily lives of speakers, preferably over a longer period of time. Sustained daily bilingual interaction does not necessarily lead to structural influence, as far as we know, but the conditions under which potential changes are realized are not yet clear. There are some more general observations to be made, which are summarized in the following.

Some components of the grammar of a language may undergo change more rapidly than others, but we do not know exactly which components these are. Likely candidates for sensitivity to change may be the tense-mood-aspect system of a language as evidenced by the hyperextension of progressives in Pennsylvania German as discussed in Section 5. Also those components of language directly involved in interpersonal interaction such as emphatic marking, politeness marking, and the encoding of information structure may be more sensitive to change, although these may be part of pragmatics rather than grammar proper. Research has also shown that some words are easier to borrow than others. Extra-clausal elements like discourse markers and content words like nouns, for example, are more likely to be borrowed than prepositions or determiners.

As far as mixed languages goes, the types of hierarchies noted in Section 3 as regards the division of labor between the two languages generally hold. Furthermore, bilingual practices such as the use of mixed languages and code-mixing cannot be studied without attention to all

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press

languages involved. In fact, the role and the effect of *all* languages involved should be studied in any contact situation.

In language change involving language contact there is a cross-generational scaffolding effect, leading to more contact-induced change over a longer period. However, some speech communities appear to show much less change than others. The reasons for this are not understood so far, but possibly they involve communities' notions of purism, whether these are below or above levels of awareness.

The challenge lies in making the general observations here more concrete. One promising avenue for future research is to bring together insights from experimental language processing studies with formal and computational approaches and observational paradigms. An example of innovative research is the study on code-switching by Goldrick, Putnam, and Schwartz (2016), and the fact that it had commentaries from different research communities points to the way forward. Only interdisciplinary approaches will be able to increase our understanding of the connections between individual bilingualism and the very shape of the languages they use.

Acknowledgments

Aalberse's work on heritage languages was supported by the Traces of Contact project (ERC #230310). Muysken's work on this chapter was supported by this same project and by the NWO Language in Interaction Consortium.

Appendix: Abbreviations and conventions used in examples

1SG	first person singular	DIR	directional
AG	agentive	F	feminine
ANT	anterior	M	masculine
BA	preposition or light verb preceding affected objects	NEG	negation
		NMLZ	nominalizer
		OBV	obviation marker
CL	classifier	PL	plural
COMP	complementizer	PLUPF	plusquamperfect
DAT	dative	X.Y	original form corresponds to combination of two glosses
DEL	delimitative		
DEM	demonstrative		

https://doi.org/10.1017/9781316831922.027 Published online by Cambridge University Press