

CODE-SWITCHING BEHAVIOR AS A STRATEGY FOR MAYA-MAM LINGUISTIC REVITALIZATION

Wesley M. Collins

Abstract

Since 1991, Fishman has carved out a “new” area of focus for research and linguistic activism—the Reversal of Language Shift (RLS)—within the general field of the Sociology of Language. In this article, I discuss a strategy of RLS employed by educated speakers of Maya-Mam, an endangered language of Guatemala. Less-educated Mam routinely code-switch to Spanish, while educated speakers categorically do not. Communication Accommodation Theory (Giles & Powesland 1975) offers a framework for accounting for this distinctive behavior through consideration of convergence and divergence strategies aimed at constructing positive social identities (Tajfel 1974). I briefly discuss this code-switching behavior, and compare people’s opinions about it as a positive or negative communication accommodation. I suggest that the initiative of Mam teachers in “purifying the language” is supportive of their overall goal of RLS and Mam revitalization.

1 Introduction

Over half of the world’s 6,500+ languages are spoken only by adults who are not passing their native language on to their children (Krauss 1992). Aside from these *mori-bund* languages, an additional 40% are considered *endangered* vis-à-vis their speakers’ socio-economic, educational, and geographic proximity to speakers of major languages

WESLEY M. COLLINS. Code-switching behavior as a strategy for Maya-Mam linguistic revitalization. *OSUWPL* 57, Summer 2003, 1–39.

Copyright © 2003 The Ohio State University

like English, Spanish, Mandarin, or other regional or area trade languages. Nettle and Romaine (2000) report that as few as 600 languages around the world are considered “safe”.

Scholars from Sapir (1931) to Hale (1992, 1998) have argued that the loss of such minority languages through social assimilation to the encroaching majority groups or by the physical death of minority language speakers deprives the world of a number of treasures, both academic and cultural. While in some sense inevitable and perhaps even “economical” in a Darwinian sense of “only the strong survive”, the demise of minority languages and cultures is indeed considered “loss” by many scholars.¹ Krauss (1992) argues that in the same way that ecological devastation deprives the world of important biological diversity and the products derived from it, so the devastation of linguistic and cultural diversity deprives the academic world of treasures we never knew we had. Hale realizes that his complaint is self-serving, but Crawford (1995) discusses different issues of language death—social justice and the right of minority peoples to their own languages and cultural distinctives—as being just as important as the academic issues.

These issues of language *shift*, *decay*, and *death* fall within the purview of the sociology of language along with issues of language maintenance, multilingualism, language planning, and bilingual education, where the internal aspects of language are not so critically in focus as is the language itself as an entity or cultural object within the larger context of society as a whole.

Since Fishman (1991), the sociology of language, and particularly the field of language maintenance and shift, has been expanded to include a new sub-discipline, the reversal of language shift (RLS), which is conceived to be a purposeful, operational response to the looming demise of minority languages around the world.² Within RLS, scholars and speakers of endangered languages have studied language loss and strategies that have proved effective in helping to restrain it. Of these strategies, those which are adopted and promoted by the speakers themselves are the most likely to have long-term success (Lastra 2001).

In this study I look at code-switching behavior among the Maya-Mam as an indicator of social identity (Tajfel 1974). I discuss Mam social identity, in turn, as a crucial element in the group’s (or a sub-group’s) decision to converge toward or diverge from (Giles & Coupland 1991) the majority Spanish language and culture. I show that less-educated Mam use Spanish-Mam code-switching as a strategy for convergence toward the majority language and culture, while the more formally educated teachers avoid code-switching altogether. I discuss both this convergence (*boundary leveling* in Woolard’s (1988) words) and divergence (*boundary maintenance*) in terms of ideology stemming from the growing realization of the socio-political oppression of the Maya by the Spa-

¹ But, see Ladefoged (1992) who, like Mufwene (2001), says that speakers will determine the future of their language behavior based on what they see as the costs of maintenance vs. the benefits of shift.

² Certainly other processes beside RLS operate on the linguistic/cultural stage comprising the life of a language. Creolization, dialect splits, and language differentiation go on at all times and are worthy of study. Nevertheless, Krauss’s point still stands; 90% of the languages spoken today in the world are in danger of being lost within a generation.

nish-speaking majority—a realization that seems to be accessible to those with several years of university training, but largely ignored by those with less education. Coming full circle, I expand on the practice of divergence as a platform for language maintenance and RLS (Fishman 1991, 2001).

1.1 Sources and investigation

Mam is a Mayan language spoken by as many as 500,000 people in Guatemala's Western Highlands (Godfrey & Collins 1987). The language is further subdivided into six major dialects (Northern, Central, Southern, Western (Tacaneco), Tajumulco, and Todos Santos). Data for this paper were gathered over a number of years among speakers of the Central dialect, centered in the town of Comitancillo, San Marcos, where I lived and worked under the auspices of the Summer Institute of Linguistics from 1980 until the late 1990s. Comitancillo has a population of approximately 50,000 people, and the 1991 census claims that the municipality was 98% Mam.

The research that I present here consists of a brief analysis of four narrative Maya-Mam texts elicited between 1980 and 2002 (Appendices 1–4) and the discussion of answers to an attitude survey I gave to approximately 140 Mam speakers during the summer of 2002 (Appendix 5). Finally, I interviewed 12 Mam men and women about their opinions regarding language vitality and code-switching, also during the summer of 2002.

2 Code-switching

Gumperz defines code-switching (CS) as “the juxtaposition within the same speech exchange of passages of speech belonging to two different grammatical systems or subsystems” (1982:59). Hudson additionally calls it “the inevitable consequence of bilingualism ...” (1980:52). It is a term that has included a number of phenomena along a multilingual continuum (see also Thomason 2001:70); at one end is simple borrowing, where L2 is fully incorporated into L1 (see also Haugen 1950); at the other end is language decay and death where L1 is swallowed up by L2 (Knowles-Berry 1987).

Romaine calls code-switching “a communicative option available to a bilingual member of a speech community on much the same basis as switching between styles or dialects is an option for the monolingual speaker” (1994:59). This reflects Weinreich (1953) who discusses multiple reasons for lexical innovation in L1. Auer (1995), following Romaine, sees CS as a robust discourse strategy where code-switches (at least for skilled bilinguals) can indicate change of participant, parenthetical comments, or a topic shift, along with other discourse features. He says that access to a second language “provides specific resources not available to monolingual speakers for the constitution of socially meaningful verbal activities” (1995:115).

Knowles-Berry, though, sees the switch to Spanish within Chontal (Mayan) discourse as the top of a slippery slope that has led to Chontal's being “criticized as a mixed or inadequate language” (1987:338). With this waning reputation (perhaps either causing the criticism or pursuant to it—most likely both), parents do not teach Chontal to their

children, fewer people learn it, and it loses more and more ground and is slowly displaced even in the domain of the home, causing a hastened decline toward language death. Fishman considers this transmission of language from parents to children as the most crucial phase in reversing language shift—a stage he calls *inter-generational transfer*.

General stages in language attrition and death are widely agreed upon. Sasse (1992) says that the process starts with L1 speakers becoming bilingual in L2 and employing widespread code-switching. Next, L1 decays as more and more individuals opt to speak L2 in domains previously reserved for L1, and L2 linguistic structures and lexicon lead to a simplified, stylized version of L1. Finally L1 is replaced by L2 in all domains. Others subdivide these three stages, but the general process seems largely self-evident. For a succinct overview of others' similar views, see Winford 2003:258ff.

In communities where people speak more than one language, choices have to be made in literally every circumstance regarding which language is most appropriate for the business or pleasure at hand. When the situation or domain determines language use—for example at home or in church, with friends or strangers—CS has been termed *situational*. When the choice of a language is a statement of the cultural and social values encoded in its use, or, in other words, when language choice is used to define (or redefine) the situation, CS is referred to as *metaphorical* (Blom & Gumperz 1972). Heller writes about metaphorical code-switching between English and French in post-Bill 101 Quebec. In instances like these, speaking French is not merely a language choice; it is a political statement and a stand against what the speaker may well see as a history of linguistic and cultural oppression (Heller 1985, 1988a, 1992, 1995).

2.1 Code-switching and borrowing

In the end, whether code-switching is or becomes an indicator of language vitality or demise will prove to be more a factor of how it is construed by the speakers themselves, rather than by the predictions of scholars, but we still must distinguish between code-switching and borrowing in order to pursue the topic at hand. I maintain that for the Mam data we inspect here, there is a difference—not always crystal clear—between the two phenomena. Yet this distinction between code-switching and borrowing is important, because I claim that more educated Mam do not code-switch (although they do use occasional borrowings), whereas less-educated Mam employ not only widespread borrowing but code-switching as well.

The basic difference between a code-switch and a borrowing is that a borrowing has an L1 history. It was originally introduced by bilinguals, but now even monolinguals recognize it as part of the language, i.e., part of the lexicon of L1—part of a single grammar. Code-switches do not have this history. They show real-time decision making of a speaker who controls two grammars—at least in part. They are brought into the stream of speech consciously, as part of L2—a speaker's second grammar.³ Empirically,

³ There would obviously be a time-lag between when the first users of a borrowed term would have incorporated it into L1 and when other, more remote (either geographically or socially) speakers would pick it up. In this sense, it could be both a borrowing and a code-switch at the same time (for different speakers). But the basic insight is that a *borrowing* is part of L1, and a *code-switch* is part of L2.

whether or not a form is part of L1 or L2 is impossible to distinguish, so scholars have tried to formalize the distinction, albeit with limited success.

Myers-Scotton (1992) claims that borrowing and CS belong on the same continuum. Borrowings start out historically as code-switches or what Poplack calls nonce or one-time borrowings (1980), which with time and usage become part of the lexicon of L1. Myers-Scotton's first and main heuristic for determining the difference between a borrowing and a code-switch is frequency. Borrowings are oft repeated. Used once, they can be used again; they can catch on, and then spread, often due to perceived "lexical gaps" in the L1, i.e., meanings for which there are no apparent L1 terms. Corollary to their repeatability is the fact that borrowings can occur in widely differing contexts, and they tend to be single words, which would be far more repeatable and versatile than extended phrases. Second, borrowings show a greater degree of phonological integration than code-switches. This is not a water-tight differentiation, however, since it seems self-evident that a Spanish speaker speaking Spanish, for example, would pronounce even clear code-switches to English with a "Spanish accent". This accent (which is plainly a type of phonological integration of Spanish into English) would not make the code-switch a borrowing. Nevertheless, borrowings are more susceptible to being adapted to L1 word and syllable shapes, phonotactics, allophony, and prosodic phenomena than are code-switches. This makes sense, since a certain amount of bilingual skill—certainly including at least minimal phonological convergence toward L2—is involved in a code-switch, whereas borrowings are used by monolinguals as just another item in the language, like the use of the French word *chef* in English. One need not be bilingual in French to incorporate this term into our own lexicon. In fact, most Americans probably neither know nor care where it comes from in the first place. A good example of a borrowing from our data would be the word *plas* in line 15 of Appendix 2, reproduced here as example (1). It comes from Spanish *plaza* 'plaza, town square'.

- (1) Ex tib'aj jun tal netz' mexh-jo
 and on a cute little table-SPEC
 o k'ayini'-y toj **plas**.
 we sold-EX in plaza
 'We sold (our things) on a little table at the plaza.'

Nouns borrowed into Mam from Spanish lose any post-onset material in the final unstressed syllable of the Spanish form, and generally maintain a CVC word-final syllable shape, ultimate stress, and final devoicing as in example (2):

- (2) mula → [mulʰ] 'mule'
 calle → [kayʰ] 'street'
 tomato → [tɔmat] 'tomato'
 domingo → [dɔmiŋk] 'Sunday'

The word *plaza* fits this phonological pattern. Plus, it is oft repeated; the term is known throughout the Mam community by young and old alike.

Third, Myers-Scotton claims that borrowings are characterized by a greater degree of morphosyntactic integration than code-switches. Borrowed roots tend to be affixed and inflected just like any other roots in the language. In Appendix 2, line 2, reproduced here as example (3), the word *t-karr* ‘his car’ is affixed for possession just like any other Mam noun (the word *karr* is borrowed from Spanish *carro*).

- (3) Oxa q-b'aj-a; o xi'-y tuk'a t-**karr** Josué.
 three our-number-EX we went-EX with his-car Josué
 ‘There were three of us; we went in Josué’s car.’

Compare this with Appendix 1, line 3, where the word *vecinos* ‘neighbors’ is brought into Mam lock, stock, and barrel:

- (4) y no solamente, casi jacula txi n-q'ma'n
 and not only almost would.be.able go I-say
 ok-qe-x jni'-qe **vecinos** ...
 just-each-AUG all-each neighbors
 ‘and not only me, I could almost say just about all the neighbors, ...’

There is no morphologically marked plural in Mam; the sense of plurality in example (4) is carried by *jni'qe* ‘all them’, which makes its NP head plural even though it is not marked on the noun as such. Plus, as mentioned above and illustrated in (2), if this word were a borrowing, all post-onset material would be lost in the final unstressed syllable. Nevertheless, the Spanish plural morpheme <-s> occurs, even though one would expect [βɛsinn-] if it were a fully borrowed form. The use of *vecinos* with neither phonological nor morphological integration leads us to the conclusion that the word is a code-switch and not a borrowing.

Syntactically, the Spanish strings in Appendix 3, lines 1 and 2 (example (5) below), and again in example (6) (from Appendix 3, lines 10 and 11), demonstrate both Spanish phonological and morphological structure, as well as Spanish word order, rather than Mam.

- (5) W-aj-a tu'n n-yolin ch'in ti'j n-ja'-y.
 I-want-EX that I-speak little about my-house-EX
 ‘I want to tell you a little bit about my house.’

En primer lugar, at kab'a n-ja'-y.
 in first place exist two my-house-EX
 First of all, there are two constructions.

- (6) Ex atzin jun-tl n-okin te **cocina**,
 and as for one-other PROG-serv for kitchen
 and the other one is a kitchen

o quiere decir ja' n-b'ant-e wab'j.
 or wants to.say where PROG-made-DUR meal
 or which means, where the food is prepared.'

These are code-switches, not borrowings. According to Myers-Scotton's Matrix Language Frame Model (1997), a code-switched string (like the underlined forms in examples (5) and (6) above) will maintain the morphology and syntax of the switched language. She would call *en primer lugar* and *o quiere decir* examples of Embedded Language *islands*, since not only content morphemes are in the L2, but the grammatical morphemes as well (1997:221), even though this *island* is surrounded by the matrix language—L1.

These three characteristics (high frequency and phonological and morphosyntactic integration) of borrowings outlined by Myers-Scotton are largely corroborated by Hill and Hill in their work on Mexicano (1986), and to these I suggest the addition of two more. First, if there is a semantic difference or *enhancement* (a term I use to refer in general to what Winford (2003) further specifies as a semantic *restriction*, *extension*, or *shift*) between a form in L2 and its incorporation into L1, the L1 form is not a code-switch, but a borrowing. For example, in Appendix 1, line 12, the borrowing *lisens* ([lise'ns]) comes from Spanish *licencia* 'license'. In Mam it means 'permission', a semantic shift. The meanings are similar, but certainly not the same. Another example is the word *familiy* ([fa'miliy]) in Mam, borrowed from Spanish *familia* 'family'. In Mam, the word has come to mean 'children,' especially 'very young children.' In fact, the English sense of the nuclear family is not expressed in Mam as such; rather it is subsumed by 'those living in your house' or 'those that you always see'. This is an example of a *semantic restriction* cited by Winford. In both cases above, *lisens* and *familiy* are borrowings by our criteria, not code-switches; they are characterized not only by frequent occurrence and wide usage, and phonological and morphological incorporation, but semantic *enhancement* as well.

A second addition to Myers-Scotton's distinctions between borrowings and code-switches (albeit not unrelated to the morphosyntactic criterion already mentioned), is that it is not uncommon for code-switches to be added on top of L1 forms as opposed to substituting for them, which is more common with a borrowing. In Appendix 3, line 12, (repeated here as (7)), FL says *entonces ex*, which is a doubling up of this kind, where the Spanish word *entonces* and the Mam word *ex* both mean 'and then', and they individually serve the same discourse function of pointing toward what lies ahead in the text.

(7) Entonces ex at jun q-chuj-a, ja' n-qo chuj-in-i'y.
 so and exist one our-sweat.bath-EX where PROG-we bathe-NONFUT-EX
 'Now then, we have a sweat bath where we bathe.'

If *entonces* were a borrowing and therefore, by our definition, part of the Mam lexicon, we would expect it to substitute for *ex*, thereby obeying Mam syntactic constraints—rather than be followed by it. Nevertheless, it is fairly common in Mam CS for

a code-switched form from Spanish to do double duty with a Mam form.⁴ Since this doubling up on discourse markers conforms to neither Mam nor Spanish syntax, I conclude that *entonces* here is a code-switch, not a borrowing.

By the various criteria set forth here, we can see that there are numerous cases of CS in Appendices 1 and 3, and absolutely none in Appendices 2 and 4. Contrast example (3) with (5) and (6) above. The pattern of Spanish incorporation in examples (5) and (6) and throughout Appendices 1 and 3 where code-switching is extensive (virtually every sentence) is very different from the minimal Spanish incorporated in example (3) and throughout Appendices 2 and 4, where the only Spanish words used are fully incorporated borrowings. The remainder of this paper explores why this is the case.

3 CS as an act of social identity

As mentioned above, Heller shows how the choice of English or French in certain parts of Canada is considered a political statement and a potentially hostile act and not an innocuous search for a common language with which to discuss the weather. McClure and McClure concur, stating that the use or not of “the code-switching register, rather than any specific switch, may be used to convey social information about the speaker . . .” (1988:35). Without question, language is the flashpoint of a long history of tension over issues of political and socio-economic power. Heller further reports how the act of speaking French to a Canadian Anglophone in an official transaction is very likely a demand for respect and social reparation after years of English linguistic and cultural dominance. The negotiation of language choice in such a situation may be as innocent as trying to address someone in a language that both can understand, but, more often than not, it hints at issues of far greater import. Whether it is a case of deep political meaning attached to CS or the mere negotiation of a common language, Heller calls for interpretive approaches to the phenomenon (1988b:265). Why do people code-switch? What does it say about the people who code-switch (or refuse to) and what does it say about the societies where CS would either be seen as a positive or negative linguistic strategy? This is what is behind Blommaert’s claim that the study of CS itself is “a type of social historiography, in which the object of enquiry is fundamentally historical in nature” (1992:63). In other words, we cannot hope to explain CS behavior in purely linguistic terms. Rather, we must refer to the specific historical relationship between the people-groups that speak the switched languages and then attempt to tease out the motivations that would promote or inhibit switching between the two codes. It is right here that we can begin to understand the strong opinions and interesting behavior of those who speak both Mam and Spanish. Suffice it to say, there is always a macro-sociolinguistic aspect to CS; Haugen was correct when he said that the use of forms from a second language “always goes beyond the actual ‘needs’ of language” (1953:373).

⁴ Weinreich (1953:34) reports a kind of *reinforcement* where bilingual speakers can indeed double up on function words in speech with a complex construction comprised of one L1 term and one from L2. He does not comment on whether this would be a borrowing or a CS. I suspect CS. If it had been a borrowing, we would expect contrastive meaning (perhaps of focus, emphasis, etc.) with the non-doubled construction. I am unaware in Mam of *entonces ex* being a new discourse marker, nor does it have a corollary in Mam which has many options for focus and emphasis—none of which fit this “doubled” construction.

3.1 Studies of language and social identity

Le Page and Tabouret-Keller claim that “language acts are acts of identity” (1985). As we have already seen, language negotiation and CS can often be seen as statements far beyond mere choice of communication medium. The work of social psychologists like Tajfel and Turner (1979) has shed much light on the concept of social identity and how it relates to language behavior. Tajfel and his followers have done a number of fascinating experiments inducing people to become part of trivial groups (1974). For example, in one experiment, people were exposed under laboratory conditions to a quick flash of dots projected on a screen, and they were then grouped according to those who thought they saw a lot of dots as opposed to others who thought they saw just a few. What the experimenters found is that even in these meaningless groupings there was a strong sense of social identity—of “us” vs. “them”. Tajfel and Turner call this “the laboratory analog of ethnocentrism” (1979) and “a remarkably omnipresent feature of intergroup relations”. What is more, decisions made by individuals within the ingroup very significantly favored their own group over the outgroup. Tajfel claims that in any environment with more than one perceived grouping, the creation of an ingroup (and therefore, by definition, a non-ingroup or outgroup) is inevitable, often based not on such trivial amalgamations as the dot counters, but on distrust or even hatred of those not on the inside.

For Tajfel, the formation of groups is based on the following three-part sequence. First, people realize that they participate in certain social categories based on their education, income, dialect, gender, church affiliation, neighborhood, work, etc. Second, the perception by individuals of this category membership together with the positive and negative values held in common with other individuals that participate in the same social category (Tajfel includes these values as a kind of shared experience) determine the social identity of this group which is then compared to the identities of others. This comparison is the third and crucial step in the process. Indeed, “my group’s” social identity assumes the identity of at least one other group over against which my group’s identity exists. In other words, in order to know what something is, we must also know what it is not. This comparison gives rise to perceptions of superiority and inferiority.⁵ And the existence of a social group assumes that members of such a group construct a sense of superiority in some way over those not in the group. If, on the other hand, a group considers itself to be inferior in some way, Tajfel claims that it has several options for amelioration:

1. It can become more like the superior group in some way(s). If an individual can actually join the superior group, Tajfel identifies this situation as ripe for *social mobility*, where people can rise socially as individuals and join the more prestigious group. When the individual cannot join the superior group—usually because of some racial or ethnic characteristic—social mobility is impossible, and the only way for individuals to ascend socially is if the entire

⁵ Although he does not present it as such, Tajfel’s model of group relations is a conflict model where superior groups have a vested interest in subjugating inferior groups. It is only the existence of the “inferior” group(s) that give meaning to the concept of superiority.

group can rise in some way. Tajfel calls this a situation disposed toward *social change*, where the inferior group as a whole (as opposed to just a specific individual) becomes more in kind like the group that is perceived to be superior (even if the superior group continues to reject the inferior group).

2. The inferior group can reinterpret those characteristics considered inferior and celebrate them. This was at the heart of the *Black is beautiful* movement of the 60's and 70's or the *Gay pride* movement of recent years.
3. The inferior group can create new group characteristics which would provide a sense of positive distinction from the superior group. If history is against a group, the group can simply rewrite it and then try to sell the new version to themselves, at least, if not to the wider culture.

3.2 Communication Accommodation Theory

It is these three strategic responses that are reflected in Speech Accommodation Theory (Giles & Powesland 1975, *inter alia*) and later, Communication Accommodation Theory (CAT), an extension or enlargement of Speech Accommodation Theory. CAT was put forward as scholars realized how versatile the theory is for explaining not only sociolinguistic style shifting, code-switching strategies, and other linguistic phenomena,⁶ but also for elucidating approach-avoidance strategies in areas as diverse as fashion, advertising, and sales, to mention just a few (see Shephard et al. 2001:41). Indeed, in Giles and Johnson's initial presentation of *Ethnolinguistic identity theory* (another extension of SAT) they acknowledge that their work "draws heavily on the influential theory of intergroup behavior by Tajfel and Turner (1979) called 'social identity theory'" (1987:70).

In CAT terms, Tajfel's first amelioration strategy above is considered *convergence*, while the next two are instances of *divergence*. Divergent accommodation strategies have been less studied than convergent accommodation, although Giles and Coupland say that divergence is really a kind of convergence, the only difference being that the one(s) being converged toward are external to the real-time speaker-hearer situation (1991:80). They define divergence as referring to "the way in which speakers accentuate speech and non-verbal differences between themselves and others" (1991:65). They define convergence, on the other hand, as:

... a strategy whereby individuals adapt to each other's communicative behaviours in terms of a wide range of linguistic/prosodic/non-vocal features including speech rate, pausal phenomena and utterance length, phonological variants, smiling, gaze and so on. (1991:63)

⁶ Giles and Coupland have revisited some of Labov's work specifically related to style as a factor of self-monitoring. Their analysis (1991) is that Labov's findings could be recon-structed as "interpersonal accommodation processes". In other words, Labov's interviewees perhaps were responding to cues in the interviewers' speech rather than the supposed formality/informality of the context itself.

Both convergence and divergence are mechanisms for achieving solidarity group-internally and maintaining positive social identity vis-à-vis outsiders. In order to determine whether and how these concepts shed light on Mam speakers' attitudes toward CS as a strategy for convergent or divergent accommodation, I designed a questionnaire that sought to identify opinions about language vitality and code-switching. The entire questionnaire (translated into English) is attached as Appendix 5. The questionnaire and its results are discussed in §3.2.1.

As CAT has been applied to various speech situations, it has been enhanced and extended. An area of special interest to us in the present study is the issue of *typicality*. Typicality refers to a situation in which a speaker converges toward or diverges from a stereotype. It is convergence toward an ideal, not necessarily toward an actual speaker. This is developed in Gallois and Callan's "Stereotypically driven accommodation" (1988), where the speaker's perception of whom he or she is converging toward or diverging from is more telling than the actual convergence facts of the situation. This cognitive/attitudinal aspect of typicality shows the true social psychological core of CAT (Thakerer et al. 1982), and it was to explore these attitudes that I developed the questionnaire seeking native-speaker opinions on Mam language vitality and code-switching behavior. After a discussion of the questionnaire, I return to a discussion of typicality.

3.2.1 An attitude questionnaire

As mentioned, the questionnaire was developed to help gauge people's attitudes on two issues, language vitality and code-switching. Some of the information requested was general in nature (age, profession, sex, religion, plus questions 1–4), and several questions were opinion questions or requests for information that did not lend themselves to easy quantification (questions 5, 11, 13, 15, 19, and 20—see Appendix 5 for details).

Questions specifically about language vitality and prestige are numbers 6 ("Are you ever embarrassed to speak Mam in front of native Spanish speakers?"), 7 (a question about switching languages in "mixed company"), 8 (a question about ranking languages by importance), 9 (a question about requiring university students to learn a Mayan language), 12 ("Do you think it's good for native Spanish speakers to learn to speak Mam?"), and 14 (about the long term survivability of Mam). The questions about code-switching are 16 ("Mixing Mam and Spanish in a single conversation is good or not?"), 17 ("What kind of people mix the two languages?"), and 18 (where interviewees are asked to agree or disagree about specific reasons for CS). Question 7 (mentioned above as a question about language vitality) also deals with CS.

In order to look for correlations between level of education and attitudes, I stratified the interviewees into three levels of education: School teachers (who have had the equivalent of roughly two years of junior college), students in their first of three years of teacher training (hereafter, *trainees*), and people with less than a high school education. This last group ranged from no years of schooling up to nine (grades 7–9 are considered "high school" in rural Guatemala).

I reduced a number of the questions to two-way (yes-no) or three-way (Mam-Spanish-both) answers, and I entered these in matrices to the statistics package, *StatXact-4*. I ran a chi-square test using Monte Carlo methodology, since, although I interviewed over 100 Mam speakers, this is still a relatively small sample of the 60,000+ speakers of Central Mam. I chose a p -value of < 0.01 .

The interviews themselves were conducted in two different ways. I was given permission to attend a local teachers' meeting and to take ninety minutes to discuss (in both Mam and Spanish) the questionnaire and the issues that it was meant to elucidate. I gave the questionnaire out to approximately 60 primary school teachers, and I discussed each question, one at a time, to make sure they were all clear. Although I discussed each question with the entire group and dealt with issues that arose, each teacher answered each question as he or she saw fit, filling out the survey form as I discussed each question.

I handled interviews for the teacher trainees in much the same way. I presented the survey to a large group (approximately 60 students, over the span of ninety minutes), I discussed each question, and I had students give their own answers on the survey form.

For the less-educated Mam, I wrote the answers myself on the survey form based on answers given to me in one-on-one interviews. These interviews lasted approximately twenty minutes each. Some of these people could read but others could not, so I filled out the forms myself while doing the individual interviews. I interviewed 30 less-educated individuals (less than eight years of schooling).

CODE-SWITCHING AND REVITALIZATION IN MAYA-MAM

	Teachers			Trainees			Less educated		
	Mam	Span	Both	Mam	Span	Both	Mam	Span	Both
What do you speak at home?	15	2	2	42	9	10	21	1	6
What do you speak in the street?	8	2	9	15	14	21	12	5	13
What do you speak at church?	5	1	10	10	10	32	2	4	12
What do you speak with your brothers and sisters?	11	2	6	38	7	13	17	3	10
What do you speak with your parents?	15	2	2	43	6	7	20	2	6
What do you speak in the municipal offices?	6	6	4	9	13	34	10	10	10
What do you speak with friends?	6	2	9	9	11	34	11	3	8

Table 1. Raw data from question 4: Language(s) spoken in different domains.

For the questions regarding domains (outlined in Table 1), there were no significant differences among the three groups. The same is true for questions 7 and 8 (Table 2). There were no significant differences among the three groups for these data.

	Teachers			Trainees			Less educated		
	M	S	B	M	S	B	M	S	B
7. If you are speaking Mam with friends and a <i>ladino</i> comes up to you, what language do you speak?	5	8	5	18	15	17	15	13	5
8. Rank languages by importance to you.	4	5	25	10	6	42	5	1	13

M = Mam, S = Spanish, B = Both

Table 2. Questions about language vitality with 3-way answers.

| Teachers | Trainees | Less
educated

	Yes	No	Yes	No	Yes	No
6. Are you ever embarrassed to speak Mam in front of <i>ladinos</i> ?	0 0%	36 100%	6 10.5%	51 89.5%	0 0%	33 100%
9. Should <i>ladino</i> university students be required to learn a Mayan language?	25 83.3%	5 16.7%	37 77.1%	11 22.9%	15 78.9%	4 21.1%
12. Do you think it's good for <i>ladinos</i> to learn to speak Mam?	18 94.7%	1 5.3%	57 100%	0 0%	37 77%	11 23%
14. Do you think Mam will be lost?	0 0%	36 100%	14 24.1%	44 75.9%	4 21.1%	15 78.9%

Table 3. Yes-No questions about language vitality.

Of the questions in Table 3, numbers 12 and 14 proved to be significant with a *p*-value of < 0.01. It appears that the clearest opinion towards language vitality is expressed in answers to question 14. Here, teachers, who categorically state that Mam will survive both linguistically and culturally, group significantly against the trainees and less-educated. In question 12, trainees group with teachers vis-à-vis the less-educated. In both questions, teachers and less-educated are significantly different. Nevertheless, I discarded 12 from further comparison because the comments on the “yes” answers to the question were actually quite different in nature between the teachers and the less-educated. Teachers said that *ladinos* should learn Mam because it is a matter of equity and fairness.⁷ The teachers had to learn Spanish; it’s only fitting that *ladinos* should learn a Mayan language. Also, teachers felt that it was good for *ladinos* to understand worldview issues and cultural matters available only through language. Less-educated Mam felt that it was good for non-native speakers to learn a Mayan language in order for them to be able to communicate better with the Maya. In other words, the less-educated see language as a communicative tool, whereas teachers tend to see it as a symbol of cultural equality. So, despite “yes” answers on the questionnaire, I considered the additional comments to be such that the answers from the two groups could not be conflated into a single category.

Table 4 shows the raw scores for people’s opinions about reasons for CS. People were asked whether or not they agreed with a number of reasons as to why native Mam

⁷ *Ladino* is a Guatemalan term for a Spanish speaker. A *ladino* might be ethnically Mayan, but has adopted Western dress and values and, most importantly, has abandoned his/her native Mayan language.

speakers would code-switch. These answers do not necessarily reflect their own opinions. Rather, their answers represent what they perceive to be the reasons code-switchers have for switching languages. The only significant reason ($p < 0.01$) is the first one, but since these are responses based not on assessment of their own usage, but on what they think others must have as motivation, I do not consider them true indicators of how people actually assess their own reasons for CS. For example, fifteen of the seventeen teachers who responded to this question said that code-switchers would no doubt argue that the lack of sufficient vocabulary in Mam is motivation for using Spanish in a Mam conversation (see row 1, Table 4). Nevertheless, it is clear that the teachers themselves do not agree that this is a legitimate reason, since they use *neologismos* (new word forms) and circumlocutions to avoid CS in their own speech and they generally frown on CS as a linguistic option. They were extremely positive in response to questions 19 (“Is it worth coming up with dictionaries of *neologismos*?”) and 20 (“Do you think people will use these *neologismos* once they are formulated?”).

In addition to these reasons that interviewees responded to, teachers added the following reasons on their own. Three said that people code-switch because they lack interest in their own language. Two more said that most speakers don’t even realize that they are switching. Two trainees added that there was a general lack of interest, while one said that CSers were not well taught in Mam, and that’s why they switch. Finally, among the less-educated, three said that switchers don’t know Mam well; three more claimed that people don’t realize that they are switching; one said that people switch because they don’t investigate how to speak Mam well; one added that the old Mam words are hard to remember; and one claimed that close contact with Spanish speakers causes people to code-switch.

Regarding views concerning CS, the one question that significantly groups trainees and the less-educated against teachers is number 16 (“Mixing Mam and Spanish in a single conversation is good, bad or neither?”). Sixty-eight percent of the teachers rejected the idea that CS was “good”, while almost 70% of both the trainees and the less-educated stated that CS was either “good” or that “it didn’t make any difference”.

In summary, there are two questions in the questionnaire that clearly differentiate groups: Question 14 about language vitality (Table 3) and question 16 about code-switching (see discussion just above). In both cases, trainees and less-educated Mam group against formally trained school teachers. This difference is developed more in §4.

	Teachers		Trainees		Less educated	
	Yes	No	Yes	No	Yes	No
The words don't exist in Mam, so using Spanish is the only way to get our meaning across.	15 88.2%	2 11.8%	16 34.7%	30 65.3%	17 100%	0 0%
Spanish is used to show-off.	4 33.3%	8 66.7%	15 38.5%	24 61.5%	10 58.9%	7 41.1%
Speakers appreciate both languages and want to use them both together.	8 50%	8 50%	33 70.2%	14 29.8%	15 93.8%	1 6.2%
Spanish is more prestigious, so speakers want to use it when they can.	12 66.7%	6 33.3%	21 45.7%	25 54.3%	11 73.3%	4 26.7%
CSers learned to talk using both languages. It isn't their fault.	15 88.2%	2 11.8%	35 72.9%	13 27.1%	15 88.2%	2 11.8%
CS is the actual and modern way to speak Mam.	7 46.6%	8 53.4%	32 64%	18 36%	14 87.5%	2 12.5%
CS speakers are lazy and don't want to do the work involved in finding appropriate Mam forms.	8 53.3%	7 46.7%	23 62.2%	14 37.8%	8 53.3%	7 46.7%

Table 4. Yes-No opinions regarding people's reasons for CS.

3.2.2 CS as stigmatized behavior

Our data show that less-educated Mam code-switch often (note underlined portions in the narrative texts in Appendices 1 and 3). When queried on this, this group says that they are so ensconced in both cultures that it is impossible for them not to code-switch. Questions in the questionnaire aimed at probing subjective judgments on this issue are numbers 16–20, particularly 16 (“Mixing Mam and Spanish in a single conversation is good or not?”) and 17 (“What kind of people mix the two languages?”). A common response by this group to the survey question on whether or not code-switching is good or bad was that “this is the way we learned to speak from our childhood. It is therefore not our fault and language mixing shouldn't be stigmatized.” Almost 70% of

those surveyed who were less educated claimed that code-switching was either good or that it didn't matter one way or the other. This compares to 68.8% of the teachers surveyed who stated that CS is unequivocally bad. Question 18 (Table 4) sought people's opinions as to the reasons why code-switching is so widespread among the Mam. There was a majority opinion across the three groups that CS was due to lexical gaps in Mam, the perceived high prestige of Spanish, and the history of routine language mixing throughout Mam-speaking society.

Despite the generally positive view of CS by the less-educated, CS is nevertheless stigmatized by two salient groups. First, many—perhaps most—monolingual Spanish speakers claim that Mam is not a *real* language at all, but rather a *dialecto* at best, which is what the Mayan languages are called in the national schools. Languages are taught as being national or transnational; dialects are regional. Some people have told me in interviews that the Mam communicate through primal gestures and grunts, “like animals”. Although this is clearly an extreme view, it is equally clear that minority languages in Guatemala are considered less than full languages. Adherents to similar views bolster their claims with the fact that most Mam speakers fluently and frequently code-switch, supposedly demonstrating the inability of the Mam *dialecto* to lexicalize important concepts which must therefore be articulated in Spanish (and often, in not prescriptively acceptable Spanish). If Mam really were an adequate language, they reason (assuming it is a language at all), speakers would not need to resort so often to Spanish to express themselves. So rural Spanish speakers stigmatize CS.

The second group that stigmatizes CS is Mam teachers, who claim that to code-switch is to buy into the idea that Mam cannot be used to articulate complex ideas, philosophies, and technologies. In their minds, code-switching supports the erroneous assertion that Mam is indeed an inferior language.⁸

But if code-switching is stigmatized, why is it so common? The survey and the interviews with code-switchers lead me to suggest that it is due to the sense of convergence to the majority culture. The Mam historically have not had the access to resources that the Spanish speaking majority has had for centuries. Wealth and political power are concentrated in the hands of *ladinos*. Education—especially higher education—is more accessible to *ladinos* than to the Mam; the radio and cable TV beamed into the Mam area are almost exclusively in Spanish, as are the daily newspapers. Many school teachers and other government officials are native Spanish speakers. In view of such a great power mismatch, CAT would suggest that these Mam code-switchers see CS as a convergence toward the trappings of power and prestige. Table 4 shows people's subjective judgments as to why CSers mix Spanish and Mam.

⁸ This sense of inferiority is combated locally by referring to the history of the Maya. Their advanced learning was put into a writing system still seen in various glyph sites throughout Guatemala, Mexico, and Honduras. The Maya were also advanced in agriculture, astronomy, and math. The fact that these technical and cultural advances were expressed via a Mayan language is clear proof of the language's ability to articulate complex ideas.

Yet the CS that they employ is stigmatized by the very people to whom the less-educated Mam would want to converge. In fact, this type of CS is common even when everyone in the audience is a Mam speaker. This shows that CS is an attempt to identify with the people (even though they are not present) that have the power and prestige—the *ladinos*.

They are trying to show to themselves and to those who will pay attention that they have feet firmly planted in both worlds—the Spanish world of power and the Mam world of Mayan values and culture. This straddling of the fence is normally a part of Mam evangelical services. Routinely, the leader will stand at the beginning of the service and speak a little in both languages. He will then ask whether the service should proceed in Mam or Spanish (since he—and those in attendance—are supposedly equally competent in both). The answer is always *los dos* ‘both’. Under the circumstances, CS is the best they can do, since their access to real power is minimal.⁹ Parents have told me that Spanish is the only hope for their children. Land is scarce and very expensive; families tend to be large, and inherited land is not extensive enough to raise the crops needed to sustain life for a family and animals. If their children are to prosper, they cannot rely on the life that Mam alone would give them.¹⁰

3.2.3 Typicality and Mam/Spanish convergence

In total opposition to these less-educated Mam, when Mam teachers speak Mam, they do not code-switch at all. The only underlined portions in Appendices 2 and 4 are borrowings—and even these are rare. Over 50% of these teachers believe that CS is a sign of laziness or lack of interest (see Table 4). Although almost 90% (15 of 17 educated respondents) think that CS can be caused by perceived lexical gaps, many of these same people voiced the opinion that it is up to the individual to purify his or her language and investigate how to say what needs to be said in Mam alone without any recourse to Spanish. In other words, these lexical gaps should be filled, if at all possible, with Mam lexical items.¹¹

⁹ This is reminiscent of Eckert's (2000) claim that women and girls seek cosmetic and symbolic power because they are denied access to real power.

¹⁰ Ed Beach of SIL (in private communication) has told me about a generation of Maya-Tectitec (a language closely related to Mam) who speak neither Spanish nor Tectitec natively. Parents had decided to speak to their children only in Spanish so that they could supposedly help them advance socio-economically, but the parents themselves were not adequate speakers of Spanish. So the children acquired a virtual Spanish-Tectitec creole, rather than native Spanish or Tectitec proficiency. Nils Hansegård calls this phenomenon *semi-lingualism*. Schaengold (in private communication) reports a comparable situation for Navajo, except that these “Navajo creole” speakers are native speakers of English, and speak a very stylized version of Navajo. I expect this is not an uncommon occurrence among minority cultures around the world.

¹¹ What Guatemalan educators call *neologismos* ‘new words’ has become a virtual growth industry among Mayan bilingual educators. This is basically about finding Mayan ways to fill lexical gaps. I’ve seen dictionaries in Mayan languages suggesting words for concepts like computer, carburetor, penicillin, bus, etc. In an interview with the local bilingual school superintendent, he said that there is a protocol for developing *neologismos*. First is function: *iqb'il xjal* ‘carrier of people’ is the suggested term for bus. Second is physical appearance: *txaq sotz'* ‘bat wing’ is a widely accepted Mam term for ‘umbrella’. Third is

So when educated Mam speak Mam without any CS, what are they doing? What they say in interviews is that they are diverging from Spanish. They are tired of being socially and culturally stigmatized, and they are fed up with the *ladino* perception that *ladinos* as a class are superior to the *indios* and that their beautiful Mam language is considered inadequate to the modern age. In an interview with a Mam college professor, he expressed interest in survey question 9 (“Do you think university students should be required to learn a Mayan language as a prerequisite to graduation?”). He told me that he has two answers to that question—one public and one private. His answer for public consumption is that *ladinos* should be required to learn a Mayan language because Guatemala is a plurilingual/pluricultural society and exposure to a Mayan language would give *ladinos* a less jaundiced view of the Mayan world. This can only be good, he said. But his private answer is much darker. He said that *ladinos* would undoubtedly find a way to take advantage of Mayans if they spoke their language. He cited parts of the country where *ladinos* routinely speak a Mayan language, and he said that those were the very parts of the country where the daily wage was lowest. So his real answer to the question is “No! Our language is the one place where *ladinos* can’t go, and let’s keep it that way.”

As stated above, Giles and Coupland claim that divergence is a kind of shifted convergence, with the targets of convergence being located outside the environs of the speech event. Assuming this to be the case, what would these non-CSers be converging toward? It cannot be sustained that they are moving toward a core group of respected Mam elders, since these very elders (like PT in Appendix 1) are among the ones who rampantly code-switch for reasons given in Table 4 above. Rather, they are converging toward a stereotype of what a “typical Mam” should be. A Mayan, after all, should certainly speak a Mayan language. Because the target of this convergence is not an actual social group, Gallois and Callan’s *accommodation to a stereotype* (1988) is highly relevant here, since it is not the case in this type of convergence that accommodation is toward the speech patterns of flesh-and-blood individuals (Giles & Powesland 1975). Rather, these educated Mam are converging toward an ideal, not an actual interlocutor.

4 Is education the big differentiator?

As mentioned in 3.2.1 above, during the summer of 2002, I surveyed three groups of people: Teachers, trainees, and less-educated. What I expected to find was basically a straight line relationship between years of education and divergence strategies of accommodation—that the more education a person has, the less he or she would code-switch. I was surprised to find that it was only the teachers who were so adamant about boundary maintenance and the rosy future of Mam language and culture.

Also, in interviews, educated Mam were careful to avoid CS in talking to me and the others present, whereas all others routinely employed it, whether they were trainees or less-educated. So teacher trainees consistently grouped with those less educated than themselves rather than with teachers.

an illustrative phrase. Finally, if none of these strategies works, the concept is borrowed from Spanish (respecting Mam phonological constraints) and promoted.

One last piece of evidence is the stark contrast documented in Appendices 2 and 4 between two similar texts offered by the same man twenty-two years apart. In Appendix 2, FL gives a talk about his house. At the time he was 23 years old and had been to school for just three years. Virtually every line of his discourse employs CS.

In Appendix 4, FL gives another talk about his house. This time, however, he is 45 years old and has a university education. This time his text has absolutely no CS. Since this is a longitudinal comparison, I attribute his change in CS behavior to be his education, rather than anything else, much of which would have been controlled for, since both texts are offered by the same person. It certainly is not just age. Less-educated Mam of FL's age code-switch routinely. Nor is it contact with me, one who tries to model (albeit unsuccessfully) "pure" Mam with no CS. Many of my Mam colleagues and co-workers speak as in texts 1 and 3, whereas the teachers (irrespective of whether they work with me) are careful to avoid CS, despite the fact that their Spanish is excellent. Indeed, they speak both languages very, very well.

What would cause teachers to be so militant about language vitality and CS as opposed to the other two groups? I suspect it has to do with two things: training opportunities and role responsibilities. First is the training itself. Throughout the first nine years of school (six years of primary school and three years of "basic" education), learning is essentially by rote memory, and oriented to facts, lists, and formulas. After these nine years, students can opt for a professional track, either education, business, or pre-university. These tracks begin with much of the same and gradually give way to more analytical, cerebral pursuits and the formation of informed opinions. It is during this time—during the second half of teacher training—that the more militant and confrontational attitudes of the teachers are formed. A main issue I see in this is that of causation. Earlier in their education, the Mam are learning "how things are"; as they continue their schooling, they come to see their socio-cultural situation not as part of "God's plan" or some kind of predetermined fatalism, a characteristic of Mayan culture (see de Landa 1566 and Martínez Peláez 1970), but as something caused by years of oppression, injustice, and restricted options. This is not an automatic realization; it comes through much discussion and orientation. I have attended many of these sessions where the practical aspects of this history are being considered, discussed, and reacted to. This kind of discussion (sharply politicized) has certainly been true of some of the rural union organizations as discussed by Nobel laureate Rigoberta Menchú (Burgos 1985), but it is discussed in a more systematic and empirical way as Mayan students are taught upper-level university courses by Mayan professors.

Aside from advanced training, once a person becomes a teacher, there are numerous seminars or "professionalization" workshops, further university training and opportunities for *concientización* ('raising of consciousness') where many of these ideas are hammered out.

The second reason I think that teachers group separately from trainees and those less-educated has to do with their role in Mam society. There are clear role distinctions between Mam men and women, and roles among people are also distinguished. Certain men burn limestone, others cut trees for lumber, some are potters, and still others become

traditional priests or folk healers (Redfield 1941). Roles and responsibilities are clear for these distinct tasks. The Mam themselves say, “Each person has his own work”. Some of the Mam are teachers. And these teachers feel a strong responsibility to model and promote a better way for their people—what they consider a way of dignity and respect.¹² In regard to CS, these teachers consider it a capitulation to the oppressors. Interestingly, although they are excellent Spanish speakers, they choose not to mix Spanish and Mam no matter which of the two languages they are speaking. When they speak Spanish, they do so articulately; when they speak Mam, they do so without CS.

I suggest that this militancy bodes well for the long-term viability of Mam vis-à-vis language shift, despite its status as an endangered language. The fact that the Mam themselves are taking up the torch in relation to their own language and culture is a very positive sign.

5 The reversal of language shift

By studying a number of minority languages in different stages of decline and endangerment, Fishman has suggested a Graded Inter-generational Disruption Scale (1991). In a sense, it is an eight-point scale of how viable a language is in specific multilingual situations around the world.¹³ The scale also maps a general strategy for language revitalization. For example, at stage eight, a language is little spoken and must be recapitulated from the memories of aged speakers, from recordings, and from other records in order for it to be learned by adults. At stage seven, there is use of the language only among adult speakers. At stage six, the language is passed down to children who speak it as a first language. At stages five and four, children are formally educated in the language; levels three, two, and one have to do with language use in daily work, higher education, and government respectively. So not only does the scale situate a language on a dimension of viability, but it also indicates steps toward improving a language’s chances for survival.

Fishman says that level six is the critical one for language vitality. If children are not learning the language from their parents in the home, there are two major problems. First, there is no *domain* where the minority language is safe from intrusion by the majority language. Second, children do not have the benefit of adult modeling of the language in daily cultural life. He says that programs promoting language and cultural revitalization must focus on stage six, and all further stages must make “inter-generational transfer” top priority. In other words, using the language in higher education or government cannot substitute for the language not being spoken as a first language by children.

¹² In addition to the linguistic strategies mentioned here, Mayan teachers have adopted Tajfel’s third ameliorative strategy—revisionism—to help create a positive sense of cultural space for their people. Many Mayan teachers deny that their forebears were animists and that they practiced human sacrifice. They reinterpret the *Popol Vuh* in those instances that discuss human sacrifice, or they claim that the text has been altered by outsiders. The *Popol Vuh* is considered by many to be the Mayan Bible, a book filled with history, folk wisdom, and Mayan mythology.

¹³ Fishman recommends RLS whether the predictors of success are realistic or not. He claims that significant progress can be made toward “saving” threatened languages even when the situation looks bleak. Thomason (2001:82) would concur on this point. She points out that speaker attitude is more important than any other factor when it comes to language maintenance.

If children are not learning the language at home from their parents, everything else is tantamount to tidying the curtains in a burning house. That this is self-evident is clear from reports of heavy national investment in programs in Ireland that endorse Irish language prestige and use (Ó Riagáin 2001). Legislation has been enacted to support Irish and the government has promoted it in the schools; it has been used throughout the media and people have looked on it with pride. Nevertheless, for those living outside of the most intimate Irish sectors, only one quarter of those who grew up speaking Irish at home are establishing Irish-speaking homes themselves. Despite the highest of ideals and heavy investment, the majority of the Irish people themselves are largely passing English on to their children—not Irish.

In a situation somewhat similar to that of a number of Mayan languages, Quechua is a minority language spoken by several million people. Although the language seems safe from demise, Hornberger and King (2001) consider it threatened in its homeland—Peru. The percentage of Quechua-speaking monolinguals is falling (from 31% of Peru’s population in 1940 to 11% in 1982), while the percentage of Spanish monolinguals is rising (from 50% in 1940 to 72% in 1982). Bilingualism is largely subtractive: i.e., Quechua speakers tend to learn Spanish and then drop Quechua. Although Hornberger and King praise the work of both governmental and non-governmental organizations in trying to help support Quechua literacy efforts and language revitalization programs, they point out that it is the speakers themselves who must ultimately decide whether to save the language or not. The inexorable encroachment of Spanish into the home and family domains puts the long-term future of Quechua on shaky ground. Promotion *for* the language community will never take the place of promotion *by* the language community. Lee and McLaughlin discuss similar findings (and a similar analysis) for Navajo (2001), as does Lastra for Otomí of Mexico (2001). (For another view on whether endangered languages should be “saved”, see Ladefoged 1992 and Mufwene 2003.)

Central Mam seems to be fairly secure at the present time. But the decisions by members of a single generation can bode ill for language maintenance (see note 10). Fishman points out that there is much overlap among stages in the Graduated Inter-generational Disruption Scale (GIDS) and that activities to reverse language shift can and should be taking place all along the scale that would promote language use in general, and, in particular, that would promote the transfer of the language from parents to children. In considering Central Mam in relation to the GIDS, there are a number of important ties.

1. Mam teachers are working to reassemble parts of the language (stage 8), rediscovering words and syntactic patterns from the elderly—or generating new terms that are Mam if at all possible, rather than Spanish. For example, numbers beyond twenty are rarely used and largely unknown by most Mam speakers, even though the Classic Maya had a robust number system. Mam (and other Mayan) scholars are attempting to reconstruct the number system rather than use Spanish numbers, which is what virtually everyone does in day-to-day speech.

2. These same teachers have also established forums where Mam should be spoken with absolutely no CS. This includes local radio announcements, signage (on both private and public buildings and other venues), and in meetings of Mam-speaking teachers. They also try to influence pastors and priests, writers, merchants, and others who they feel are shortchanging the language by mixing it with Spanish. These are basically stage 7 activities where adults are encouraged to speak Mam.
3. Fishman's stage 4 has to do with government schools; stage 4a is the use of the minority language in schools that are under the control of native speakers of the minority language. Local Mam speakers are now the supervisors of bilingual education throughout the area and most teachers in area schools are native speakers of Mam. This has been a stated goal of local teachers—taking responsibility for area schools.
4. Local leaders have promoted the use of Mam on the radio; they are supporting the publication of books and a local newspaper. They are the ones that are the most vocal in support of requiring university students to study and speak a Mayan language (question 9, Appendix 5). These are stage 2 and 3 activities on the GIDS.
5. Mayan speakers have lobbied for the recognition of their languages as official or nationally recognized codes of communication. Although the latest constitution did not grant official status to any language other than Spanish, Mayans were able to include supportive language that recognized the historical and traditional importance of languages other than Spanish. The constitution has been translated into most of the nation's languages and there is provision for translation services in court; and early childhood education is guaranteed in a student's native language. These guarantees look good on paper, but they usually are not carried out. Nonetheless, they give the promoters of RLS a legal leg to stand on, and they represent stage 1 on Fishman's GIDS scale.

5.1 Inter-generational transfer

All of these are important strategies and are supportive of Mam revitalization, but none of them will matter if stage 6 is not respected, that of passing the language on to children in the home. The survey questions on language domain show that across the board, a high percentage of Mam speak only Mam at home and with both older and younger family members (see Table 1).

In a situation like this, it seems reasonably safe to say that Mam will continue strong for another generation—at least in Comitancillo. But as experience in other Mam dialects (Godfrey & Collins 1987) and in multilingual situations around the world (Fishman 2001) has shown, it really takes only one generation to seal the fate of a minority language. That said, it is far easier to maintain a minority language than to resurrect it. This is why Fishman's stage 6 is so critical. Strategies centered around the other stages are supportive of language maintenance, but stage 6 is operational.

Over the last decade or so Mam teachers have been carving out a positive social identity for their own people via various means—through a revisionist history, through a “purified” language, and through advanced education. And they are in the process of defending this newly tailored identity in excellent Spanish to groups across the country and around the world. Whether these strategies will result in the mainstay of RLS—inter-generational transfer—remains to be seen.

5.2 RLS in context

Languages are not maintained in a vacuum. Mufwene points out that in multi-lingual situations, languages are not so much abandoned as deliberate action, but rather as “the cumulative consequence of repeated communicative acts” (2002:387). These acts are benefit-driven. If a person needs English or Spanish or Mandarin to make a living, it is unlikely that the native language will long be able to withstand these major languages’ intrusion into the domain of the home, which will certainly affect a language’s position on the GIDS. So the larger language context is always important. (For an interesting and surprising discussion of the *extension* of language domains, see Hartman Keiser 2003.)

In Guatemala, a number of factors have converged to help promote the possibility of language and culture revitalization:

1. The awarding of the 1992 Nobel Peace Prize to Rigoberta Menchú, a Mayan woman, has brought international attention to the plight of Central and South American indigenous groups. Mufwene considers politics the second bulwark of language maintenance (economics is first).
2. The National Bilingual Education Program has been funded for over twenty years. Well over a million Mayan children have received primary school instruction in their native language. This has heightened interest in native language literacy and literature production.
3. Three national universities have established applied linguistics programs that have helped train hundreds of Mayan teachers and professionals.
4. In 1991, the Academy of Mayan Languages became an autonomous national institution (for details see England 1998:106). This grass-roots organization provides a forum for Guatemalan Mayan people to discuss and to resolve issues related to Mayan life and particularly Mayan languages. Despite many challenges, the Academy has been a focal point for Mayans promoting native language and cultural maintenance.
5. The signing of the Peace Accords in December, 1997 have led to greater respect for human rights (including indigenous rights) throughout the nation.
6. Tourism has become the main motor of the Guatemalan economy, replacing coffee as the nation’s greatest dollar earner. Tourism officials realize that much of what tourists come to see is Maya-related, so the government has a vested interest in Maya language and culture maintenance.

7. Comitancillo, although open to outsiders, is still populated almost exclusively by insiders. One need not speak Spanish at the market or in most churches. Local government offices have Mam-speaking attendants. Mam books are available. A Mam radio station was established in 2002. A monthly newspaper funded by local businesses is in the works. Although Spanish is highly valued, Mam is valued as well.

In other words, there is a support structure for Mayan revitalization. In Mufwene's words, there is the "concurrent mobilization of the political and economic machineries" (2002:390) that enables the good start by Mam teachers to be buttressed by a context in which being Mayan and speaking a Mayan language is benefit-driven, not only politically and economically, but socially and educationally as well.

Perhaps this many-pronged attack—especially since it shows strong initiative by the Mam themselves—will enable the language not only to survive, but also to prosper. Whether this happens, or whether the movement is reduced to an elitist notion promoted by a handful of teachers, will depend on the daily language decisions of the masses.

6 Conclusion

In this paper I have applied the insights of Communication Accommodation Theory to an understanding of code-switching behavior among the Mam. It was found that the Mam masses use CS promiscuously, apparently as a strategy of convergence toward the prestige and power of the dominant Spanish-speaking culture. Mam teachers, on the other hand, avoid CS, claiming that it is a slap in the face of Mam language and society. Rather, they pursue a policy of *Mam only* when speaking Mam—an act which has here been analyzed as a strategy of divergence from the dominating culture and convergence toward an idealized notion of Mayan-ness, an icon of linguistic and cultural egalitarianism among all cultures. The behavior of both groups reflects a concern for positive social identity.

The masses understand that the power structure of Guatemala is oriented toward Spanish. They see that they have been left behind politically, educationally, and economically. Their convergence strategy is an attempt at a bigger piece of the pie. Teachers, on the other hand, have proved that they can be successful in engaging the majority culture—at least in educational terms. They are graduates of the *ladino* education system. They have come to realize that the racist nature and history of Guatemalan society makes social mobility impossible.¹⁴ If change is to come at all it will be community wide in the sense of *social change*, or "a rising tide raising all ships". Their strategy of social change is two-fold. It looks inward toward Mayan culture and attempts to remake it as an equal partner to the dominant culture—witnessed by their ability to speak fluent and unmixed

¹⁴ Anecdotes of this latent and not-so-latent racism abound. In one interview, a Mam teacher told me he was the featured speaker at a university function, but when he showed up in simple dress and dark skin, he was denied entry to the auditorium. He waited patiently outside until the event organizer went to search for him. This same man told me that a Mayan doctor would never be allowed to treat a *ladina*. When all is said and done, "we are still just despised *indios* in the eyes of members of the dominant culture".

Mam. At the same time, it also looks outward in an attempt to promote this equality—in excellent Spanish—to the world beyond the village.

What I've described here is basically just one part of the mosaic—the *no CS* strategy of Mam teachers—and how they see a purified Mam fitting into the larger picture of the reversal of language shift.

Appendices 1–4. Abridged texts.

Abbreviations used:

AUG	augmentative	NONFUT	non future
CAUS	causative	PAS	passive
DUR	durative	PERF	perfect aspect
EX	exclusive	PROG	progressive aspect
FUT	future	PUNCT	punctual
IMP	imperative	REC	recently completed aspect
LOC	locative	SG	singular
NEG	negative	SPEC	specifier

Alphabet:

In the appendices that follow, orthographic conventions are largely taken from Spanish, with the following modifications:

b'	implosive bilabial stop
ch'	glottalized alveo-palatal affricate
j	uvular fricative
q	uvular stop
'	glottal stop
k'	glottalized velar stop
ky	palatal stop
ky'	glottalized palatal stop
q'	glottalized uvular stop
tx	retroflexed alveo-palatal affricate (backed <ch>)
tz	alveolar affricate
tz'	glottalized alveolar affricate
xh	alveo-palatal fricative
x	retroflexed alveo-palatal fricative (backed <sh>)

Appendix 1. PT, no formal schooling, 68 years old. Text taped 1982.

1. Entonces n-qo yolin ti'j-jo mejeb'lin ojtxa.
so.then PROG-we talk about-SPEC marriage long.ago
So we are talking about marriage in the old days.
2. Atzin we xi n-li te xhin jaw ti'jin
that me go I-saw when I up grew.up
Now as for me, when I grew up,
3. y no solamente, casi jacula txi n-q'ma'n
and not only almost would.be.able go I-say
ok-qe-x jni'-qe vecinos
just-each-AUG all-each neighbors
and not only me, I could almost say just about all the neighbors,
4. pero manera que mina,
but in.a.way that no
but in a way, no,
5. porque ex ma chin meje-tz-a pues.
because and REC I married-then-EX then
since I am married
6. Entonces atzin te costumbr te t-xmoxin ojtxa;
so.then that the custom when he-convince long.ago
qa ti'j xmoxin pues,
if regarding convincing well
Now then, the custom of when a boy would find a wife long ago, the issue of courting,
7. Entonces ma qo aj q-kanin, como at jun respeto,
then REC we go.home we arrived since there a respect
If regarding the taking of a wife, well then, when we arrived, since there is a kind of respect,
8. nejku aj-tzin q-xi xmoxil pues,
first when-then we-go to convince then
Since there was a kind of respect, first when we would go to take a wife,
9. n-xi q-q'ma'n te q-tata, awo ichin-qo o q'a-qo, te q-nana
PROG-go we-say to our-father we men-us or boys-us to our mother
We men, or more correctly, we boys would tell our father and our mother (that it's time that I look for a wife).

10. Entonces, “Ku txi'y” qa chi,
 so.then IMP go if said
 Then, “Go,” they would say.
11. Entonces ma n-qo-x-tz
 so.then REC PROG-we-there-then
 So we would go then.
12. “Ku txi'y qanil lisens tzma ja tuk'a manb'aj,
 IMP go ask license to house with father
 “Go to her house and ask permission of her father.”
13. qa kub' meje ch'in twutz manb'aj ex-qe txub'aj.”
 that down kneel a.little before father and-each mother
 That you go a kneel a little (in supplication) before her father and mother and
 her relatives.

Appendix 2. HL, school teacher, 26 years old, taped in 1992.

1. Jun qlixje, o xi'-y toj jun q-b'e-y k'ayil.
 one morning we went-EX in a our-trip-EX to sell
 One morning we went on a trip to sell (at a market).
2. Oxa q-b'aj-a; o xi'-y tuk'a t-karr Josué.
 three our-number-EX we went-EX with his-car Josué
 There were three of us that went in Josué’s car.
3. Atzi'n q-b'aj-a, a Julián ex-sin ayi'n-tz-a.
 that our-number-EX that Julián and-then me-then-EX
 So our number included (Josué), Julián and me.
4. Atzi'n te q-xi'-y, tb'anil-x ch'in n-b'e
 that when we-went-EX nice-AUG a.little my-road
 te q-kanin Twi'muj.
 when we-arrived Twimuj
 Now when we went, the road that took us as far as Twimuj was pretty nice.
5. Noq-tzin tu'n-tz-jo te q-xi' ch'il-tz-a
 only-then regarding-then-SPEC when we-went little-pues-EX
 tu'n q-kanin Triunfo, nya-xix wen b'e
 that we-arrived Triunfo not-AUG good road
 Regarding this (trip) when we went (from Twimuj) to Triunfo, the road was
 not very good.

6. Ma nin-x jul, ex-sin manyor quq-x-tz.
quite big-AUG holes and-also much dust-AUG-then
There were big holes and a lot of dust.
7. Te q-xi'-y Txolja, o-taq tz'ok wajxaq
when we-went-EX Txolja PERF-PAST entered eight
tajlal te qlixje,
its count of morning
When we left Txolja (Comitancillo), it was eight in the morning.
8. Ex q-kanin o-taq tz'ok lajaj te qlixje.
and we-arrived PERF-PAST entered ten of morning
And we arrived (at Triunfo) at ten in the morning.
9. Ponix-jo tal najb'il, ja' o qanin-tz-a.
nice-SPEC little place where we arrived-then-EX
The place where we went was very nice.
10. A-tzin t-xilin ulne iky-jo tze'n-ku
that its-essence coming similar-SPEC like-down
b'e n-tzaj xkye tzma-x Twi' Chlub'.
road PROG-come begin at-there Tuichilupe.
The way we went was similar to the road that comes toward us from
Tuichilupe.
11. Atzi'n te q-kanin-tz-a,
that when we-arrived-then-EX
ma nin-x xjal n-k'ayin-taq Triunfo.
very big-AUG people PROG-sell-past Triunfo
When we arrived, a lot of people were selling there in Triunfo.
12. N-we' karr ja'lin,
PROG-stop car now
Now the car stops,
13. ex b'e'x xi' te Julián k'ayil.
and PUNCT went he Julián to sell
and Julián went off to sell.
14. Ex iky-x-jo qe, oxe qe k'ayil
and similar-AUG-SPEC us three us seller
ti'j k'axhjl-a iqin-taq.
regarding merchandise-EX carry-past
And we as well, (being as) the three of us were salesmen of the stuff that had
been brought.

15. Ex tib'aj jun tal netz' mexh-jo o k'ayini'-y toj plas
and on a cute little table-SPEC we sold-ex in plaza
We sold (our things) on a little table at the plaza.

Appendix 3. FL, 3rd grade education, 23 years old, taped 1980.

1. W-aj-a tu'n n-yolin ch'in ti'j n-ja'-y.
I-want-EX that I-speak little about my-house-EX
I want to tell you a little bit about my house.
2. En primer lugar, at kab'a n-ja'-y.
in the place exist two my-house-EX
First of all, there are two constructions.
3. Por supuesto, nya' we-ku'-y n-junal-a
of course not my-personal-EX my-only-EX
Of course, it's not mine only.
4. Sino, casi antza intin-k-xi'y toj ky-ja,
rather almost there I.LOC-put-EX in their-house
Rather, I am there in their house;
5. Toj ja ite' n-tat-iy ex n-nan-iy
in house exist my-father-EX and my-mother-EX
ex también jni w-it'zin.
and also all my-younger.siblings
in the house are my father, my mother and all my younger siblings.
6. Entonces pues, atzin q-ja'y pues, kab'a c-b'aj;
so then as-for our-house-EX then two their-number
So then, we have two houses.
7. Jun repeyar-in t-wutz o repeyar-in pared te
one plastered-NONFUT its-face or plastered-NONFUT wall for
One has its surfaces plastered, or its walls,
8. ex atzin jun mina.
and that one no
and one doesn't.
9. Atzin jun n-okin te k'u'b'l
that one PROG-serve for storage
Now one of them is for storage,

10. Ex atzin jun-tl n-okin te cocina,
and that one-other PROG-serve for kitchen
and the other one is a kitchen
11. o quiere decir ja' n-b'ant-e wab'j.
or wants to.say where PROG-made-DUR meal
or which means, where the food is prepared.
12. Entonces ex at jun q-chuj-a, ja' n-qo chujin-i'y.
so.then and exist one our-sweat.bath-EX where PROG-we bathe-EX
Now then, we have a sweat bath where we bathe.
13. Ex también atzin toj twi' q-jay kykab'il xk'o'n
and also that in roof our-house both tile
toj ky-wi' tok-x
in their-roof affix-there
And also the roofs of both houses have tile installed.
14. O quiere decir que a' xk'o'n, a b'inchin tu'n
or wants to.say that that tile it made by.means.of
tzaqb'aj tx'otx'
clay earth
In other words, the tile is made from clay.
15. Ex at jun q-pila twi'.pe'n,
and there.is one our-sink patio
And we have a sink in the patio,
16. O sea, “agua potable” tb'i twi' pe'n.
or as it were “water potable” its.name patio
Or as it were, (we have) what is called “potable water” (in Spanish) in the patio.
17. Ex atzin q-tx'otx'-a nya' nim, b'alaqa quince ech.
and that our-land-EX not big maybe fifteen cuerdas
And our land isn't extensive, perhaps about 15 cuerdas (about 1 1/2 acres)

Appendix 4. FL, supervisor of teachers, 45 years old, taped 2002

1. Atzin n-ja'-y, ayin Filiberto López
that my-house-EX I Filiberto López
As for my house, I, filiberto López,

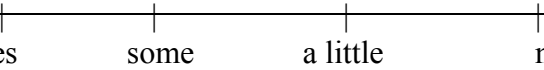
2. Atzin n-ja'-y ate ta' toj tnam.
 that my-house-EX LOC is in town
 My house is in town,
3. Tzalu'n toj tnam te Txolja te tnam Chman
 here in town of Txolja in town Grandfather
 here in the town of Comitancillo in the department of San Marcos
4. Atzin ila'y-x ab'q'e n-kub' n-te'n toj n-ja'-y
 that various-AUG year PROG-down I-LOC in my-house-EX
 It has been several years since I've been in my house.
5. Ex ila'-ku n-k'wal-a ja'lin;
 and various-dispersed my-child-EX now
 and I have a number of kids now;
- at qaq ky-b'aj
 there.are six their-number
 there are six in all.
6. Ex atzin q-ja'y nya'-xix t-b'anil, noq-x ch'in
 and that our-house-EX NEG-AUG SG-nice just-AUG little
 And as for our house, it isn't the best, but it's pretty nice
7. Ex ate ta'ye tzalu'n toj tnam.
 and LOC is here in town
 And it's right here in town.
8. Te junjun alumj ite' ja
 for some animals there.are houses
 And there are pens for some of the animals.
9. At jun ky-ja eky'.
 there.is one their-house chicken
 And there's a chicken house.
10. At chujb'il
 there.is sweat.bath.place
 And there's a seat bath.
11. ex k-ajwil q-ij noq toj maq'maj a'
 and FUT-serve us-to just in warm water
 and also a (sanitary) service with warm water

12. mo toj xb'ajin a'
or in tepid water
or we could say tepid water.
13. Atzin maq'maj a' noq b'inch-it tuk'a q'ij maq'-te
that hot water just make-PAS with sun heat-it
Now the hot water is heated by the sun.
14. Atzin ja, ja' n-b'ant-e wab'j nya-x q'ilnin
that house where PROG-made-HAB food NEG-AUG expensive
Now the kitchen isn't ornate,
15. qu'n tu'n nim ch'in pwaq t-aj
because since big little money 3SG-want
tu'n t-b'ant jun ja nim
for 3SG.make one house big
because it takes quite a bit of money to build a large house,
16. ex-si'n tu'n t-nim-ix te jun-tl.
and-also to 3SG-big-CAUS it one-other
or even to add on to one that's already built
17. atzin tx'otx' nya'-xix nim t-elnin
that land NEG-AUG big 3SG-dimension
Now as for our land, it isn't extensive.

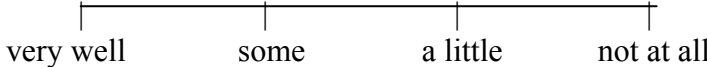
Appendix 5. English version of questionnaire.

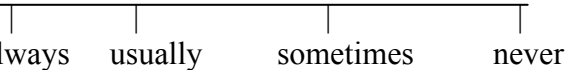
Date: Profession: Religion:

Age: Sex: Birthplace:

1. Do you read and write Mam? 

2. How well do you consider that you speak Spanish?



3. Do you always understand the radio? 

4. What language(s) do you speak at home? _____ in the street? _____ at school? _____
with your siblings? _____ with your parents? _____ at City Hall? _____
with your best friends? _____

5. If you are talking with a friend, what factors determine whether you speak Mam or Spanish?

6. Are you ever embarrassed to speak Mam in front of Latins?

7. If a single Latin comes up to you and your friends who are having a discussion in Mam, would you speak Spanish or Mam?

8. Put a number over each language according to its importance for you personally.

Spanish Mam English Quiché French German Chinese Portuguese

9. In order to be successful academically, you've had to learn Spanish. Do you think that Latins should be required to learn a Mayan language as a requirement for university graduation?

10. Do you think Mam language and culture will prosper or diminish? Why?

11. Do you know any Latins that speak Mam? Do they speak well?

Why did they learn?

12. Do you think it's good or not for Latins to learn to speak Mam?

13. What do we need to promote the use and prestige of Mam?

14. Some futurists say that within 100 years there will only be a few dozen world languages left over from the 6,000 spoken today around the world. Do you think that Mam will be lost? Why?

15. If Mam were to disappear, would you consider this something positive or negative? Why?

16. Mixing Mam and Spanish in a single conversation is good or not?

fine neither good nor bad bad

Why?

17. What kind of people mix the two languages?

18. Tell if you agree or not to the following possible reasons as to why someone might mix Spanish and Mam.

The words don't exist in Mam for a Spanish concept (plane, skyscraper, etc.).

To show-off before others that perhaps don't speak as well.

"Mixers" appreciate both languages and want to speak them both together.

"Mixers" believe that Spanish is more prestigious than Mam and they want take advantage of the perceived benefits of Spanish.

That's the way they learned to talk. It isn't their fault for speaking the way their parents taught them to speak.

CS is actually updated Mam. This is the modern way to speak Mam and it shouldn't be considered bad.

"Mixers" are lazy and don't want to go through the work of investigating how to speak correctly.

Other reasons?

19. Is it worth while to come up with dictionaries of *neologismos*?

|-----|-----|
yes whatever no

Why?

20. Do you think people will use these *neologismos*?

Why?

Other comments.

References

- AUER, PETER. 1995. The pragmatics of code-switching: A sequential approach. *One speaker, two languages*, ed. by Lesley Milroy and Pieter Muysken, 115–35. Cambridge: Cambridge University Press.
- BLOM, JAN PETER, and JOHN GUMPERZ. 1972. Social meaning in linguistic structures: Code-switching in Norway. *Directions in sociolinguistics*, ed. by John J. Gumperz and Dell Hymes, 407–35. New York: Holt, Rinehart & Winston.
- BLOMMAERT, JAN. 1992. Codeswitching and the exclusivity of social identities: Some data from Campus Kiswahili. *Codeswitching*, ed. by Carol M. Eastman, 57–70. Clevedon: Multilingual Matters.
- BURGOS, ELIZABETH. 1985. *Me llamo Rigoberta Menchú y así me nació la conciencia*. Mexico City: Siglo Veintiuno Editores.
- CLÉMENT, R. 1979. Ethnicity, contact and communicative competence in a second language. *Language: Social psychological perspectives*, ed. by Howard Giles, W. Peter Robinson, and Philip M. Smith, 147–54. Oxford: Pergamon Press.
- COULMAS, FLORIAN (ed.) 1997. *The handbook of sociolinguistics*. Oxford: Blackwell.
- CRAWFORD, JAMES. 1995. Endangered native American languages: What is to be done, and why? *Bilingual Research Journal* 19.17–38.
- DE LANDA, DIEGO. 1566. *Relación de las cosas de Yucatán*, trans. by William Gates. New York: Dover Publications.
- ECKERT, PENELOPE. 2000. *Linguistic variation as social practice*. Oxford: Blackwell.
- ENGLAND, NORA C. 1998. Mayan efforts toward language preservation. In Grenoble & Whaley, 99–116.
- FISHMAN, JOSHUA A. 1991. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon: Multilingual Matters.
- FISHMAN, JOSHUA A. (ed.) 2001. *Can threatened languages be saved?* Clevedon: Multilingual Matters.
- FRASER, COLIN, and KLAUS R. SCHERER (eds.) 1982. *Advances in the social psychology of language*. Cambridge: Cambridge University Press.
- GALLOIS, C., and V. J. CALLAN. 1988. Communication accommodation and the prototypical speaker: Predicting evaluations and status solidarity. *Language and Communication* 8.271–84.

- GILES, HOWARD, and NIKOLAS COUPLAND. 1991. *Language: Contexts and consequences*. Pacific Grove, CA: Brooks/Cole Publishing Co.
- GILES, HOWARD, and PATRICIA JOHNSON. 1981. The role of language in ethnic group relations. *Intergroup behaviour*, ed. by John C. Turner and Howard Giles, 199–272. Oxford: Blackwell.
- GILES, HOWARD, and PATRICIA JOHNSON. 1987. Ethnolinguistic identity theory: A social psychological approach to language maintenance. *International Journal of the Sociology of Language* 68.69–99.
- GILES, HOWARD, and PETER F. POWESLAND. 1975. A social psychological model of speech diversity. *Speech style and social evaluation*, ed. by Howard Giles and Peter F. Powesland, 154–70. New York: Harcourt Brace.
- GODFREY, THOMAS JAMES, and WESLEY M. COLLINS. 1987. *Una encuesta dialectal en el area mam de Guatemala*. Guatemala City: Instituto Lingüístico de Verano.
- GRENOBLE, LENORE A., and LINDSAY J. WHALEY (eds.) 1998. *Endangered languages*. Cambridge: Cambridge University Press.
- GUMPERZ JOHN J. (ed.) 1982. *Language and social identity*. Cambridge: Cambridge University Press.
- HALE, KEN. 1992. Endangered languages. *Language* 68.1–42.
- HALE, KEN. 1998. On endangered languages and the importance of linguistic diversity. In Grenoble & Whaley, 192–216.
- HARTMAN KEISER, STEVE. 2003. Pennsylvania German and the “lunch pail threat”: Language shift and cultural maintenance in two Amish communities. In Joseph et al., 3–20.
- HAUGEN, EINAR. 1950. The analysis of linguistic borrowing. *Language* 26.210–31.
- HELLER, MONICA. 1988a. Strategic ambiguity: Codeswitching in the management of conflict. In Heller 1988b, 77–96.
- HELLER, MONICA (ed.) 1988b. *Codeswitching: Anthropological and sociolinguistic perspectives*. Berlin: Mouton de Gruyter.
- HELLER, MONICA. 1992. The politics of codeswitching and language choice. *Codeswitching*, ed. by Carol Eastman, 123–42. Clevedon: Multilingual Matters.
- HELLER, MONICA. 1995. Language choice, social institutions, and symbolic domination. *Language in Society* 24.373–405.

- HILL, JANE H., and KENNETH C. HILL. 1986. *Speaking Mexicano*. Tucson: University of Arizona Press.
- JOSEPH, BRIAN D.; JOHANNA DE STEFANO; NEIL G. JACOBS; and ILSE LEHISTE (eds.) 2003. *When languages collide: Perspectives on language conflict, language competition, and language coexistence*. Columbus: The Ohio State University Press.
- KNOWLES-BERRY, SUSAN. 1987. Language decay in Chontal Mayan: The speech of semi-speakers. *Anthropological Linguistics* 29.332–41.
- KRAUSS, MICHAEL. 1992. The world's languages in crisis. *Language* 68.4–10.
- LADEFOGED, PETER. 1992. Another view of endangered languages. *Language* 68.809–11.
- LASTRA, Y. Otomí language shift and some recent efforts to reverse it. In Fishman 2001, 142–65.
- LEE, TIFFANY, and DANIEL MCLAUGHLIN. 2001. Reversing Navajo language shift, revisited. In Fishman 2001, 166–94.
- LE PAGE, ROBERT B., and ANDRÉE TABOURET-KELLER. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
- MARTÍNEZ PELÁEZ, SEVERO. 1970. *La patria del criollo*. Mexico City: Ediciones en Marcha.
- MCCLURE, ERICA, and MALCOLM MCCLURE. 1988. Macro- and micro-sociolinguistic dimensions of code-switching in Vingard (Romania). In Heller 1988b, 25–51.
- MILROY, LESLIE, and PIETER MUYSKEN (eds.) 1995. *One speaker, two languages*. Cambridge: Cambridge University Press.
- MUFWENE, SALIKOKO S. 2001. *The ecology of language evolution*. Cambridge: Cambridge University Press.
- MUFWENE, SALIKOKO S. 2002. Colonization, globalization and the plight of “weak” languages. *Journal of Linguistics* 38.375–95.
- MUFWENE, SALIKOKO S. 2003. Language endangerment: What have pride and prestige got to do with it? In Joseph et al., 324–45.
- MYERS-SCOTTON, CAROL. 1992. Comparing codeswitching and borrowing. *Code-switching*, ed. by Carol Eastman, 19–39. Clevedon: Multilingual Matters.
- MYERS-SCOTTON, CAROL. 1993. Common and uncommon ground: Social and structural factors in codeswitching. *Language in Society* 22.475–503.

- MYERS-SCOTTON, CAROL. 1997. Code-switching. *The handbook of sociolinguistics*, ed. by Florian Coulmas, 217–37. Oxford: Blackwell.
- NETTLE, DANIEL, and SUZANNE ROMAINE. 2000. *Vanishing voices: The extinction of the world's languages*. Oxford: Oxford University Press.
- Ó RIAGÁIN, PÁDRAIG. 2001. Irish language production and reproduction 1981–1996. In Fishman 2001, 195–214.
- POPLACK, SHANA. 1980. Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics* 18.581–618.
- POPLACK, SHANA. 1988. Contrasting patterns of codeswitching in two communities. In Heller 1988b, 215–44.
- REDFIELD, ROBERT. 1941. *The folk culture of Yucatan*. Chicago: University of Chicago Press.
- ROBINSON, W. PETER, and HOWARD GILES (eds.) *The new handbook of language and social psychology*. Somerset, NJ: John Wiley and Sons.
- ROMAINE, SUZANNE. 1994. *Language in society*. Oxford: Oxford University Press.
- SASSE, HANS-JÜRGEN. 1992. Theory of language death. *Language death: Factual and theoretical explorations with specific reference to East Africa*, ed. by Matthias Brenzinger, 7–30. Berlin: Mouton de Gruyter.
- SHEPHARD, CAROLYN A.; HOWARD GILES; and BETH A. LE POIRE. 2001. Communication Accommodation Theory. In Robinson & Giles, 33–56.
- TAJFEL, H. 1974. Social identity and intergroup behaviour. *Social Science Information* 13.65–93.
- TAJFEL, H., and J. C. TURNER. 1979. An integrative theory of intergroup conflict. *The social psychology of intergroup relations*, ed. by W. G. Austin and S. Worchel, 33–47. Monterey, CA: Brooks/Cole.
- THAKERER, JITENDRA N.; HOWARD GILES; and JENNY CHESHIRE. 1982. Psychological and linguistic parameters of speech accommodation theory. In Fraser & Scherer, 205–55.
- THOMASON, SARAH G. 2001. *Language contact*. Washington, DC: Georgetown University Press.
- WEINREICH, URIEL. 1953. *Languages in contact*. The Hague: Mouton.
- WINFORD, DONALD. 2003. *An introduction to contact linguistics*. Oxford: Blackwell.
- WOOLARD, KATHRYN. 1988. Codeswitching and comedy in Catalonia. In Heller 1988b, 53–76. Berlin: Mouton de Gruyter.

DEFINING THE OUTCOME OF LANGUAGE CONTACT: OLD ENGLISH AND OLD NORSE

Hope C. Dawson

Abstract

The English language throughout its 1500 year history has been impacted by socio-historical developments and changes. One such development took place in Old English: the invasion of England by Norse tribes from c. 800-1000 A.D. was a series of events which had a significant and lasting impact on all areas of the English language. The nature of that social situation and the linguistic outcome is of interest in contact linguistics; in particular, the application by some of terms such as *creolization* and *creole* to this process and its outcome has been controversial. In this paper, I examine the English-Norse contact situation and its effects on English and propose that the linguistic outcome of this contact was a koine, and show that this account can better describe the effects of this contact situation on the English language.

1 Socio-historical background

A series of Norse invasions of England from c. 800–1000 A.D. resulted in language contact between Old English (OE) and Old Norse (ON).¹ These invasions can be

¹ The term “Norse” in this paper refers generally to the people groups which inhabited the Scandinavian peninsula and Denmark and which were involved in the raids on the British Isles. Distinctions made
HOPE C. DAWSON. Defining the outcome of language contact: Old English and Old Norse.
OSUWPL 57, Summer 2003, 40–57.
Copyright © 2003 The Ohio State University

divided into three periods, each of which differed in nature and thus in impact on the language (Baugh & Cable 1993).

1.1 First period: 787 to c.850 A.D.

A period of early raids began in 787 A.D., as recounted in the Anglo-Saxon Chronicle, and continued with some intermissions until c. 850. These were simply plundering attacks on towns and monasteries near the coast, including the noted sackings of the monasteries at Lindisfarne and Jarrow in 793 and 794. A forty year respite from the attacks followed these sackings, but this ended in 834 with renewed attacks along the southern coast of England, in East Anglia.² The attacks in this period were apparently the work of small isolated bands. The size of the invading force and the swiftness of the raids indicate that no significant language contact took place in this period.

1.2 Second period: 850 to 878 A.D.

Much more widespread plundering by large armies marked the second period, and this resulted in extensive settlements and the establishment of the Danelaw and Norse institutions in parts of England. It began in 850 with the arrival on English soil of a fleet of 350 Norse ships; the Norse spent the winter on the island Thanet, which is on the Kentish lip of the Thames estuary, and moved upriver in the spring, capturing Canterbury and London and ravaging the surrounding countryside. Nonetheless, the Norse showed no clear-cut attempt at permanent colonization for several years, being “concerned with loot and sporadic raids rather than systematic probing of defenses with a view to stable settlement” (Loyn 1977:56).

This changed in 865 when a great Norse army arrived in East Anglia; they plundered the area in 866, and captured York, the capital of Northumbria, in 867. The invaders then turned south to Mercia, and again attacked East Anglia, London, and Wessex. They established a base in Bernicia in northernmost Northumbria and set up an overlordship in the Tyne region. The army’s fifteen years of fighting in England culminated with “the colonisation by the Danes of extensive tracts of northern and eastern England and, consequently, in the first implanting on English soil of the Norse language” (Geipel 1971:40). Many of the Norse attackers remained in Northumbria, making a home for themselves and “the first permanent settlement of Danes in England” (Geipel 1971:41).

These attacks had left the eastern part of England largely in the hands of the Norse, but King Alfred (871–899) took the throne of Wessex and resisted the Norse rule. After seven years, he led his people to victory over them at Ethandun in 878, and the Treaty of Wedmore was signed by the English and Norse in 886. The Norse swore by the treaty to leave Wessex alone and to “confine activities to areas east of Watling Street and

between “Norwegians” and “Danish” by some of the authors quoted here do not correlate to the modern-day usage of these terms.

² See the map in the Appendix.

north of the Thames” (Geipel 1971:42), or east of a line from Chester to London,³ the area delimited by this treaty became known as the Danelaw. Distinctively Scandinavian institutions were established in the North and in the East Midlands, and Norse legal and monetary systems eventually replaced their English counterparts in the Danelaw. The system of land measurements and administrative districts and their governance was also replaced by agrarian settlements and a retention of military organization, as “[I]and settlement and the introduction of immigrants were achieved under the discipline of armies which maintained fortified headquarters at Northampton, Cambridge, Tempsford, Thetford and Huntingdon” (Loyn 1977:60). The Norse presence was particularly strong in the North, where “York dominated the whole area, rapidly developing into a powerful Scandinavian fortified market ... [and] emerged as the political heart of a vigorous colonising movement in Northumbria” (Loyn 1977:60). The large numbers of invasions and settlements and the establishment of a permanent and influential Norse presence in the northern and eastern parts of England in this period were significant for language contact.

1.3 Third period: 878 to 1042 A.D.

Political adjustment and assimilation marked the third period. Two large Norse fleets landed in Kent in 892; from there, the invaders struck inland towards Wessex, and they were joined by many of the Norse who were already living in England. King Alfred, who had remained watchful of the Norse after the Treaty of Wedmore, renewed the fight against them, finally prevailing after four years in the summer of 896. The Norse dispersed to Northumbria, East Anglia, and Normandy, where they continued to be put on the defensive under Alfred’s successors, the Wessex kings Edward the Elder (900–925) and Athelstan (925–939). When a powerful force of Vikings arrived in Yorkshire, the Norse living in England “now stood to suffer as much from any further Viking irruptions as did their Anglian neighbors”, and “the inhabitants of eastern England, Angles and Danes alike, [took] up their weapons and rall[ied] to King Aethelstan’s side” (Geipel 1971:47). Nevertheless, the Vikings captured York and ruled for some years, but the English gradually reclaimed much of the land of central and east England, including all of Northumbria, which had been under Norse control.

Almost all of England was again under English control by the middle of the tenth century, but Norse influence was still strong in the northern and eastern areas. In the re-taking of Norse lands, “[t]he colonists were nowhere extirpated, they seem to have offered scant resistance to the reclamation of their lands, and their absorption into the fabric of the English nation appears to have taken place without undue violence” (Geipel 1971:47). While maintaining some aspects of their cultural identity, the “Danish farmers, settled and often Christianised, came to realise that their best hope of peaceful future lay in acceptance of the overlordship of the West Saxon dynasty” (Loyn 1977:63).

A series of new invasions began in 991, however, when Viking fleets attacked the southern coasts of Wessex from Dorset to Cornwall. The Norse made their way north towards York, and they were joined by many second- and third-generation Norse inhabitants on their way through the Danelaw. The Wessex King Athelred (978–1016) was

³ See the map in the Appendix.

angered by this betrayal and ordered the killing of all foreigners outside the Danelaw. In retaliation, the Norse King Sveinn led a great fleet of warships to East Anglia in the spring of 1007, and the invaders swept inland from East Anglia. Reinforcements arrived from Denmark in 1009, and Sveinn stepped up the attacks, invading Northumbria and scorching Oxford. Aethelred abandoned the throne and fled to Normandy in 1013, and having taken the capital Winchester, Sveinn seized the throne of Wessex in 1014. He died shortly thereafter, however, and Aethelred returned briefly “to deal, with malicious brutality, with the vociferous pro-Danish element in the north and east” who wanted Sveinn’s son Canute to return from Denmark to claim the throne (Geipel 1971:50).

Canute landed with a fleet in 1015, and “in a matter of months, the whole country, save for London, was in Danish hands” (Geipel 1971:50). After the death of Aethelred in 1016 and his successor Edmund shortly thereafter, Canute was proclaimed king of all England; he ruled over an empire consisting of England, Denmark, and Norway by 1028. Many of his followers “elected to remain on English soil, becoming, as had their predecessors, farmers, landowners and traders—not merely in the Danelaw, but also further to the south and west” (Geipel 1971:51). This Anglo-Norse state ended, however, with William of Normandy’s conquest of England in 1066, and Norse resistance to William led to the “Harrying of the North, in which large areas were depopulated and scorched black; ... placename evidence suggests that much of the northern Danelaw was eventually repopulated by settlers of mixed Scandinavian/Irish parentage ... [and there were] no further attempts by the Danes to reestablish the lost portions of the [Danelaw]” (Geipel 1971:51–52). The substantial numbers of Norse who settled within the northern and eastern parts of England during this period and their shifting political loyalties and cultural integration resulted in significant contact between English and Norse speakers.

2 The impact on the English language

This situation of extended language contact between English and Norse had considerable impact on all aspects of the English language, particularly those language varieties which were spoken in the northern and eastern areas of England. Some of the effects were lost, but many survived, and features of this language variety were later diffused into the dialects which would become the foundation of “Standard English”, so that many of these effects can be seen in Modern English.

2.1 The lexicon

Norse lexical influence on English is still readily apparent in the Modern English lexicon, even though some lexical effects which were found in the Northern ME dialects were subsequently lost. Many of the lexical items which show the influence of Norse on English are “new” words (i.e. ones for which there was no OE parallel), such as *steak* < ON *steik*; *reindeer* < ON *hreindyri*; *snare* < ON *snara*; *sprint* < ON *spretta*; and *flat* < ON *flatr*, all of which are of Norse origin. In other cases, the Norse word replaced an OE word; for example, *window* < ON *vindauga* “‘wind eye’”; *window*’ took the place of OE *eyethurl* “‘eye hole’”; *window*’; *take* < ON *taka* replaced the OE *niman*; and *sky* < ON *ský* replaced the OE *ūprodor* and *wolcen*.

Other cases are not so clear-cut; because the English and the Norse language varieties were similar in many aspects and had many roots in common, the contact also resulted in more subtle influences on the English lexicon. The origin of particular words for which there was an ON and OE common root can be determined by knowing the outcome of certain phonological developments that distinguish the two language varieties. The cluster **-sk-*, for example, had been palatalized in OE to [ʃ] (orthographically <*sc*>) but remained [sk] in ON. These separate developments indicate, therefore, that words such as *sky* < ON *sky'*, *skin* < ON *skinn*, and *whisk* < ON *visk* are of Norse origin, and words such as *shall* < OE *sceal* (cf. ON. *skal*) and *fish* < OE *fisc* (cf. ON *fiskr*) are of English origin. This development also gave rise to Modern English word pairs as *shirt* (from OE *scyrte*) and *skirt* (from ON *skyrta*), where distinctive semantics now distinguish two words which are etymologically the same. Other word pairs of this type include *no – nay* (ON *nei*), *whole – hale* (OE *hál*, ON *heill*), and *rear – raise* (OE *ræran*, ON. *reisa*).⁴ Similarly, the differential development of OE and ON [k] and [g] in certain contexts reveals that *egg*, *kid*, *get*, and *give* owe their current phonetic shape to the Norse influence; the OE pronunciations *eyren* ‘eggs’⁵ and *jefa* ‘give’ were eventually replaced in standard English.

Other Norse influences can be found in the semantics of lexical items, an effect which is particularly salient in those cases where the phonetic shape could be derived directly from either OE or ON. Modern English *bloom* (flower), for example, could represent the normal development of either OE *blōma* or ON *blōm*, but its OE meaning of ‘ingot of iron’ leads to the conclusion that its modern use must have been influenced by the ‘flower, bloom’ meaning of the ON cognate. In other examples, both phonetic and semantic influence can be seen; for example, the modern word *gift* indicates Norse influence in its phonetics, with the initial [g] contrasting with the OE cognate’s initial [j], and in its semantics, where the meaning reflects ON ‘gift, present’ rather than the OE cognate ‘payment for a wife’. Finally, this lexical influence resulted in the development of “compromise forms” which cannot be traced directly to either OE or ON exclusively; for example, the ME *werse* ‘worse’ shows influence from both ON *werre* and OE *wyrsa*, and the ME *whaare* ‘where’ and *thaare* ‘there’ were influenced by both the ON *hwar* and *θar* and the OE *hwēr*, *θēr*.

2.2 Morphology

English morphology also reflects Norse influence, both in its derivational and inflectional affixes and in its function words. The phonetics of the ME derivational prefix *umbe-* ‘around’ indicate influence from the ON *umb-* rather than the normal development of OE *ymbe-*; similarly, the ME suffix *-leik* ‘-ness’ reflects ON *-leik-r* rather than OE

⁴ Note that while *rear* originally was mostly synonymous with *raise*, it has become quite limited in its usage; as a child, I was taught the semantic distinction of *rear a (human) child* vs. *raise cattle (sheep, etc.)*, but this has largely been lost in favor of the use of *raise* in both senses.

⁵ As commented on by William Caxton in the preface to his English translation/paraphrase of the *Aeneid*.

–*lāk*. In the inflectional domain, the ME suffix of the present participle –*ande* can be compared with the ON –*ande* versus the OE –*ende*.⁶

Independent morphemes (or function words) also reflect Norse influence. The ME preposition *til* ‘to’ < ON *til* is still found in the Modern English *till*, which exists alongside the more common *to* from OE *tō*. Similarly, while Modern English *from* developed from OE *fram/from*, the more limited form *fro*, as in the phrase *to and fro*, reflects ME *fraa/froa* ‘from’ from ON *frā*. Perhaps the most significant and lasting area of Norse influence on English is seen in the personal pronoun system, in the third person pronouns *they*, *them*, and *their*. These forms clearly demonstrate the influence of the Norse. In particular, the initial [ð] of the Modern English forms can be traced to Norse; compare *they* with the ON *þei-r* and OE *hīe*, *hēo*, and ME *theim* ‘them’ and *theire* ‘their’ with the ON *þei-m* and *þei-ra* and the Northumbrian OE *him* and *hira* or Mercian OE *heom* and *heora*.

2.3 Morphosyntax

The ME of Northern England, and later of more geographically wide-spread varieties of English, is marked in comparison to OE by a fairly dramatic shift in the morphosyntax from a highly synthetic system to one more analytic. While changes in the inflectional system were underway before the Norse contact occurred and can be attributed to factors such as phonological change (e.g., a reduction of unstressed vowels, loss of word-final consonants), the overall impact of these changes was accelerated in the areas in which Norse-English contact took place.

Morphosyntactic features of OE c.850 A.D. included a noun system which had three basic noun classes, the strong masculine, strong feminine, and weak nouns; these were inflected for singular and plural number, and nominative, accusative, genitive, and dative case. Adjectives were indefinite or definite, and were inflected for singular and plural number; masculine, feminine, and neuter gender; and nominative, accusative, genitive, dative, and instrumental case. Demonstratives were similarly inflected for case and number, and for gender in the singular. The verbal system inflected for number and for three persons, as well as tense, voice, and mood. The ME morphosyntactic system, in contrast, bore greater similarity to that of Modern English, with nouns being marked only for singular or plural, and a genitive case marking in the singular; adjectives were no longer inflected, the demonstrative had been reduced to a single form *the*, and verbs distinguished only the third person singular in the present tense. While these changes cannot be directly attributed to Norse influence in the same way that morpho-lexical effects can be, the correlation between the acceleration of these changes and the geographical location of the Norse settlements leads to the conclusion that English-Norse contact played a role.

⁶ These elements and others in which the Norse influence is apparent are clearly and thoroughly documented by Thomason and Kaufman (1988:293–95).

3 Analysis

The contact between English and Norse had a lasting impact on the English language, as has been shown. What can be said about this situation from a theoretical standpoint, and particularly within the field of language contact studies? An analysis that can account for the linguistic effects of the contact situation, while fitting the socio-historical situation appropriately, is needed.

3.1 Koines

The term “koine”, while used for many years, has not always been well-defined linguistically. The definition of koine that I adopt here is that given by Siegel (1985), and it can be broken down into three sections. First, a koine is defined by the language contact situation in which it developed: It is “the stabilized result of mixing of linguistic subsystems such as regional or literary dialects” (Siegel 1985:363). Second, a koine is defined in terms of how it was used socio-historically: “It usually serves as a lingua franca among speakers of the different contributing varieties” (ibid.). Third, it is defined by linguistic characteristics with respect to the language varieties from which it developed: It “is characterized by a mixture of features of these varieties and most often by reduction or simplification in comparison” (ibid.).

Can these three aspects of the definition of koines be applied to the result of the English-Norse situation? The first criterion is that the contact situation involve the “mixing of linguistic subsystems”, which Siegel goes on to define by saying that:

Two or more different linguistic varieties may be considered subsystems of the same linguistic system if they are genetically closely related and thus typologically similar enough to fulfill at least one of two criteria (1) they are mutually intelligible (2) they share a superposed, genetically related linguistic system, such as a national standard or literary language. (Siegel 1985:365)

The English and Norse language varieties involved here were genetically closely related and are generally believed to have been mutually intelligible. For example, Thomason and Kaufman state that Norse and English at the time of their contact were structurally and lexically close enough that “it was relatively easy to understand the other language without learning to speak it” although “one could never be in doubt which language was being spoken” (1988:303).

The criterion of usage as a “lingua franca” among the speakers of English and Norse is difficult to prove definitively because of a lack of direct evidence. The facts that are known about the socio-historical situation, however, support such a scenario. The third criterion of a koine containing a “mixture of features” and being characterized by “reduction or simplification” in comparison to the varieties from which it developed closely parallels the features of the variety of English that developed from that spoken in northern England during the time of contact, such as the reduction and simplification of

the nominal and verbal inflectional systems, the development of “compromise forms”, and the influence of Norse on personal pronouns; see §§2.2-2.3.

3.2 Koineization

The facts of koineization, or the process of koine formation, lend further support to naming the outcome of this contact situation a koine. Siegel’s definition of koineization builds on the socio-historical criteria given above (§3.1). Koineization is “a gradual process which occurs only after prolonged contact between speakers who can most often understand each other to some extent” (Siegel 1985:372). Here the necessity of some mutual intelligibility is reiterated, but with an additional focus on the processual aspect of koineization. Koine formation is not an abrupt process, but rather a gradual one, growing out of prolonged contact between speakers. The contact between English and Norse took place over a period of 200 years, from c. 865 to 1066 A.D., with some contact taking place before then and certainly continuing to a certain extent even after the Norman conquest.

The process of koineization can be divided into three stages. The first or “pre-koine stage” is “the unstabilized stage of the beginning of koineization” during which “various forms of the varieties in contact are used concurrently and inconsistently”. At this stage, “[l]evelling and some mixing has begun to occur, and there may be various degrees of reduction, but few forms have emerged as the accepted compromise” (Siegel 1985:373). Similarly, Trudgill (1986:107) says that “there may be an enormous amount of linguistic variability in the early stages” of (dialect) contact situations. Evidence of such a period in which “[t]he Scandinavian and English words were being used side by side” (Baugh & Cable 1993:98) can be seen in the basic nature of many of the Norse-origin words in English. In the English-Norse contact situation, this stage likely occurred in the late ninth or early tenth century, or at the end of the second period and the early part of the third period of contact (§§1.2-1.3), when Norse settlements were being formed and institutions being established in England, and some more or less regular contact was taking place.

The second stage of koineization results in a “stabilized koine”. In this stage, “[l]exical, phonological, and morphological norms have been distilled from the various subsystems in contact, and a new compromise subsystem has emerged”. This stabilized system is “often reduced in morphological complexity compared to the contributing subsystems” (Siegel 1985:373). This stage is one in which “focusing ... takes place by means of a reduction of the forms available” (Trudgill 1986:107), which is the process which Trudgill particularly calls “koinéization”, “which consists of the levelling out of minority and otherwise marked speech forms, and of simplification, which involves, crucially, a reduction in irregularities” (Trudgill 1986:107). This stage in koineization would have occurred in the third period of English-Norse contact, as assimilation and adjustment was taking place socially and politically between the Norse and English.

The third stage is that in which an “expanded koine” may appear, “often accompanied by linguistic expansion, for example, in greater morphological complexity

and stylistic options” (Siegel 1985:373–74). This may happen concurrently with the social expansion of the koine; for instance, it may become the literary language or the standard language of a country. This third stage is where we find “[t]he result of the focusing associated with koinéization [which] is a historically mixed but synchronically stable dialect which contains elements from the different dialects that went into the mixture, as well as interdialectal forms that were present in none” (Trudgill 1986:107–8). This stage likely occurred late in the third period of English-Norse contact and in the following years.

What brings about the formation of koines in general, and how does this apply to the English-Norse contact situation? “The contact status quo may end with certain political, social, economic or demographic changes which cause either increased interaction among speakers of various linguistic subsystems or decreased inclination to maintain linguistic distinctions” (Siegel 1985:366). Norse was certainly spoken in the Danelaw and other Norse settlements for some time, but Thomason and Kaufman believe that it was lost within two generations of an area’s reintegration into English control.⁷ The change of political and social status of Norse may well have caused such a change in the interaction between the speakers as well as in attitudes towards the native language varieties. Kerswill and Williams’s finding in their study on modern-day koine formation that “focusing occurs in either the second or third generation (the children or grandchildren of the migrants)” (2000:71) further confirms that the timeline proposed by Thomason and Kaufman fits the koineization hypothesis.

The English and Norse language varieties were both maintained in areas under Norse control, with perhaps some “bilingualism” in the communities, but with most interactions able to be accomplished using the original, mutually intelligible languages. After areas were returned to English control, the social situation changed, with more interaction between the groups, and with no longer as much impetus to maintain the distinctions between the languages. By the second or third generation of this changed social situation, the children had developed a new, compromise language variety, or koine, and Norse was lost. I could also propose that the dialects of English original to the areas in question were lost as well, as we have seen the differences between the language varieties found in these areas compared to the previous forms of English found there. The fact that these northern dialects were later influential in the formation of London standard English resulted in the spread of many of the “Norse” features from the northern koine into the other dialects of English.

3.4 Outcome of koineization

Linguistically, a koine “is characterized by a mixture of features of these varieties and most often by reduction or simplification in comparison” (Siegel 1985:363). The

⁷ “Norse began to go out of use in any area when the area was reintegrated (through conquest) to the English polity, and was effectively defunct within two generations ... of this reintegration” (Thomason & Kaufman 1988:284). Therefore, Norse began to go out of use earliest in the southern parts of the Danelaw, or the Midlands, from c. 920–980, and was lost in the north from c. 955–1015 (1988:337).

mixture of English and Norse features was seen in the lexical effects of the English-Norse contact (§2.1). Northern Middle English's lexicon had many words of Norse origin, particularly words that were very "basic" in nature. Other cases in which words show the influence of the Norse in their semantics also reveal this mixture. Even more telling are those words which are apparent blends of the original English and Norse words. For example, some English words were

modified, taking on some of the character of the corresponding Scandinavian word. *Give* and *get* with their hard *g* are examples, as are *scatter* beside *shatter*, and *Thursday* instead of the OE *Thunresdæg* ... [and also note the] survival of such hybrid forms as *shriek* and *screech*. (Baugh & Cable 1993:99)

We also saw this in ME *werse* 'worse' from ON *werre*, OE *wyrsa*, as well as ME *whaare* 'where' and *thaare* 'there' from ON *hwar* and *þar*, OE *hwēr*, *þēr* (§2.1; Thomason & Kaufman 1988:294).

The early stage of variation is followed by a period of leveling and focusing, during which "forms that are not removed during koinéization ... will tend to be re-assigned according to certain patterns. ... retained variants may acquire different degrees of formality and be reallocated the function of stylistic variants" (Trudgill 1986:110). This can be seen in word pairs where "[o]ccasionally both the English and Scandinavian words were retained with a difference of meaning or use ... [e.g. the English – Norse pairs] *no – nay*, *whole – hale*, *rear – raise*, *hide – skin*, *sick – ill*" (Baugh & Cable 1993:99).

Reduction or simplification in comparison to the original linguistic varieties is clearly seen in Northern ME. Kerswill and Williams describe "simplification" as referring to "an increase in morphological regularity, an increase in invariable word forms, and a decrease in the number of morphological categories. In addition, 'simplification' covers morphological and lexical transparency" (2000:85). The northern dialects of ME showed an increase in analyticity, with loss of grammatical gender, loss of case markings on nouns, and loss of some verbal inflections, all of which can be explained as the expected outcome of koine formation.

4 Previous analyses

The English-Norse contact situation discussed here has been treated in other language-contact studies, such as in creole and second language acquisition studies. How well can these other analyses account for the facts of this situation and how do they compare with the analysis presented here?

4.1 Borrowing

Thomason and Kaufman (1988) discuss this language contact situation within their framework of degrees of borrowing correlated with the intensity of language con-

tact. They put the English-Norse contact effects “on the borderline of types 2 and 3 of our borrowing scale” (1988:307), and say that the outcome “suggests an intense contact situation, either category (3) borrowing or considerable influence through shift, or (more likely) both” (1988:281). They emphasize the “normal” transmission of Old English to Middle English, concluding that “the available evidence puts ME squarely in the large group of normally transmitted languages, not in the smaller group of mixed languages which (in our view) have no genetic affiliations” (1988:312).⁸ They focus on the identity between OE and ME so that “the Middle English of the Danelaw, in spite of its Norse component, its greater phonological and morphological simplicity, and its other regional peculiarities neither simple nor Norse, is **English**” (1988:280). For this reason, they affirm that “in the contact between Norse and English no case can be made for anything other than rather heavy linguistic borrowing by English from Norse” (1988:310).

Thomason and Kaufman particularly address the issue of the morphosyntactic changes that took place from OE to ME because of the importance that has been attributed to these changes in language contact studies. They emphasize the fact that these changes were already underway in OE before the arrival of the Norse and conclude that while the language contact situation may well have helped or accelerated these changes, they would have taken place anyway. The lexical and morphological influences of Norse are mostly the result of borrowing, in this viewpoint.

Borrowing certainly played a major role in this contact situation, but this process cannot explain all of its effects. The borrowing framework is weak in dealing with intense language contact situations because it is too general. Once “borrowing” can be invoked to account for all manner of changes, it loses its explanatory power and suffers from a lack of limitations. For example, while the presence of such Norse words such as *steak*, *sky*, and *window* in the ME lexicon can unproblematically be explained as the result of borrowing, the explanation of the more subtle effects of Norse influence on the lexicon is more complicated. For instance, in the cases of phonetic and semantic influence, was the whole lexical item borrowed, or merely the phonetics or the semantics? What would lead to the borrowing of words which already had a close cognate in OE, giving word pairs such as *shirt-skirt*? How does borrowing account for compromise forms which show the influence of both languages?

Other questions about the borrowing framework relate to its explanation of the morphosyntactic effects. Why were derivational and inflectional morphemes borrowed? These borrowings are seen as evidence of and therefore explained by fairly intense contact, but this lacks explanatory force. Similarly, the change in the morphosyntactic system does not receive a good explanation in this account. In summary, while borrowing was likely involved here, it does not provide an explanation for all of the aspects of this language contact situation.

⁸ This is in particular a reaction against the creolization hypothesis of Bailey and Maroldt (1977); see §4.2.

4.2 Creolization

Bailey and Maroldt's (1977) proposal that the outcome of the English-Norse contact situation was a creolization of English was the impetus for Thomason and Kaufman's emphasis on borrowing and "normal" language transmission. Bailey and Maroldt define creolization as "a gradient mixture of two or more languages" (1977:21), and they state that "[i]t cannot be doubted that it [Middle English] is a mixed language, or creole" (1977:22). They further define a creole as "the result of mixing which is substantial enough to result in a new system, a system that is separate from its antecedent parent systems" (1977:21), which they base particularly on the criterion that "creoles often have special identifying traits, such as morphological (derivational and inflectional) simplification—or rather 'analyticity' in the morphological sense" (1977:21). While their particular focus is on the contact between "Anglo-Saxon" and Old French, they also attribute the "creolization" of Middle English to the earlier contact between Old Norse and "Anglo-Saxon", so that "the infusion of Old Norse elements led to that sort of linguistic instability which linguistic mixture generally creates, and thus prepared the ground for even more substantial foreign creolization afterwards" (1977:26). The borrowing of such basic concepts or lexico-morphological items as "*die, give, take, are, ... they, their*" from Scandinavian "strongly supports the assumption of an Old Norse/Anglo-Saxon creolization prior to French influence" (1977:27).

This viewpoint was also advocated by Poussa (1982), who focused even more on the role of the Norse contact, whereas Bailey and Maroldt's focus was more on the later contact with the French. She states that "the fundamental changes which took place between standard literary OE and Chancery Standard English: loss of grammatical gender, extreme simplification of inflexions and borrowing of form-words and common lexical words, may be ascribed to a creolization with Old Scandinavian during the OE period" (Poussa 1982:84).

The treatment of the English-Norse contact situation as a case of creolization is problematic in many respects. The notion of "creolization" itself and the definition of a "creole" are not without controversy. The term creole is typically used with languages that meet certain structural and/or socio-historical criteria (see, e.g., McWhorter 1998, DeGraff 2003 as examples of these competing viewpoints). Both Bailey and Maroldt and Poussa focus on structural considerations in the presentation of their hypotheses, but even if one accepts a structural definition of creoles, it is worth noting that the outcome of this contact situation does not match well with "creolization criteria" as they have been defined in previous studies. For example, McWhorter says that the clustering of three structural traits distinguishes a creole language: "little or no inflectional affixation", "little or no use of tone to lexically contrast monosyllables or encode syntax", and "semantically regular derivational affixation" (1998:798). This third structural trait can be further clarified: he claims that "in languages known as creoles, derivation is generally semantically transparent; ... evolved semantic idiosyncrasy ... is unknown" (1998:797). He, in fact, does specifically address the question of the creolization of English,⁹ noting

⁹ In particular, as his comments are in response to Bailey & Maroldt 1977, he is dealing with the question with regard to the outcome of the English-French contact subsequent to the Norman Conquest.

that "... while rather low in inflection, English does retain eight inflectional affixes. In addition, however, it is crucially distinct from creoles in its semantically evolved derivation" (1998:798). From this perspective, the contact between English and Norse (and subsequently French) likely added to the semantic idiosyncrasy and opacity of derivational affixes in English in that many of what were formerly semantically transparent and derivationally productive affixes became at most marginally productive, but remain in the language in certain forms where their affixal status might be clear but not their semantics.

Danchev (1997) employs a different set of structural criteria but comes to the same conclusion. He notes that the loss of OE short and long diphthongs and umlaut vowels, the accelerated loss of case endings in ME, and the loss of gender marking which is seen in ME adjectives and nouns can in a limited sense be seen as fulfilling creolization criteria. However, other criteria, such as prevalingly open syllable structure, no morph-syntactically marked passive, preverbal tense marking, reduced use of *be* copulas, the use of the same verb for possession and existence, lack of non-finite verbal forms, and lexical circumlocution, are not found in Middle English. Danchev concludes that while "[a]pproximately half of the more outstanding (weightier) creoleness (or creoleness-like) features occur in Middle English ... [t]hese are the features that have been attributed to more general factors ... defined as *universal language communication strategies*" (1997:97). These general changes, while "matched by similar or even identical changes ... in pidgins, creoles, and learner interlanguages", are found "in many other languages (related and unrelated ones)" (1997:98).

Wallmansberger (1988) similarly says that while "on the one hand reductions in surface morphology and the incipient, but quite noticeable trend towards analyticity correspond to factors in any creolization index, on the other hand the criteria that would constitute conclusive evidence for creolization are absent" (1988:29). Even for those who hold to structural criteria of creoles, therefore, the facts of English do not support the creolization hypothesis; nor does the contact situation between Norse and English fit the socio-historical definition as it has been applied to Caribbean and other creoles.

4.3 Interlanguage

Realizing the problematic aspects of describing the outcome of the English-Norse contact as a creole, some studies have applied the term "interlanguage", taken from second language acquisition, to this case. Fisiak proposed that "what must have emerged was an interlanguage. The formation of the interlanguage must have resembled the process of pidginization but it is doubtful whether it ever underwent any further development towards creolization" (Fisiak 1977 in Danchev 1997:80). Danchev further states that this interlanguage "developed first in the areas of Anglo-Scandinavian community bilingualism and then gradually spread over most of the country" (Danchev 1986:248). He sees interlanguage as "an apparently convenient alternative choice" because it is a "broader and more neutral blanket notion" than creole, but it also "covers most of the features shared by Middle English with pidgins and creoles" (1997:98).

Danchev concludes in his discussion of Fisiak's interlanguage hypothesis that the term does not really fit the ME situation since "while an interlanguage is a more or less developed approximation of an easily identifiable target language, for Middle English no such comparison is available" (1997:99). This is the most serious problem with the proposal: second language acquisition implies a source and target language, but which would be which in this situation? Also problematic is the fact that "interlanguage" implies that the changes in the language are the result of imperfect acquisition; the linguistic effects of the Norse contact on English are not of this type.

4.4 Neutralization

O'Neil (1978) deals with the effects of the contact on the grammar or morpho-syntax of English. In particular, he compares and contrasts the morphosyntactic changes which occurred in Northern English with those that occurred in other English dialects, as well as those that occurred in other Germanic languages. One of the characteristics of Northern Middle English (§2.3) was a "simplification" of the inflectional system, which led to an increased dependence on word order and other syntactic factors. This reduction has been important but controversial in discussions of this situation as we have seen, with creolization advocates claiming it as strong evidence for their position, but Thomason and Kaufman (1988) emphasizing instead that simplification had begun prior to the contact, and stating that the contact merely accelerated a process that would have taken place anyway.

O'Neil (1978) makes a specific distinction between what he refers to as "simplification" and "neutralization", with simplification being what was happening to the morphosyntactic system prior to and removed from the contact situation, and neutralization being what happened in the area of English-Norse contact. He notes that neutralization "is always rapid change and change involving very closely related languages (or dialects). And it is relatively superficial aspects of the languages (inflections, stress, tone, etc.) that are neutralized" (1978:248–49). He applies this notion specifically to the English/Norse contact situation as a key example of this outcome: "the complex inflectional system of Old English was largely and rapidly neutralized on contact with the complex inflectional system of Old Norse" (1978:249).

The focus of O'Neil's study is on the effects of contact on the morphosyntactic system of a language, but within this context he does mention some of the morphological effects as well. In particular, he explains the "borrowing" of the Norse forms of the third person plural pronouns (§2.2) as being "...presumably related to the fact that (a) distinctness between the plural and singular forms of the third person pronoun was lost or significantly reduced and perhaps not attended to at all by foreign ears, and (b) verb inflections marking singular from plural forms were also lost" (1978:261), thereby relating it to other changes in the language at the time. He makes a clear distinction between these morpho-lexical effects and the effects on the grammatical system, however: "...the inflectional simplicity is not borrowed ... What we have instead of borrowing is a neutralization of the inflections brought about by the speakers of the two languages in

their reaching for the inflectional common denominator by means of which they could communicate..." (1978:261).

The idea of neutralization presented by O'Neil fits well with the overall theory of koineization and specifically with the effects of the English-Norse contact on the English language. A crucial characteristic of a koine, as seen earlier, is that it is "often reduced in morphological complexity compared to the contributing subsystems" (Siegel 1985:373), which could also be described as "neutralized". The sociohistorical criteria for the contact situations are similar as well, with Siegel's "linguistically related subsystems" in koineization relatable to O'Neil's "two closely related languages differing for the most part only in superficial aspects of their grammars (inflections, accent, tone, etc.)" (1978:283), which he proposes as the inputs to neutralization.

Neutralization, however, deals mostly with the effects on the grammar of this language contact, while koineization presents a larger picture of the effects on the whole system. In other words, neutralization fits well as one part of the koineization process, and as one characteristic of the resulting linguistic system, but does not add much in the way of explaining lexical and morphological effects of the English-Norse contact on the English language.

5 Conclusion

The outcome of the English-Norse contact situation can best be analyzed within the framework of koineization. This analysis fits the sociohistorical context and the linguistic effects on the language varieties spoken in the northern and eastern parts of England, many of which later spread into standard English and are thus found in the Modern English. The creolization hypothesis as proposed by Bailey and Maroldt (1977) and Poussa (1982) is particularly problematic and the term "creole" should not be applied to the result of this contact situation. The other hypotheses that have been proposed to account for this situation, however, also have difficulty in adequately capturing the effects of the contact situation. The term "koine" takes into account the genetic and typological closeness of the language varieties involved in the English-Norse contact situation, and the koineization account explains the types of linguistic effects which are seen as resulting from it.

Appendix. Map of England.



References

- BAILEY, CHARLES-JAMES, and KARL MAROLDT. 1977. The French lineage of English. *Langues en contact - pidgins - creoles - languages in contact*, ed. by J. M. Meisel. Tübingen: TBL Verlag.
- BAUGH, ALBERT C., and THOMAS CABLE. 1993. *A history of the English language*. 4th edn. Englewood Cliffs: Prentice Hall.
- DANCHEV, ANDREI. 1986. Interlanguage simplification in Middle English vowel phonology? *Linguistics across historical and geographical boundaries*, ed. by Dieter Kastovsky and Aleksander Szwedek, 239–52. New York: Mouton de Gruyter.
- DANCHEV, ANDREI. 1997. The Middle English creolization hypothesis revisited. *Studies in Middle English linguistics*, ed. by Jacek Fisiak, 79–108. New York: Mouton de Gruyter.
- DEGRAFF, MICHEL. 2003. Against Creole exceptionalism. *Language* 79.391–410.
- GEIPEL, JOHN. 1971. *The Viking legacy*. Newton: David & Charles.
- GORDON, E. V. 1957. *An introduction to Old Norse*. Oxford: The Clarendon Press.
- GÖRLACH, MANFRED. 1986. Middle English – a creole? *Linguistics across historical and geographical boundaries*, ed. by Dieter Kastovsky and Aleksander Szwedek, 329–46. New York: Mouton de Gruyter.
- KERSWILL, PAUL, and ANN WILLIAMS. 2000. Creating a New Town koine: Children and language change in Milton Keynes. *Language in Society* 29.65–115.
- LOYN, H. R. 1977. *The Vikings in Britain*. London: B. T. Batsford.
- MCWHORTER, JOHN. 1998. Identifying the Creole prototype: Vindicating a typological class. *Language* 74.788–818.
- O'NEIL, WAYNE. 1978. The evolution of the Germanic inflectional system: A study of the causes of language change. *Orbis* 27.248–86.
- POUSSA, PATRICIA. 1982. The evolution of Early Standard English: The creolization hypothesis. *Studia Anglica Posnaniensia* 14.69–85.
- SIEGEL, JEFF. 1985. Koinés and koineization. *Language in Society* 14.357–78.
- THOMASON, SARAH GREY, and TERRENCE KAUFMAN. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

TRUDGILL, PETER. 1986. *Dialects in contact*. Oxford: Basil Blackwell Ltd.

WALLMANNBERGER, JOSEF. 1988. The “creole hypothesis” in the history of English. *Historical English: On the occasion of Karl Brunner’s 100th Birthday*, ed. by Manfred Markus, 19–36. Innsbruck: Institut für Sprachwissenschaft.

**FROM TAXONOMY TO TYPOLOGY:
THE FEATURES OF LEXICAL CONTACT PHENOMENA
IN ATEPEC ZAPOTEC-SPANISH LINGUISTIC CONTACT**

Craig Hilts

Abstract

In this paper, I begin with an examination of what constitutes a borrowing from one language to another with particular reference to lexical borrowing. I develop a set of three aspects of words/lexemes that can serve as features within the context of borrowing and as a model for their representation to be used to account for lexical contact phenomena, and compare them with characteristics used in previous descriptions of these phenomena. I then apply a featural analysis to the currently accepted taxonomy in order to demonstrate its lack of consistency in arbitrarily excluding a part of the lexical results of cultural contact and in failing to distinguish crucial differences in the agentivity of change. I argue that, by using these features, the full scope of lexical contact phenomena can be described. Using a derived and coherent terminology, I apply the features to the results of Atepec Zapotec (AZ)-Spanish (Sp) contact and conclude with a discussion of possible uses of this typology in terms of other areas of contact linguistics.*

* An earlier version of this paper was presented to an Ohio State Linguistics Departmental colloquium January 6, 2001, and portions were presented at the Kentucky Foreign Language Conference, April 20, 2001. My thanks to Rich Janda, Jen Muller, Tom Stewart, the Changelings socio-historical linguistics study CRAIG HILTS. From taxonomy to typology: The features of lexical contact phenomena in Atepec Zapotec-Spanish linguistic contact. *OSUWPL* 57, Summer 2003, 58-99.
Copyright © 2003 The Ohio State University

1 Contact reaction choices

When a new item or concept is encountered by the speaker of one language coming into contact with another place or culture and its language, one of several reactions may occur.

- a. The first might be to simply ignore the item, due to its lack of cultural or personal salience, as might occur with an animal species that is apparently useless, even in an esthetic sense, and threatless, e.g., a type of gray moth such as *Hemeroplanis punitalis*, found in the Southwest US, with no common name.
- b. A second might be to use the resources of one's native language to create a linguistic sign for the item, as with American English 'yellow-bellied sapsucker', a small woodpecker with a yellow breast that pecks holes in the bark of trees in order to drink the sap for its nourishment.
- c. A third might be to modify an existing native resource by the addition of a new meaning to a word, e.g. English *polecat* (from ME *poul-* as in *poultry* + *cat* in the sense of 'small hunting mammal'), a European relative of the weasel which has a foul odor it uses for marking territory, expanded its semantic range to include 'skunk'. Another option would be to change the referent completely, as occurred with *robin*, which changed red-breasted bird species with its change of continent (from Europe to North America).
- d. A fourth could be to learn, or at least to approximate, the word from the language spoken by the other culture, as with *skunk*, from Abenaki *seganku*.
- e. A fifth would be to use some combination of the second and fourth approaches, as with *woodchuck*, from Algonquian *wuchak*.

The last three approaches are what is involved in lexical borrowing, but the second approach should also be considered in listing the lexical results of contact as evidence of the type and the history of the contact situation. Not all borrowing occurs under the impetus of novel phenomena, but these are a starting point for an examination of the lexical results of contact.

Table 1 below gives a listing of terminology used to describe the lexical results of linguistic and cultural contact as offered by four sources, with Haugen's (1953) system the (apparent) primary source for the rest. The numbers given for each category correspond to the approaches given above.¹

group, and particularly to Brian Joseph and Don Winford, without whom this paper would have stopped before it started down the runway. All pilot errors are my own.

¹ Naturally, the first category will not be represented in the table.

Sources

Haugen 1950, 1953	Type	Hock 1991, Hock & Joseph 1996	Romaine 1988	Weinreich 1953
loanwords	4	loanwords	loanwords	loanwords
loan shifts	3	loan shifts	loanshift/ semantic transference	semantic extension
loan homonyms	3			homonymy
loan synonyms				polysemy
semantic displacement	3	<i>semantic shift</i>	semantic extension	
semantic confusion	3	loan shift	<i>semantic extension</i>	
loan translations (calques)	5	loan translations (calques)	loan translations (calques)	loan translations (exact) loan renditions loan creations (loan mapping)
loan blend		loan blend	loan blend	hybrid compound
stem derivative compound	5 5 5			
creations induced	2		part of loan translations	loan creations
creations (non-borrowings)				
hybrid creations	5		loanblend	

Table 1. Terminology of the lexical results of linguistic/cultural contact.

2 'Patterns' in borrowing

Haugen (1950) begins the discussion of borrowing by assuming that “every speaker attempts to reproduce previously learned linguistic patterns ... among which ... are those of a language different from his own ... [which may be reproduced] not in the context of [that] language”, and he defines borrowing as “the attempted reproduction in one language of patterns previously found in another” (1950:212). Haugen is never fully explicit in his definition of what constitutes a *pattern*, but by analyzing the descriptions he offers for the types he creates, we can abstract them at an elemental level.

One of these “patterns”, the form, is obvious from the outset. Phonological form is the most transparent, and therefore the prime, indicator of linguistic contact. We can use the phonological/phonetic forms as a feature once they are related to their respective source languages.

The meaning, or referent (see below), of the form is not referred to by Haugen as a pattern, since it is (relatively) fixed in the world, but it may be culturally specific (e.g., *taco*), and as such, it may be considered to have an attributable source and thus be a feature.

In addition to those two types of feature, a third figures into the description according to Haugen. In his description of the taxonomy, with particular regard to calques, he refers to a “pattern of morphemes”. He never explicitly defines what he means by “morpheme”, but because he does characterize loanwords as “imported morphemes”, we can suppose that what he is referring to as a “morpheme” is the combination of the phonological form and the meaning. The pattern he speaks of is that of the combination of morphemes (the juxtaposition of forms and meanings) that signifies the (otherwise) idiosyncratic meaning of the calqued compound. What he seems to be talking about is what maps the relationship between the form and the meaning within a language.

If we approach this relationship between the form and the meaning from a semiotic point of view, it becomes somewhat clearer.² Following the de Saussurean tradition, we can refer to the thing signified (the meaning or referent) as the **significatum**, and to the word(s), more exactly the phonological form, signifying it as the **significans**. Lyons (1977:96) refers to a scholastic maxim *vox significat [rem] mediantibus conceptibus*: “the word signifies [the thing] by means of **mediating concepts**”. It is the mediating concept that allows a combination of [skai] ‘sky’ and [skreipə] ‘scraper’ to signify ‘very tall building’. If we were to use (near-)synonyms [hevinz] ‘sky’ (“the heavens poured down rain”) and [greizə] ‘scraper’ (“he grazed his knuckles”), the compound *heavens grazer* does not carry the same meaning; it is the exact juxtaposition of particular forms and meanings that acts as the mediating concept to give the meaning ‘very tall building’ to otherwise semantically unrelated forms.

The relationship itself is arbitrary (with certain sound symbolic exceptions, e.g., ‘cuckoo’), but fixed. It can change, but not arbitrarily, not without a catalyst of some sort. The relationship is also ad hoc, a product. There is no term for the mediating concept that maps the relationship between form and meaning, but to use Haugen’s “pattern” is to use his generic term (referring to phonological form, meaning, etc.) for a subordinate category, and fails to distinguish it from a general term with too much possible polysemy within linguistics. I use the term **mapping** to mean the relationship that exists between two otherwise arbitrarily associated entities (form and meaning) to form a word.

3 Mappings

The simplest example of a relationship would be that of a monomorphemic form such as [plætəpʊs],³ which would have one meaning, namely ‘platypus’, the only aquatic egg-laying mammal. A model of the relationship, the mapping, between the form and the

² My thanks to Thomas Stewart for suggesting this approach.

³ The issue of the awareness of etymological morphology for a words like *platypus* or *conduct* is also beyond the scope of this paper, since most native speakers are unaware of the polymorphemic etymology of words like these, and second language learners would be even less aware of it (see *hoosegow* below).

meaning of this word, could be done in a linear representation, as in Figure 1, between form (F) and meaning (M):

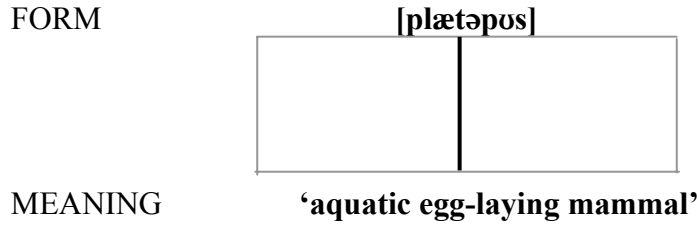


Figure 1. One meaning-one form mapping.

In this case, the mapping is represented as one-dimensional by the bold single vertical line between the form and the meaning, each of which is represented as a point on the respective lines of all possible forms and meanings.

At the monomorphemic level of relationship there are also single meanings with varying forms and, conversely, single forms that have semantically unrelated meanings, polyphony and polysemy respectively, as can be seen in Figure 2. In order to be able to demonstrate polyphony and polysemy with this model, we have to add a dimension to make it planar. Note that it is the surface relationship between form and meaning (i.e., apparent to the listener) that concerns us here rather than, e.g., whether polysemy applies to a single lexical entry.

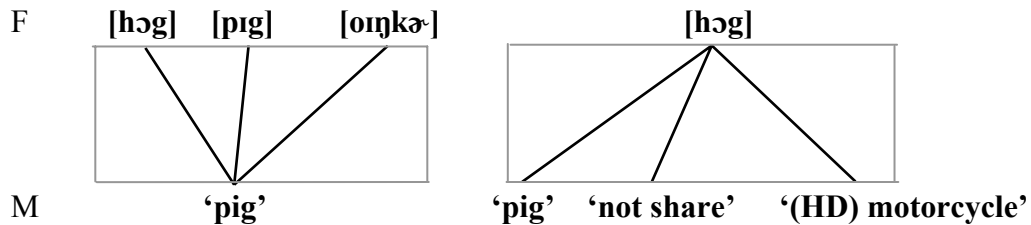


Figure 2. Polyphony and polysemy.

For the meaning 'pig' (porcine mammal), there are (at least) three English forms that correspond, of which two are monomorphemic (*hog*, *pig*) and one is bimorphemic (*oink-er*), a 'function describer'. Conversely, the mapping associated with the form [hɔg] in American English correlates that form with three meanings, two of which are nouns, and one of which is a verb. Two of these meanings are metaphoric extensions of the leftmost meaning in the representation of the mapping, but in terms of semantic fields have little in common.

Other hierarchical levels of meaning can reasonably be considered within the strictly semantic, or non-pragmatic, realm. In an extremely simplified sense, the first of these would be the compositional meaning, as in compounds, which can be derived by the (more or less) straightforward combination of the meanings of each morpheme. A

polymorphemic case, strictly compositional, would be *bluebird*, the mapping for which can be seen in Figure 3.

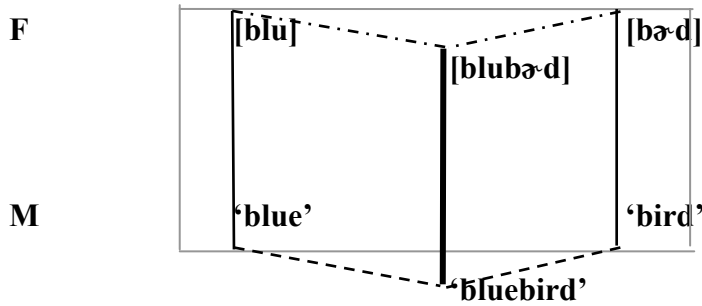


Figure 3. Compositional compound.

As can be seen here, not only is there a form-to-meaning relationship for the individual morphemes, but also, as the two types of dashed lines indicate, their forms and meanings are all contributive to the resultant compound form and meaning. To capture the aspect of a pattern of combination, another dimension is added to the representation. In this case, the combination is essentially compositional, with both form and meaning combining to describe “a bird that is blue”.⁴

The next level of meaning is more idiosyncratic and is also of compounds; an example is *monkey wrench*, modeled in Figure 4, which is a specific type of wrench, but which is neither used nor made by or for monkeys.

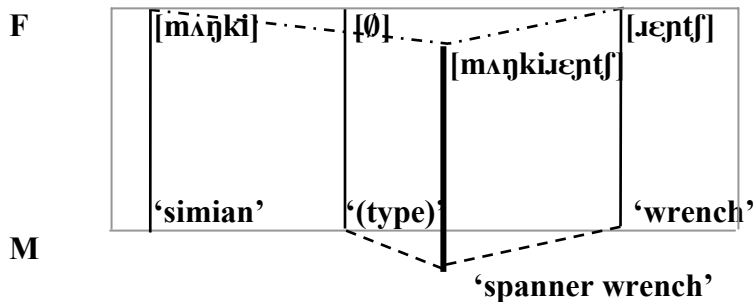


Figure 4. Idiomatic compound.

In this case, the form [ɹɛntʃ] and the meaning ‘wrench’ correspond, and each are part of the composition of the compound. The form [mʌŋki] is part of the composition of the form and is related to the meaning ‘simian’, but the meaning ‘simian’ is neither compositionally nor apparently metaphorically⁵ connected (and hence not representationally connected) to the meaning for the type of wrench, and the meaning of type has no compositionally relevant form.

An even more purely metaphoric meaning can be seen in idiosyncratic compounds such as *skyscraper*, in Figure 5, in which the strictest compositional meaning has

⁴ This is meant as an example, and ignores the prosodic differences between “blue bird” and “bluebird”.

⁵ Again, this is an issue of opaque etymology for the listener.

no referent, since *scrape* is not something that can be done to the sky (as a non-surface) in any but a metaphorical sense.

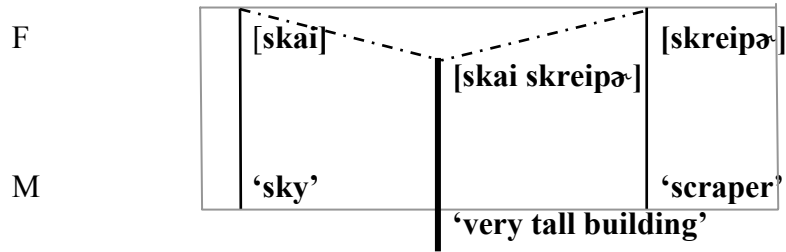


Figure 5. Metaphoric compound.

While the form is compositional here, the meaning is non-compositional (although metaphoric), unlike the case above for *bluebird*; and this is represented by the lack of connecting lines on the meaning plane. Metaphor and the idiosyncrasy of semantic relationships to each other are outside the scope of the mapping of form and meaning in this paper.

This idiosyncrasy and metaphoric meaning can also be represented phrasally, as in Figure 6. In the example of *kick-ed the bucket* ‘died’, the four morphemes ($V_{\text{ROOT}} + \text{PAST}$, DET, N) are each analyzed independently by the speaker (and, we may assume, by the hearer, regardless of whether or not the idiomatic meaning is understood), and because of that, the phrasal verb would not be given as **kick the bucketed*.

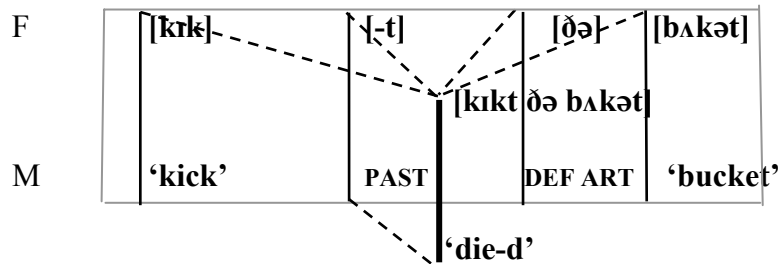


Figure 6. Phrasal compound.

Here we can see the PAST morpheme adds meaning compositionally, but the other forms do not contribute their associated meanings.

We can now look at how these relationships can be a part of the borrowing process. If we begin with the most basic case, we can show the relationship between the forms, the meaning, the mapping, and the languages. The model below in Figure 7 consists of two vertical planes representing the form/meaning relationships of two languages, in this case Spanish and English, separated by the differences in form and mapping, and their connection is the plane of meaning. A meaning is assumed to be a single entity irrespective of what language is involved and is represented as a point in the “line” of meaning at the bottom of the mapping representation. For this model, as with the representations above, we assume that the universe of possible meaning is represented by a single line of which the representation is only a part; we ignore any possible geometry of inter-

relationships and interconnectivity between meanings. For each representation given of two languages, the meanings are aligned along that line in the same sequence; thus, each language contains the same line of meaning. For any point of meaning, there may or may not be a mapping to relate one or more forms, so that for ‘kazoo’ there is a mapping in English, but there is probably no mapping for ‘kazoo’ in Xingu (Brazil). If we extend all the points of meanings along one dimension to make each meaning a cross-linguistic line, we create a plane of meaning, completely congruent between languages. As noted above, in terms of reference, meaning is fixed and constant in the world, regardless of language; an elephant by any other name would smell as trunkily. Connections of meaning between the languages are assumed to be parallel with the ends of the meaning plane. There is no explicit plane of form because, unlike the meaning, it is arbitrary.

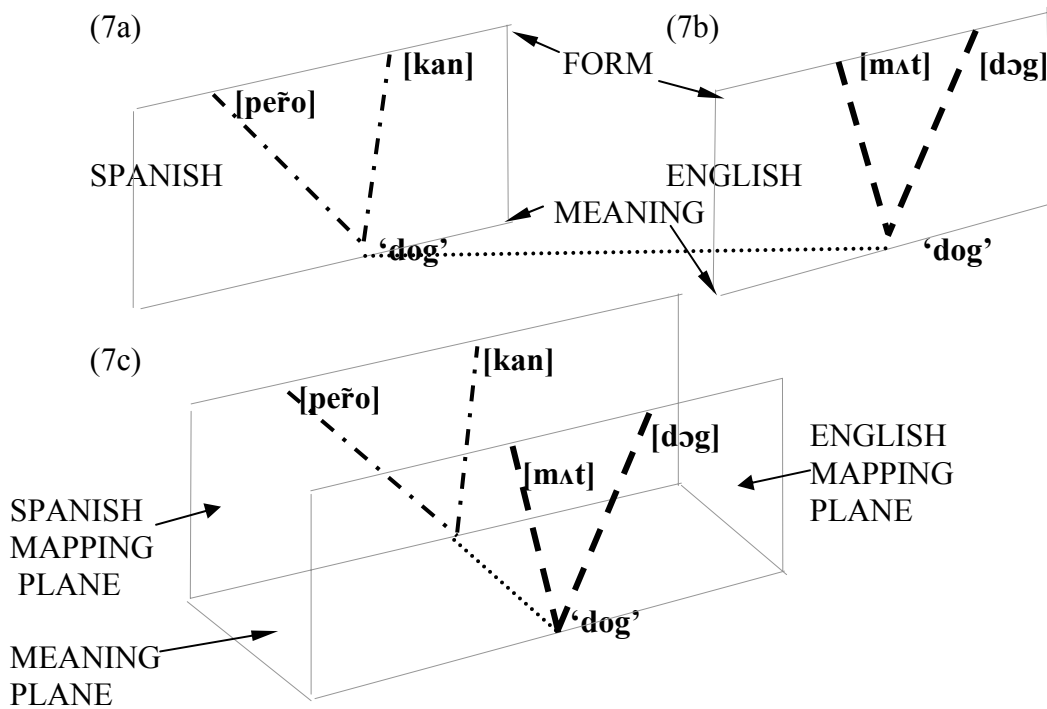


Figure 7. Spanish and English mappings of ‘dog’.

In Figure 7 above, we see parts of the planes that represent English and Spanish in (a) and (b) respectively, and we see how they relate to one another in (c).

When we observe the representation in (c) of the Spanish and English mappings as though sighting along the line of meaning of ‘dog’, we get the result in Figure 8, with the Spanish mapping as dash-and-dot lines and the English mapping as dashed lines.

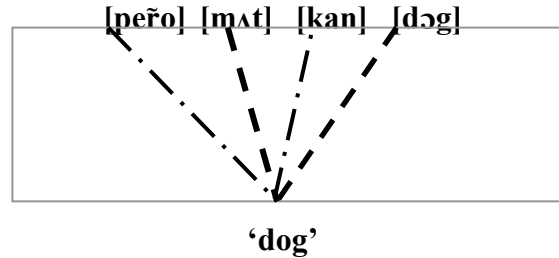


Figure 8. Spanish-English mapping comparison.

It is readily apparent that the two mappings are not congruent, since the points on the line of form cannot be close, given the fact that none of the forms have even one phoneme in common, although both mappings are polyphonous and originate from a single meaning. Just as the polyphony within a single language must be represented as divergent, so also must the polyphony between languages.

In the example in Figure 9, *taco*, the phonological forms differ in allophonic variation only, but for many other loans, e.g., *burrito*, there are, arbitrarily, differences in a number of phonemes: Spanish [buɾiɾo] vs. English [bæˌɪrɪtʊ], which may not even be consistent from token to token. The arrows given indicate only the direction of the change, and are *not* intended to indicate agentivity as defined by van Coetsem (1988:10). In this case, because the form and the meaning both come from Spanish, the form from the language and the meaning from a culture that uses it, the mapping (by default) is also Spanish. We can see that the mapping is the same for both languages because they have exactly the same form-to-meaning relationship. In this case, both the form and the mapping have been borrowed, which is signified by the dashed line for the form and by the triple lines for the mapping.

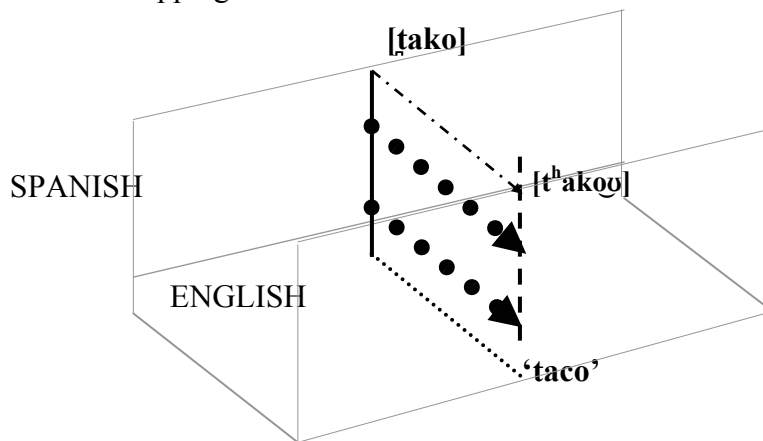


Figure 9. Pure loan.

We can now look at an example of the creation of a polyphony due to the borrowing of a form that loses its mapping in connection with a change in the meaning, probably due to a misapprehension at the time of borrowing: *hoosegow* 'jail' from Spanish *juzgado* [xus'ɣa(ð)o]⁶ 'judged' in Figure 10. A more recent borrowing from English

⁶ Often pronounced with a very lenited or omitted [ð].

into Japanese is “Viking” [baikiŋgu] ‘buffet’, probably semantically from ‘smorgasbord’ or from the name of a restaurant chain.⁷

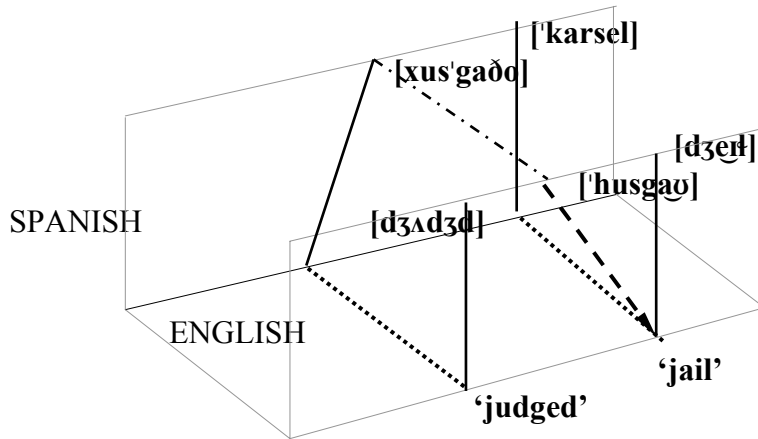


Figure 10. Polyphony through borrowed form.

Here the Spanish form has been roughly (phonemically) replicated into English, and then mapped to one possible consequence of being judged, ‘jail’. What we can see here is the creation of a polyphony due to the borrowing of a form. The mappings here are not equal, since the meaning ‘jail’ relates to only one form in Spanish but to two in English; therefore the mapping is clearly not borrowed.

We can also find a loan of polysemy in the American Portuguese (AmP) borrowing from English of the mapping of the form corresponding to ‘cold’ (AmP [frio]) and the meanings of ‘low temperature’ and ‘viral disease’, seen in Figure 11.

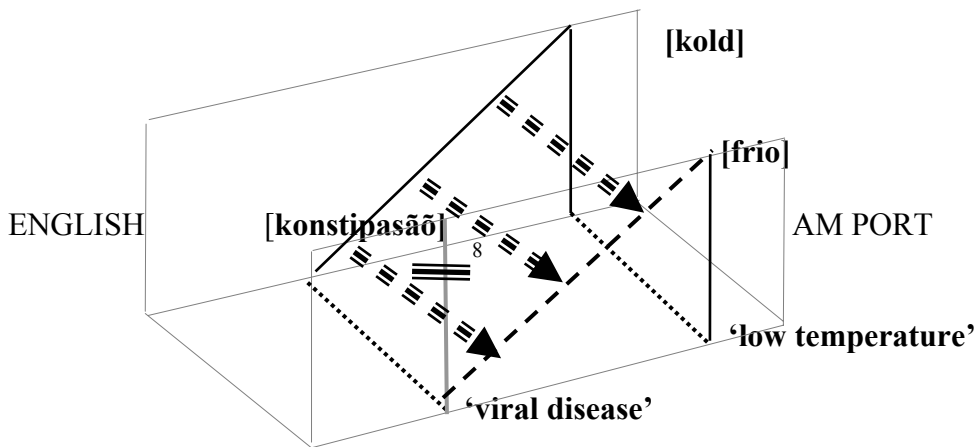


Figure 11. Borrowed mapping.

Here the AmP mapping that previously associated the form [konstipasãõ] with the disease meaning has been dissociated, and, crucially, *only* the English polysemous mapping has

⁷ My thanks to Kaoru Yoshida for this example.

⁸ Note that the triple line signifies the loss of the form-meaning mapping here.

been borrowed. We cannot say that the meaning of ‘viral disease’ has been borrowed, since its referent already existed in the culture, and that meaning was mapped onto a different form in the language.

In the AmP loan translation *estar dereito* ‘to be right’, we have a biplanar borrowing of the mapping that again has no connection with the phonology, seen below in Figure 12a,b. In Figure 12a, the borrowing of the lexical mapping is shown, as occurred above in Figure 11, remapping [dereito] to ‘correct’, an addition to the native Portuguese (P) meanings of ‘opposite of left’ and the “right” of ‘civil rights’. Figure 12b demonstrates that the morphemes *est-a-r* (ROOT ‘be’ - thematic vowel - INFINITIVE) are still analyzable and take part in normal morphosyntax, but the borrowed polyphony of *dereito* ‘right’ as “correct” occurs only in this phrasal frame.

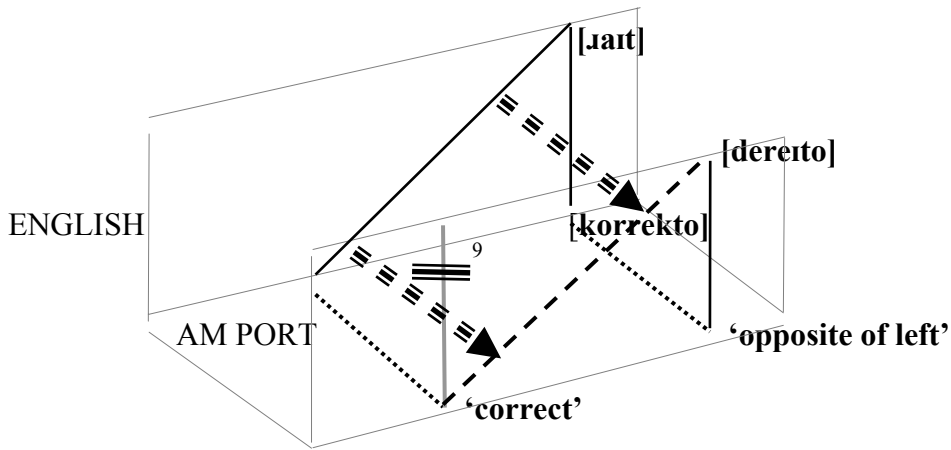


Figure 12a. Borrowed mapping.

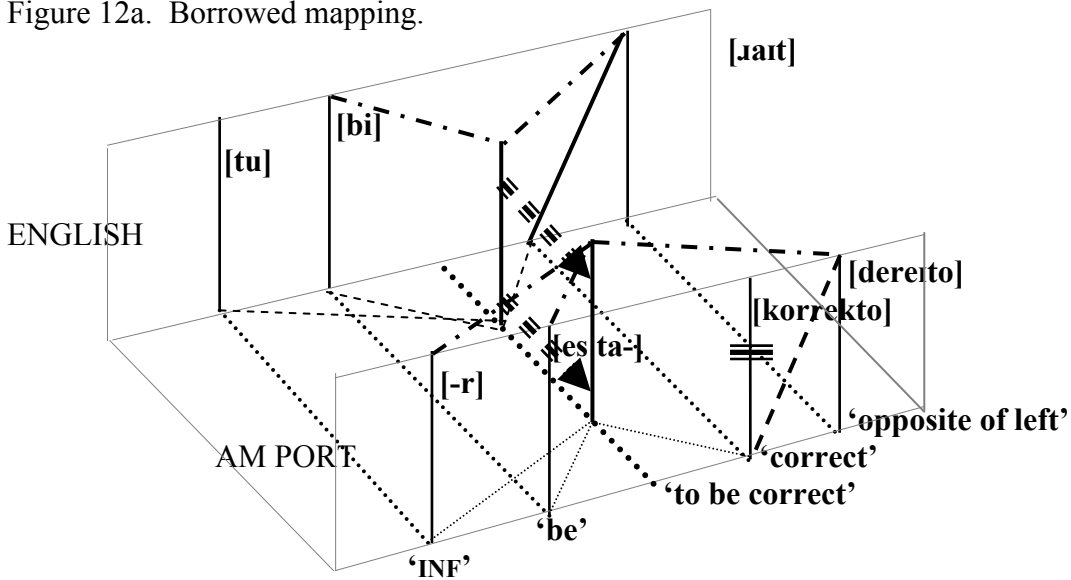


Figure 12b. Multi-faceted borrowing of mapping.

⁹ Note that the triple line signifies the loss of the form-meaning mapping here.

As can be seen, the order of morphemes is dependent on the language and makes no difference in the mapping, which must be typologically correct for each language. In this case, the use of this English mapping forces the polysemy of *dereito*. Here, there are two planes of borrowed mapping. The first is the association of the meaning of ‘correct’ with the form associated with the meaning of ‘opposite of left’. The second is the association of the phrase ‘to be X’ with the other mapping without which the association of *dereito* with ‘correct’ does not occur. In Portuguese, the normal form for a person ‘to be right’ is *ter razão* ‘to have reason’, while for an answer ‘to be right’ is *estar correcto*.

We can now look at the canonical calque, *skyscraper*, to see how the mappings of English and Spanish combine to produce the Spanish form.

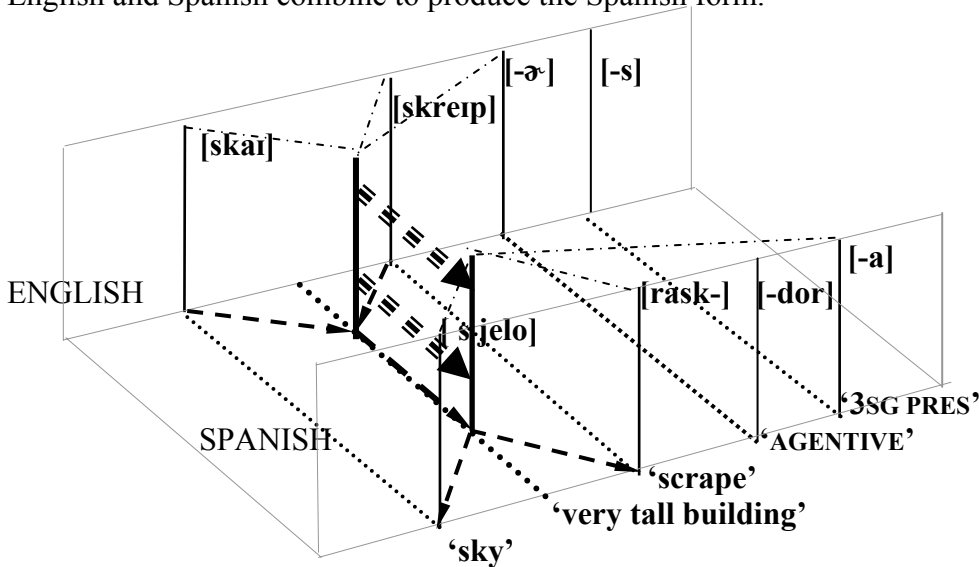


Figure 13. Partial mapping borrowing.

This is thus a partial calque, in which the pattern is not wholly emulated, but elements of the English mapping have been used (i.e., the relation of the noun and verb root forms), as has part of a Spanish mapping, which uses a “description of function” as opposed to the English “agent of function” morphosyntactic pattern. If we observe the correlation between combinatory mapping patterns from “above”, perpendicular to the plane of meaning in Figure 13 above, it becomes clearer, as can be seen below in Figure 14.

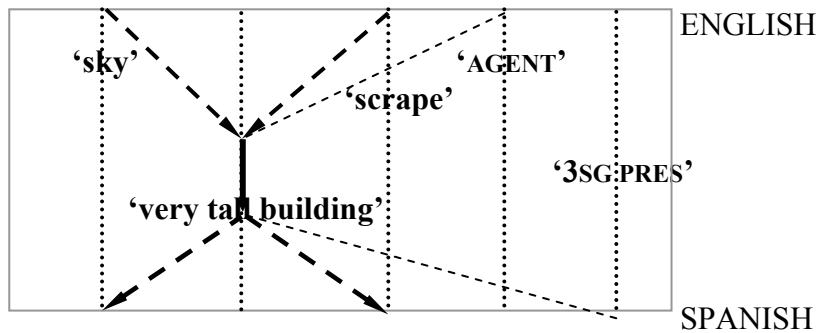


Figure 14. Partial mapping congruence.

These representations are only a sample of mappings in some of the various levels of complexity that occur in lexical contact phenomena.

ATTRIBUTION OF FEATURE ORIGINS

As a shorthand method of offering the same kind of information shown above in the diagrams, we need a method of labeling the source(s) of the features found in the results of lexical contact. We could use Haugen's (1950) terminology of *source* to refer to the language borrowed from and *recipient* as the borrowing language, but the use of *source* in its more general sense ("the source(s) of the features" above) vs. its particular sense here is confusing and seems to be more problematic than using the more vernacular but inherently more iconic terms **lender** and **borrower** languages.¹⁰ We can combine this aspect with the phonological forms and meanings of the words to create features, e.g., [lender form], [borrower meaning].

For a single word, each feature (e.g., form) can have aspects from both languages, as with Haugen's example of Pennsylvania German (PaG) *bass-ig* 'boss-y' (*bass* < English *boss*, *-ig* 'having the qualities of', as with E *-y*), which combines morphemes (meanings and their associated forms) from English and PaG, respectively.¹¹ This word would thus have the features [lender form] for English [bɒs]¹² and [borrower form] for [-ig], as well as [lender meaning] for English *boss* and [borrower meaning] for PaG *-ig*. These features are privative. Binary features would leave a logical impossibility as a category if both sources for forms were [-] (e.g. [-lender form, -borrower form]), since there would be no connection of form to either language.

For *bassig*, then, the list of features would also include [lender mapping] and [borrower mapping]. I abbreviate these feature components as follows: lender, **L**; borrower, **B**; form, **F**; meaning, **M**; mapping, **A**. Thus, *bassig* has [borrower form (BF)] and [lender form (LF)] based respectively on *-ig* and *bass*; [borrower meaning (BM)] and [lender meaning (LM)], based on the concept of "having the characteristics of an overseer" existing in both cultures; and [borrower mapping (BA)] equal to [lender mapping (LA)], based on the equality of the derivational morphosyntax of both languages. In order to delineate the status of each feature in the word, they are unmarked for those with no change or non-equality, [+] for those that were added, [-] for those subtracted (loss of), and [=] for those already equal. The complete feature bundle for *bassig* would thus be [BF, LF, BM = LM, BA = LA], based on the facts that the form is a mixture and the meaning is the same in both languages, as is the mapping. *Bassig* is represented in Figure 15, which has a single mapping for the bimorphemic lexeme, since the manner in which

¹⁰ Haugen defines these categories as "lending" and "borrowing", with the usual explanation that these terms are not literal, for which reason he offers *source* and *recipient*.

¹¹ The existence of *baas* 'boss' in Afrikaans (from Dutch *baas*, Middle Dutch *baes*) was brought to my attention by Ilse Lehiste. This fact may shed some doubt on the original source of the form/meaning conjunction, although there was also English contact in South Africa. That controversy aside, however, this is Haugen's example and serves to illustrate the point being made.

¹² Note that the form referred to here is the basic phonological shape, without consideration of adaptation towards (or in the case of hyper-foreignism, away from) borrower language phonological distinctions.

PaG and English form this kind of adjective is exactly the same, and in this case, unlike that above in *estar dereito*, there is no difference in the order of the morphemes.

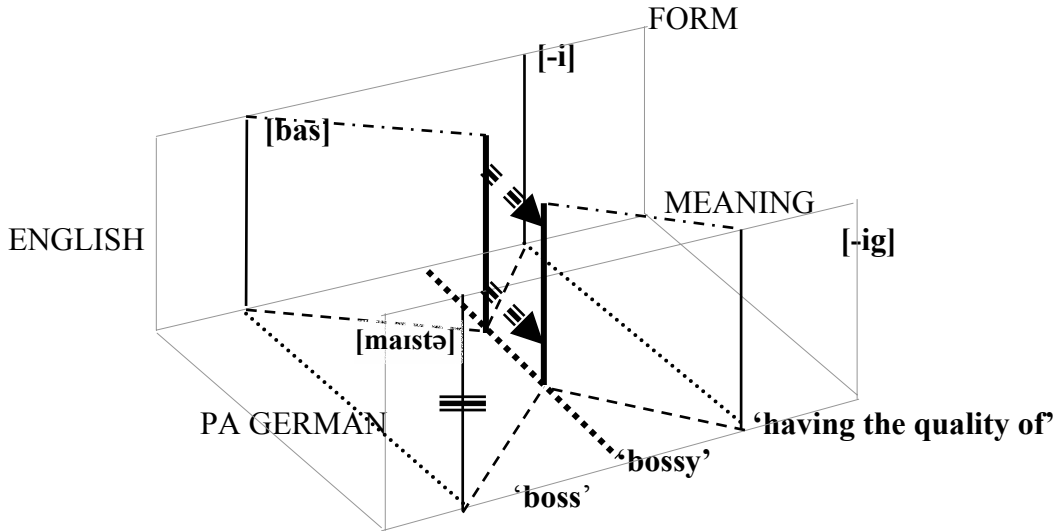


Figure 15. Super blend.

The representations of the processes involved in borrowing above are for clarification of what the features are and how they operate. The abbreviations for the features will be used for the rest of this paper now that the relationship(s) between the three have been demonstrated.

The basis for assigning a source to the phonological form of a word will be the origin of the phonemic, unadapted form. We should also note here that the genetic distance to the branching (if any) of dialects that separated the languages will have an effect on what kinds of correlations may be found in the forms. Many of the examples used by Haugen and analyzed below have relatively “short” genetic distance, e.g., Portuguese and French (for the AmP borrowing of the meaning ‘library’ from English on a homophonic basis), and because of that, experience reconvergence of meanings after divergent semantic drift from the parent language. This accounts for Haugen’s preoccupation with homophony, which in unrelated languages would be completely random and therefore of low frequency, and more so when combined with any similarity of semantic content, a basic assumption of genetic linguistics. For instance, there are no examples of homophony between Spanish and Atepec Zapotec other than interjections equivalent and more or less homophonous to English *Oh!* or *Ah!*, as would be expected.

The basis for assigning the source of the meaning of a word will be the existence of a form associated with the meaning within the language. Using AZ and Spanish as a contact example, where the Spanish were the impinging culture, a physical entity or cultural concept will have the feature of borrower meaning [BM] if it is native to the Zapotec territory (e.g., *coyote*, *blowgun*, *earth spirit*), lender meaning [LM] if it is imported

from outside that territory (e.g., *horse, hour, godparent*¹³), and both [BM = LM] if it is native to both cultures' home territories (e.g., *dog, marriage, red*).

The source of a mapping will be the language of the original phonological form associated with a particular meaning, but only as a default for monomorphemic monosemous words. Any word containing more than one morpheme will of course have a language-specific mapping, as will any polysemous word, in which any counterevidence will override the default. Linguistic categories, e.g., conjunctions, will be considered to exist in both languages, since the meaning will be possible to communicate through one means or another, and here the mapping will be the determining feature, e.g., through compounding, periphrasis, intonation, or a word/morpheme. An example from AZ would be the borrowed conjunction *sin kíʔni lǎ* 'without' ← Spanish *sin* 'without' + *kíʔni lǎ* 'so that NEG'. The "lacking" aspect of 'without' must have been conveyed previously (pre-contact) by means other than a word.

It should be noted here that not all feature combinations are possible. For the form, only the following five bundles would have values:

- a. [LF] English *taco* lender form
- b. [BF = LF] English and French phonemically equivalent forms *library* and *librairie*
- c. [BF] English *polecat* for 'skunk'
- d. [BF + LF] English [wɒdʒək] for Algonquian [wučak] 'woodchuck'
- e. [LF – BF] Atepec Zapotec [xaywar (LF)] from Spanish *jaguar*, lost its native form [–BF].

The combination [BF – LF] would be undetectable, with the native language form and the loss of the lender language form, and the combination [+LF = BF] is contradictory, with a lender form added to an equal borrower form; both are vacuous in this situation. For meaning, there are six possible combinations:

- a. [LM] English 'skunk'
- b. [BM] AZ 'coyote'
- c. [LM = BM] 'woman'
- d. [LM – BM] American *robin* from the English to the North American species
- e. [BM – LM] AZ turning Spanish *duende* [dwende] from 'goblin' to 'bad/evil'
- f. [BM + LM] English *buffalo* adding 'bison'.

The mapping can have four possible combinations:

- a. [LA] American Portuguese *frio* including 'viral disease'
- b. [BA] English *yellow-bellied sapsucker* as a purely English creation
- c. [LA = BA] PaG *bassig* as noted above
- d. [LA + BA] for Spanish *rascacielo* based on the mixture of mappings.

¹³ The Spanish concept of *compadrazgo* 'godparenthood' was different enough from the Zapotec version to give the Spanish names to the participants.

Thus, in theory, there are $5 \times 6 \times 4 = 120$ possible featural combinations.

4 Features and Haugen's taxonomy

As seen below in Table 2, Haugen's taxonomy of changes to the lexicon due to linguistic/cultural contact has as an axis the poles of lexical **borrowing** and what Haugen calls lexical **creation**, using only native resources. Lexical borrowing for Haugen (1953) is a two-part distinction: *loanwords* and *loanshifts*. These categories represent continua of a decreasing progression of the relative amount of lender language phonological/phonetic forms included in the borrowing, and can essentially be situated between the two poles of the 'borrowing' of meaning (**loan meanings**) and the borrowing of a phonetic shape (**loan forms**) as separate, although not exclusive, aspects of borrowing. Table 2 demonstrates Haugen's description of the taxonomy, and I include each type listed according to its characteristics.

The aspect of meaning must be considered in an intuitive manner for some aspects of this typology. There is considerable difference of opinion as to what constitutes a semantic category, what synonymy is, and how to determine the relative semantic closeness of two words. These arguments are beyond the scope of this paper; we can proceed with a general sense of semantic relatedness without the necessity for a complete formal set of distinctions for our purposes. Thus we can say that the informal connections we make between categories of things are as sufficient for this task now as it has been in previous analyses. The same kind of generality must be applied to the phonological form, which can vary greatly from the influence of adaptation due to borrower phonology (see above description of Figure 9), the incidence of bilingualism, the duration and intimacy of contact, prestige vs. solidarity considerations, and other extralinguistic factors (cf. Thomason & Kaufman 1988). In the case of what Joseph, Janda, and Jacobs (1999) refer to as hyper-foreignisms, e.g., *lingerie* [lɛ̃zəʁi], commonly pronounced in Standard American English (SAE) with a correctly nasalized but incorrectly placed first vowel and a "French" ending [lɛ̃zəʁei], the phonological variation corresponds to the borrower language speakers' impression of lender phonology, which would still have a borrower mapping. The French-ified pronunciation of 'party' [paʁtei] to connote "elegance" is evidence of the existence of English speakers' ideas about French phonology. This process is most likely an after-borrowing occurrence in any case. What must be considered in terms of a featural typology is the original phonemic form.

Haugen (1953)	Description/stimulus	Example
LOANWORDS	= phonology and morphemes	
Pure loanwords	wholly morphemic import	Spanish → E <i>taco</i>
Loanblends	partial morphemic import	
stem	mixed monomorpheme	E→AmN <i>kårna</i> ‘corner’
derivative	mixed types of morphemes	E→PaG <i>bassig</i> ‘bossy’
compound	mixed free morphemes	E→PaG <i>blauməpai</i> ‘plum pie’
LOANSHIFTS	borrower word changes meaning	
Extensions	semantic loan (added meaning)	
homologous	stimulus	E→QFr <i>librarie</i> = ‘library’
homophonous		E→AmP <i>grosseria</i> = ‘grocery’
synonymous		E→AmP <i>frio</i> = ‘cold (disease)’
Creation (calque)	imported arrangement	
literal	identical arrangement	E→AmP <i>estar dereito</i> ‘to be right’
approximation	approx. arrangement	E→Spanish <i>rascacielo</i> ‘skyscraper’
CREATIONS		
Induced creation	imported meaning only	Pima “downward tassles” ‘oats’
Hybrid creation	imported meaning, mixed form	Yaqui <i>lios nooka</i> “God speak” ‘pray’

Table 2. Haugen’s taxonomy and category descriptions.

‘*LOANWORD*’ is Haugen’s first major category, and his first division of that category is the *pure loanword*. As Haugen notes (1950:214), it is normally applied to those words in which the form is borrowed, with more or less complete phonemic substitution. As noted above, the mapping is borrowed, too, associating that form with the same meaning as in the lender language, as in AZ *bezhu* from Spanish *peso*, a coin denomination. That is the sense in which we will be using the term. The features for this would therefore be [LF, LM, LA]. (For a visual representation, see Figure 9 above.)

‘*Loanblend*’ is Haugen’s second subcategory of loanwords, and is characterized by conjoining native and borrowed phonological forms and/or meanings to form a word, irrespective of the level of phonological borrowing.

The *blended stem* is monomorphemic: Nor. *hyrna* + English *corner* → Am.Nor. *kårna* ‘corner’. (The blended segments from each form are in boldface.) He cites a rhotic pronunciation of the English agentive suffix *-er* as evidence that this monomorphemic word is not merely a Norwegian pronunciation of an English word. It is likely that in most, if not all, cases of stem blending, apparent blending will be a case of phonological substitution rather than an actual blending of forms. The features for this are: borrower

and lender form, equal meaning and, by default, equal mapping [BF + LF, BM = LM, (BA = LA)].

We can turn now to multi-morphemic units, which will often contain some aspect of the mapping. While Weinreich (1953) and Haugen both stipulate the necessity of some sort of bilingualism for borrowing to occur, Diebold (1980) makes the point that in the earliest stages of contact, what he calls *incipient bilingualism* can occur, in which the knowledge of the lender language is “atomistic”, consisting of words and formulaic phrases, with little knowledge of morphology or syntax, such as would occur in the first stages of pidgin creation.

Jackendoff (in prep.) discusses the semantic aspects of compounds and points to “metaphorical compounds”, such as *lady finger* (= ‘x that is like (a finger that is part of a lady)’) and *birdbrain* (= ‘someone who has as a significant part (a brain that is like that which is part of a bird)’), noting that “such composed compounds are of course semantically more complex ... and therefore more difficult to learn” (ms. p. 21). He goes on to speculate that compounding is an aspect of Bickerton’s (1990) protolanguage, but that the protolanguage is not a step on the way to “fully fledged language”, but rather the scaffold upon which fully fledged languages are built, with noun-noun compounding being a relic of the protolanguage with “only rudimentary grammatical structure, ... highly dependent on the pragmatics of the words being combined and on the contextual specifics of use”. This explanation reduces the otherwise implicit stipulation of a high degree of bilingualism that would be necessary for the mapping to be borrowed, as a higher level (syntactic/pragmatic) part of the grammar,¹⁴ in that there is a common ground cross-linguistically for the pattern of certain types of compounds, e.g., noun-noun compounds like Spanish *puerco espín* ‘porcupine’ (English < Fr.), literally “spine pig”. That aspect of the mapping may have no bearing on a borrowing. The idiosyncrasy of the association of meanings and forms, however, is an aspect of the mapping that is significant and does not fall under this simplistic part of the grammar.

The *blended derivative* is one type of morpheme substitution, exemplified by PaG *-ig* for English *-y*, giving *bassig* for ‘bossy’. Here again, the existence of a mapping within the lender language is definitive for inclusion in the category, although in this case, it is impossible to distinguish between the mappings, since they are completely congruent. The features for this would be mixed form, equal meaning and mapping [BF + LF, BM = LM, BA = LA] (see Figure 15 above). Note that this is the same as the **reverse substitution** below, (*ge-kick*), with a blending of morphemes. The only possible difference here lies in whether the morpheme is a free or bound morpheme. In either case, these are creations in which a loan root is “inserted” into a frame, or a native affix is appended to a loan root.

Blended compounds are Haugen’s last type of loanblend. The need for the existence of a mapping holds true although in some ways the source of the mapping is much more difficult to ascertain for words like PaG *blauməpai* ‘plum pie’, given not only the

¹⁴ Cf. Thomason & Kaufman 1988 and Weinreich 1953 for discussion of the relative ease or frequency of borrowing at different grammatical levels (e.g., lexicon, morphology, syntax).

transparency of the compound, but also Jackendoff's characterization above of some types of N-N compounds. The features for these are mixed form, equal meaning, and possibly equal mapping [BF + LF, BM (= LM), BA]. Here again, what his example shows is the insertion of a loan root morpheme into a borrower mapping pattern.

In all of Haugen's loan blends, what is significant is that the mappings are equal, a product of typological congruence between the languages involved. As will be seen below, for blends in Atepec Zapotec, the mappings are almost all from the borrower language. There are, however, loan blends in unrelated languages which have mappings from the lender language, as can be seen with the Japanese use of *-ade* to signify 'drink made from', borrowed from English *lemonade*, but whether the Japanese user sees this as an affix or a compound member is unclear.

LOANSHIFTS is Haugen's superordinate category for words that use semantic changes in native words to deal with the results of cultural and linguistic contact. These have borrower (or equal) forms, added lender meaning, possible loss of borrower meaning, and either borrower or lender mapping (see below for differences) [BF(= LF), (-)BM + LM, BA/LA]. He refers to this as 'substitut[ing] native morphemes' (1953:402).

Loan homonyms have equal forms, sometime loss of borrower meaning and added lender meaning [BF = LF, (-)BM + LM]. These have no semantic aspects in common with the native word, as with AmP *grosseria* 'rude remark', which is also now associated with the meaning 'grocery' for Portuguese-English bilinguals, based solely on the similarity of forms.

Loan synonyms (which Weinreich (1953) refers to as *polysemy*), have two subtypes, which add only a new distinction of meaning to the native word. The use of "synonym" here is misleading; although there may be some semantic overlap between the two meanings, there must necessarily also be some difference.

The first type of loan synonym is *semantic displacement*, and it is categorized as such on the basis of a high degree of similarity between the new and native phenomena. This is one kind of **semantic shift**, a less confusing term from Hock (1991) and Hock and Joseph (1996) that better describes the process of a form being mapped onto a new meaning and losing its original meaning. We can then contrast this with **semantic extension**, which retains the original meaning (see below). The features of a semantic shift are: borrower form and changed source of meaning [BF, LM – BM]. Haugen's example of this is the AmPort use of *pêso* 'weight' from Spanish *peso* to mean 'dollar' (although the use of *peso* for 'unit of money'¹⁵ or 'coin'¹⁶ is a nearly universal Iberian language usage). This is what I will call a **creation shift**, because the borrower language provides the mapping of possible polysemy [BF = LF, LM – BM, BA]. We must assume, because he does not explicitly say so, that *pêso* no longer retains its 'weight' meaning. A clearer example of the creation shift is Mayan *çih* 'deer' becoming 'sheep' through a process termed

¹⁵ Terrence Kaufman (p.c.).

¹⁶ Dicionários PortoEditora.

“marking reversal” by Witkowski and Brown (1980). A general problem with the category of creation shift lies in whether the shift occurred at the time of the borrowing, as could easily be the case with, e.g., American English *hoosegow* ‘jail’ from Spanish *juzgado* ‘judged’ (see Figure 10 above), or was the result of semantic drift after the initial borrowing as with Mayan *çih*.

The difference between creations and loans is partially consistent with Haugen’s categories and is easily distinguished, as can be seen in the feature bundles for **creation extensions** [BF, BM + LM, BA] and **loan extensions** [BF, BM + LM, LA], which differ only in the origin of the mapping. As an example of a creation extension, we can consider the case of AE *polecat*, which first applied to a large member of the weasel family, then went on to apply to ‘skunk’. The creation extension includes the subtype *loan homophone* which provides the stimulus of having equal forms, assuming an equal monomorphemic mapping (see above for discussion of homophony). Haugen’s example is AmP *grosseria* ‘rude remark’ adding the meaning ‘grocery’ [BF = LF, BM + LM, +BA]. The *loan homologue*, or loan extension, as with AmP *frio* ‘cold’, adds the meaning of the illness on the basis of the polysemy of English *cold* (borrower form, added lender meaning on basis of lender mapping) [BF, BM + LM, LA]. The difference in mappings here is that for *grosseria*, the addition of the meaning ‘grocery’ is based solely on the English form’s similarity to a Portuguese form and is strictly a creation with no connection to the English mapping. For *frio/cold*, however, the mapping is one of polysemy; two semantically unrelated meanings, ‘temperature’ and ‘disease’, are associated with the form, a mapping in the lender language which is then emulated in the borrower language (see Figure 11). Haugen (1953:400) ignores the origin of the mapping in this case, citing the polysemy of *cold* as the cause with no further analysis.

Semantic confusion, with the loss of borrower form, the addition of lender meaning and a borrower mapping [LF – BF, LM – BM, BA], is described by Haugen as the case when a native morpheme, on a homophonic basis, adds to its original meaning and the “native distinctions¹⁷ are obliterated through the influence of partial interlingual synonymy”, as in AmP *livraria* ‘bookstore, home library’ coming to include the meaning of English ‘library’ (*biblioteca* in Port.). Since what changes here is not the meaning, but rather the mappings of forms to meanings, I believe a more iconic name would be a **form shift**, to signify a meaning extension of a native form based on a lender mapping, with the concurrent loss of the original borrower form. Another example of this phenomenon, with no homophonous stimulus, is AmP *frio* ‘cold’ adding the meaning of the illness on the basis of English *cold*, and losing the Portuguese form *constipação*. In both cases, the mapping of the borrower form to the meaning is lost, and the (polysemous) lender mapping is substituted.

Loan translations (or **calques**) (borrower form, lender meaning and mapping) [BF, LM = BM, LA] are another type of loanshift, according to Haugen, and are defined as the importation of a particular structural pattern in the form of a non-compositional combination of two semantic elements. The idiosyncrasy of loan translations is an impor-

¹⁷ The distinction here is apparently one of the mapping of form-to-meaning, in which the form *biblioteca* is “obliterated”.

tant distinction to make clear, since the inclusion of transparent constructions, e.g., English *fat boy* ‘boy who is overweight’ as translated from S *niño gordo* (NOT the nuclear burden of the Enola Gay “Fat Boy”), would render this category nearly universal and thus generally vacuous. This is a case where a non-transparent mapping in the lender language is a defining factor. Haugen’s example of a calque is the canonical *skyscraper* → S *rascacielos*. It is important to note in this instance that there is a difference in the composition of the Spanish version, which does not use the agentive form, as the English *skyscraper* does. The Spanish version, literally translated, means “scrapes sky”; as a strict translation from E, it would be *rascador del cielo*. Thus, while this is a mapping borrowing, associating particular forms and their meanings, it is a blend of mappings (**calque blend**), since it is not free of derivational morphological trappings. Weinreich calls this a **loan rendition** (borrower form, loss of compositional borrower meaning, lender meaning and mixed mapping) [BF, LM – BM, BA + LA], based on the mapping, but not an exact **loan translation (calque)** (same features except for pure lender mapping) [BF, LM – BM, LA], for which he offers AmP *estar direito* ‘to be right’ after E, for which it should be noted, as above, that “right” as ‘correct’ would be idiomatic in Portuguese.¹⁸ (See Figure 12.) Weinreich (1953) adds the category **loan creation**, which to avoid confusion of terminology I call **loan mapping** (borrower form, equal meaning, lender mapping) [BF, BM = LM, LA], used to match designations in the contact language, such as Yiddish *mitkind* “fellow child” for ‘sibling’, where only the concept of a single word form for the meaning is borrowed.¹⁹ Haugen claims that there is only a difference of degree between a loan(blend) with a single borrowed element, as with Pa. German *blauməpai* ‘plum pie’, and a compound borrowing (calque) (1950:214). While it may appear that a difference in the number of what he terms “morphemic substitutions” is all that is taking place in these cases, it is also clear that for the calque, the compound is more definitively based on the lender mapping, while that of *blauməpai* could be either PaG or default universal (per Jackendoff).

‘**CREATIONS**’ is given by Haugen as a distinct category, separate from the borrowing process, coming into the borrowing language not as *direct* imitations of some item(s) in the lender language, but as innovations dealing with stimuli from the lender culture. Romaine (1988:56) categorizes Haugen’s term “creations” as a subset of loan shifts, and says that (unspecified) others have labeled these “loan translations” or “calques”. As shown above, there is a subset of loanshifts (the creation shifts) that do indeed share the definitive feature of this category, the borrower mapping. If these were calques, however, they would have a lender language mapping, and Haugen’s stipulation for inclusion in the category of creations is that the mapping and form (not his words) are from the borrower language. His (1956) example for a pure or **induced creations** (borrower mapping and form, lender meaning) [BF, LM, BA] is Pima “having downward [grain] tassels” for ‘oats’.²⁰ Another Pima creation is *wuhlo ki’iwia* “burro eats” for

¹⁸ Dicionários PortoEditora.

¹⁹ Note the caveat above (fn. 1) regarding the opacity of morphemic composition to speakers. Thomas Stewart (p.c.) also points out the utility of having a single word, which I regard as a stimulus in the same way that homophony might be a stimulus, with the utility of being close to a native form.

²⁰ From Herzog 1946. No Pima forms were given by Haugen.

‘oatmeal’.²¹ If these were calques, we would expect either a Spanish or English mapping, and neither exists. He also refers to *reverse substitutions* (mixed form, with added borrower meaning based on a borrower mapping) [BF + LF, BM + LM, BA], in which loan roots are filled into native mappings, as with PaG *Ge-kick* (*Ge-HAB* + English *kick*) ‘habitual kicking or objecting’. This is the same thing as the blend derivative, which must also qualify as a creation, except that the addition is of a root rather than a derivative morpheme.

Haugen further offers the term *hybrid creation* (mixed form, lender meaning, borrower mapping) [BF + LF, LM, BA] for Yaqui *lios-nooka* ‘pray’, from Spanish *dios* ‘god’ + Yaqui *nooka* ‘to speak’, to distinguish it from induced creations, noting that it “cannot have come into being as [an] imitation”(404). This is a **creation blend**. There must be an assumption here that the loan form has not become completely integrated into the borrower language to the point that it is no longer seen as a foreign word; this is the problem of delineating the moment/decade/generation in which a loanword is no longer perceived as “foreign” in order to accurately distinguish hybrid creations from loan blends. Although there are other Yaqui words, such as *hiosia nooka* ‘read’, literally “paper speak”, which demonstrate that this is based on a Yaqui mapping, the difference in the agentivity of the verb makes it plausible that *lios-nooka* was not a native creation, but was a loan concept, created on an imperfect Yaqui syntactic/semantic model by a Catholic priest to distinguish Christian from “pagan” prayer, in which case it might be a true **loan blend**.

5 Why features?

If we apply these features to Haugen’s 1953 taxonomy, as above in Table 2, it becomes clear that the use of features does not simplify the typology. On the contrary, it shows that Haugen’s categories lack clear distinctions. Table 2 above showed his descriptive criteria for each term. Table 3 below lists the features of each category to show the lack of coherence in the taxonomy.

The criteria change between divisions, using form and meaning in the first half of the chart (and the first half of Loanshifts), through “synonymous loanshifts”. For the remaining categories, the “arrangement” (mapping) becomes the defining factor and the others are essentially ignored. In many ways, the only consideration given is to the lender language, as though the borrower language were of secondary importance, the not particularly noteworthy vessel for these otherwise fascinating phenomena.

²¹ Saxton & Saxton 1969.

Haugen (1953)	Description/stimulus		Features
LOANWORDS	= phonology and morphemes		
Loanwords A	wholly morphemic import		LF, LM, LA (default)
Loanblends	partial morphemic import		
stem B	mixed monomorpheme		BF+LF, MB=LM, BA=LA
derivative C	mixed types of morphemes		BF+LF, MB=LM, BA=LA
compound D	mixed free morphemes		BF+LF, MB=LM, BA=LA
LOANSHIFTS	borrower word changes meaning		+LM
Extensions	semantic loan (added meaning)		BM+LM
homologous E	stimulus	= form, ≈ meaning	LF=BF, BM+LM, BA
homophonous F		= form	LF=BF, BM+LM, BA ²²
synonymous G		≠ form, ≈ meaning	LF-BF, BM+LM, BA
Creation (calque)	imported arrangement		LA
literal H	identical arrangement		BF, BM=LM, LA
approximation I	approx. arrangement		BF, BM=LM, LA+BA
CREATIONS			
Induced creation J	imported meaning only		BF, LM, BA
Hybrid creation K	imported meaning, mixed form		BF+LF, LM, BA

Table 3. Features applied to Haugen's taxonomy and category descriptions.

Haugen's description of the difference between the superordinate category of Loanshifts and the subordinate category of Induced Creations is that of the difference between changes in the meaning of the borrower word and a borrower word (form) having only an imported meaning. What we have here is a change in a form's association with one or more meanings, which is a quantitative (gradient) distinction. On the opposite side of the "contrast", within Loanshifts, there are two types of change in meaning. One type borrows an idiosyncratic polysemous lender mapping. The other type creates an idiosyncratic polysemous (borrower) mapping for a borrower form based on the semantic closeness of a novel phenomenon to something already associated with the form. There is thus overlap between Haugen's Loanshifts and Creations categories.

It is also possible for a novel meaning to be associated with a lender form but with a borrower mapping, which uses lender forms and their meanings in a novel way to form a sort of reverse calque. An example in AZ is *lasu kabrestu* 'halter (for horses)' ← Spanish *laso* 'loop' + *cabestro* 'halter', with a native superordinate-subordinate (head-first) compound using loan words for both parts. (This is equivalent to *tuna fish* in English, with its head-last syntax.) It is not a loanword, but is a creation using loanwords. Neither Haugen's Loanwords nor Creations categories offer a clear place for this type of construction.

²² Note that the homologous and homophonous categories have equal feature specifications. This is a result of the fact that Haugen makes a gradient distinction between "similar" and "same" meaning.

There are also cases where a lender form changes its meaning association, as happened with *hoosegow* (cowboy English for ‘jail’ from Spanish *juzgado* [xusgao] ‘judged’), where it is clearly a loanword, but also a shift, in which case there is not a clear boundary between these two categories either. On these bases, then, there are no clear delineations between any of the superordinate categories, which means that there is no significant categorial advantage to the use of Haugen’s typology.

A second problem with Haugen’s typology is that the characteristics used to define each type are not consistent throughout the typology. As noted above, meanings, in and of themselves, cannot be imported. Haugen appears to equate morphemes with what we would have to call the morphological structure *in combination* with meaning. This can then be combined with the phonology to create the word. The form is specifically important to loanwords and loanblends, but apparently only peripherally so if it is native, and is ignored for calques (although perhaps only as a default). There is no clearly defined difference between full phonological and partial morphemic import (loanblends), and imported meaning and partial form (hybrid creations). Perhaps most importantly, what Haugen refers to as the “arrangement”, which must be an aspect of the mapping, is used only in reference to the calques although it is important in distinguishing between types of loanshifts.

A third problem lies in the gradience of his characteristics. In several of his categories, he uses descriptions like “wholly, partially imported” and “partial form vs. root form”, and particularly “equal vs. approximate”, all of which present problems of degree. The features given here, as noted above, are privative, and thus lend themselves to non-gradient application in statistical analysis, if not for definition.

The distinctions made within his superordinate categories are useful only within those categories. Furthermore, because the categorial boundaries are inexact, these distinctions must also be inexact. In contrast, not only do many of these distinctions fall out naturally from a featural analysis, as with the difference illustrated below between literal calques and loan renditions (lender vs. mixed mapping), but more distinctions of import are drawn, as shown with the difference between extension types, loan and creation, based again on the mapping. In addition, Haugen’s categories as such are shown to have incomplete featural coherence; even his two types of blends are separated across an arbitrary division. Features, however, work across all of Haugen’s superordinate categories.

If we assume a default model of the mapping, like that given above for *taco* (one form to one meaning), it is reasonable to limit it to a one-dimensional representation. If we keep in mind the (generally) emic nature of borrowing, a one-dimensional representation is adequate for the vast majority of loan phenomena; in the earliest stage of contact a form will be borrowed according to the phonotactics and phonemic inventory of the borrowing language. If, however, a borrowing exhibits evidence of a more-than-one-dimensional mapping, it must be the result of a greater familiarity with the lender language than obtains with what Thomason and Kaufman (1988:74) call casual contact. Those forms that become associated with an additional meaning based on the lender

mapping (AmP *frio* adding the disease meaning) will be evidence of a high degree of bilingualism and will not be from the initial stages of contact. Mixed mapping is a grouping independent of form or meaning which is absent in Haugen’s typology, but which is significant in demonstrating evidence of non-linguistic factors with strictly linguistic evidence.

Haugen’s taxonomy (Table 3 above), with changes and additions as noted above, is represented in Table 4 below as a three-dimensional featural matrix. Those terms followed by a letter represent his terms as shown in Table 2. The terms without a letter designation are designations of words with feature bundles that do not appear in his taxonomy, and they are discussed below.

It is important to note that because of the three dimensions used in this typology, there is no reasonable binary division to be made, as there was with Haugen’s taxonomy. In keeping with the idea of using features, we can draw parallels to the idea of natural classes, i.e., all items with lender mapping, or all items with added meanings, etc. Just as with a phonemic inventory, any division that attempts to divide by a single dimension can create a class, but it ignores the other classes that can be formed across the division. Voicing is binary, but a typology split along that dimension would ignore the similarities between alveolar consonants or between fricatives.

		FORM	MEANING		
			Lender	Mixed	Borrower
MAPPING	Lender	Lender	loanword A		
		Mixed		loan compound D loan derivative C	
		Borrower	calque J	<i>loan extension</i> H <i>loan shift</i> I	
	Mixed	Lender		loan homologue G	
		Mixed	loan homophone F	blended stem B analogue E (loan homonym)	
		Borrower		loan rendition K	<i>loan mapping</i>
	Borrower	Lender		semantic confusion	
		Mixed	<i>created blend</i> M (hybrid creation)	reverse substitution N <i>creation compound</i> M (blended compound)	
		Borrower	induced creation L (<i>creation</i>)	<i>creation extension</i> H <i>creation shift</i> I	native vocabulary

Table 4. Featural distribution of Haugen’s types.²³

²³ Note that the bold italicized terms are mine; the bold terms are Weinreich’s.

When we compare the distribution of the types of phenomena in terms of their features with their distribution in terms of Haugen's categories (from either 1950 or 1953), no clear pattern of features-to-categories emerges; the features for extensions and shifts are divided into two mapping groups, and blended forms occur with both lender and equal meanings (and will occur below with borrower meanings). This distribution is a clear demonstration of the lack of coherence of the categories given in previous typologies.

We can now turn to applying this type of categorization to examples from AZ to demonstrate its usefulness in examining the results of lexical contact within a particular language.

6 Atepec Zapotec origins and contact with Spanish

AZ is a variety of Zapotec, a member of the Zapotecan branch of the Otoman-guean family of languages of Mesoamerica. AZ is generally a head-first language (VSO, NA, NPoss), and as such, its nominal syntax generally agrees with that of Spanish. Verbs have TMA prefixes, but the system is relatively straightforward. As of this writing, only one Spanish verb has been found that is fully morphologically incorporated into the language. These facts will have a bearing on the determination of which language supplies the mapping for each contact-induced phenomenon.

Atepec is a village high in the Sierra Juárez mountain range in the northern part of the state of Oaxaca, Mexico. While it has had contact with the Spanish language and aspects of Spanish culture for at least 450 years, its relative isolation before 1957, when a road was built through the Sierra Juárez, had served to prevent the kind of contact that fosters widespread bilingualism. A historical distrust of anyone from outside the pueblo was another factor in this.

The current situation has changed, with compulsory education in Spanish for all children. One woman there with teenage children of her own complained that this was detrimental to their ability with AZ. The acquisition of electricity has also brought increasing contact with television and radio. Announcements made over the village's loudspeaker system are now given in Spanish. I have also found evidence of a generational difference of simplification in the morphophonology of the deictic proximal suffix.

7 AZ lexical contact phenomena

We now discuss lexical contact phenomena in AZ in a demonstration of the featural typology given above, following the types of phenomena through the three features in a geometric fashion. As was demonstrated in Table 3, there are many featural combinations which have not been discussed in the previous literature. For those new types which appear below, I offer a coherent terminology to represent them in as iconic a manner as possible. Of the three dimensions available, the one that offers the greatest

challenge to the taxonomies given above is that of the mapping, so on that basis, I begin with the categories given under the lender (Spanish) mapping.

7.1 Lender (Spanish) mapping

7.1.1 Lender form

Within lender forms with lender mappings, the first category expected would be that of the simple loanword. As noted above, monomorphemic forms (or for the borrower language speaker, apparently monomorphemic forms) must be assumed to have equal mappings, and could therefore fall under the category of mixed mapping. However, I will assume that the fact that the lender word offers a monomorphemic mapping is the reason for the equality of the mappings and include loanwords in the category of lender mapping. Given that, we begin with lender forms, moving through mixed to borrower forms, and following the same progression within each form category with the meanings. We can begin with the sampling of representative early loanwords in Table 4.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
gázhú ²⁴ [ɣáʒú]	<i>ajo</i> [áʃo]	garlic
kuléká	<i>clueca</i> [kluéka]	brood hen
ánjeli [ánxeli]	<i>angel</i> [ánxel]	angel
kumárí	<i>comadre</i>	godmother
gutzílu	<i>cuchillo</i> [kučilo]	candle
kurûtzi	<i>cruz</i> [krus]	cross

Table 4. Loanwords [LF, LM, (LA)].

In these words, the meanings are Spanish (all are imports), and the forms also are originally completely Spanish, disregarding AZ phonological adjustments, which are an important aspect of borrowing but outside the scope of this paper (however, see fn. 17).

Next, we can move to those loanwords in Table 5 that occur because of what Weinreich (1953) called “necessity”, i.e., to fill an apparent gap in the lexicon, upon which basis I call them **gap loans**. It should be noted here that “necessity” is a misnomer, since it is clear that means other than borrowing (i.e., creations) can be used to fill the gap.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
gwinda	<i>guinda</i>	burnt red color (of livestock)
demasiádú	<i>demasiado</i>	very much, a whole lot
dilijensia	<i>diligencias</i> ‘investigation’	investigation
krióyú	<i>criollo</i>	native (to the town)

Table 5. “Gap” loans [LF, BM, =LM, LA].

²⁴ In these forms, several processes of adaptation are occurring: ‘garlic’ [aʃo] *gázhú* has acquired an initial [ɣ]; ‘brood hen’ *kuléká* shows metathesis of the non-native diphthong and consonant cluster, and ‘godmother’ *kumárí* has simplified the consonant cluster, while non-low final vowels are uniformly raised.

From these examples, we can see that although the *significatum* for each existed for AZ, the sign for each is either an AZ form or mapping replacement, or a novelty.

The set of examples in Table 6 is for lender forms which replace or compete with borrower forms, which we can designate the **loan form shift**.²⁵

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
kampaníyú	<i>campanillo</i> ‘little bell’	brown-backed solitaire (bird)
kuyóté	<i>coyote</i>	coyote (AZ <i>tzawiyó</i>)
jagwar	<i>jaguar</i>	jaguar

Table 6. Loan form shifts [LF, (–BF), BM, LA].

For these words, the fact that all of these animals have a distinct and salient characteristic (e.g. the brown-backed solitaire has a “distinctive bell-like call”²⁶) means that they are easily recognizable and would therefore have been named. It may well be that gap loans and replacements should be included under one category, since for *demasiádu* ‘too much’ (or any of the others in Table 6), there may have been some other word/construct that constituted a sign for it. In some cases, it might be possible to look for related forms in related languages to determine whether it is likely to be a replacement.

The example in Table 7 is a form that does not compete with a borrower form for a particular meaning but replaces the superordinate term (in this case ‘green’ *yáʔá*) with the Spanish form.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
berde limón	<i>verde limón</i>	lime green

Table 7. Gap loans [LF, –BF, BM, LM, LA].

This example is one in which the form is completely Spanish, and in which the mapping for the syntax is indeterminable; it is equally Spanish and Zapotec, however, for different two reasons, the first of which is the fact that the lime is an import, and the second of which is the use of ‘lime’ (an import) to describe a certain shade of green. For colors, other than five primary names (‘red’, ‘yellow’, ‘white’, ‘black’, and ‘green’ in AZ), the differences between languages in color terms and what might be included in the range of any particular term constitute a mediating concept (cf. Berlin & Kay 1969).

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
pítu kanúá	<i>pico canoa</i> ‘canoe beak’	green toucan

Table 8. Form shift calque [–BF, LF, BM].

²⁵ To keep the terminology consistent, the term “loan form shift” will automatically designate the loss of the borrower form.

²⁶ Schoenhals 1987.

The example above in Table 8 is an example of a **form shift calque**, with a Spanish mapping.²⁷

7.1.2 Mixed forms

We can now turn to the category of mixed forms with a lender mapping, which in AZ contains no words of Spanish-only meaning. We begin, therefore, with those in which the meanings are equal, on the assumption given above that the *significata* exist in both cultures.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
kustiyú kia?	<i>costilla</i> ‘rib’ + ‘my’	my wife
laya? jwísíú	‘tooth’ + <i>juicio</i> ‘wisdom’	wisdom tooth
wê? ...kwerda	‘to give’ + <i> cuerda</i> ‘cord’	to encourage s.o. to speak
tsi?nu ora	‘twelve’ + <i> hora</i> ‘hour’	noon

Table 9. Calque blends [BF, LF, BM, =LM].

Of the **calque blends** in Table 9 the first three are clearly metaphors of varying abstractness. The last, *tsi?nu ora* ‘noon’ ‘twelve hour’, is clearly based on the Spanish concept of time, and the fact that it exists in conjunction with an AZ word *lawi? tsá*, ‘middle [of] day’, is a clear demonstration that a lexeme existed previous to Spanish contact, but the synonymy of these two forms may differ in punctuality.²⁸

We can now look at one type of extension of meaning (**loanblend extension**).

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
ébèkkíá ... fotu	‘to take out’ + <i>foto</i> ‘photo’	to take a photograph

Table 10. Loanblend extension [BF, LF, BM, +LM].

In this case, the mapping seems to be Spanish, and the meaning of *ébèkkíá* has been extended to include the taking of pictures, perhaps on the basis of the similarity of the meanings of Spanish *sacar* ‘to take out’, although it is possible that the metaphor of ‘take out’ (after being put in) is responsible for a parallel use in this case. Although French and English both use the same lexeme for general ‘take’ and ‘take (a picture)’, ‘take out’ *vyázo* is also used in modern Greek,²⁹ and *снять* ‘remove’ is used in Russian, so this metaphor may actually be a common means of communicating the idea of taking pictures cross-linguistically.

²⁷ The phonemic replacement of /k/ with [t] is probably indicative of a phonological constraint against [kVk] which appears in only two words in the language, at least one of which is sound symbolic (onomatopoeic) *kukuí* ‘nightjar’ (bird).

²⁸ Mitla Zapotec (Stubblefield & Stubblefield 1991) and Isthmus Zapotec (Pickett 1965) both have similar native forms “middle of day”.

²⁹ Brian D. Joseph (p.c.).

There are also cases, as in Table 11, in which the lender meaning of the word can be lost, either through misunderstanding of the meaning, as likely occurred with *hoose-gow* (see Figure 9 above) or through semantic drift after borrowing.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
gútè?...kwenta	‘to give’ + <i>cuenta</i> ‘report’	to turn in; to betray; to accuse

Table 11. Loanblend shift [BF, LF, BM, –LM].

The final word in the category of mixed forms with lender mappings is the **calque blend** for a native meaning, as in Table 12.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
kuttsí ettsé?	<i>cochi</i> ‘pig’ + ‘spine’	Mexican porcupine

Table 12. Calque blend replacement [BF, LF, BM].

The mapping for this word must be Sp. *puerco espín* ‘spine pig’, since *kuttsí* is a loan word ‘pig’ from *cuchi/cochi*, Mexican Spanish for ‘pig’, probably from Peninsular Sp. *cochino*. English uses the same mapping, from French, which we can contrast with ‘hedgehog’, an English native creation.

7.1.3 Borrower forms

The remaining category of form is that of native forms. I found only one type of AZ native form that used a Spanish mapping, and it is a loan extension, essentially the same as the loanblend extension given above, except that the complement of the verb in the blend is the only one that fits the meaning of that extension. In all of those in Table 13, the meaning range of the borrower word is extended to match (part of) the meaning range of the lender word, using the mapping of more than one meaning to a single form.

<i>AZ form</i>	<i>Original meaning</i>	<i>Spanish meaning(s) added to AZ form</i>
íchittha	to raise, lift (<i>levantar</i>)	to conduct (a census); to give/bear (false testimony); to take minutes/document
íthella	to command (<i>mandar</i>)	to send (a letter)
lla?	leaf (<i>hoja</i>)	sheet (of paper)
kíxá ... ló	to lay down	to establish (law)

Table 13. Loan extension [BF, BM, +LM].

As a recapitulation of the categories within the lender mapping, Table 14 shows the two-dimensional matrix of types. Note that of the twelve types listed, only the two shaded (loanwords and loan mapping) appear in previous taxonomies, and only one (loan concept) has no apparent AZ representative. The lack of words with a Spanish mapping, an AZ form, and a strictly either Spanish or AZ meaning is certainly not surprising, although in part, this could be due to the default assumption of form and meaning of monomorphemic words being from the same source. In a lender language that uses stative verbs instead of adjectives, for instance, there would be a clear difference between

mappings. In this case, however, it would be unlikely, bordering on impossible, for the form to be strictly borrower, since with the change in grammatical category, we would expect some morphological accoutrements to accrue to the borrower word, thereby mixing, if not entirely shifting, the mappings.

FORM	MEANING			
	Lender	Mixed (=BM)	Mixed (+BM)	Borrower
Lender	Loanwords	Gap Loans Form Shift	Loanword Extension	Form Shift Calque
Mixed		Calque Blend	Loan Blend Extension	Loan Calque Blend Loanblend Shift
Borrower		Loan Concept	Loan Extension	

Table 14. Lender mapping categories [LA].

7.2 Blended or indeterminate mapping

We can now turn to the general category of words with a blended or indeterminate mapping. The first member of this could be loanwords, but as given above, unless there is a compelling reason to believe that the mapping is from the borrower language, e.g., prefixation, I assume that the mapping is the lender language. None of the words of this class have a form that is strictly of lender or borrower, although it might be possible. In AZ, two of the words in Table 15 with an indeterminate (=) mapping are strictly of Spanish meaning, the **loanblends** ‘barbed wire’ (‘wire’ itself is a loanword) and ‘wheat’.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
alambre yettsé?	<i>alambre</i> ‘wire’ + ‘spine’ ³⁰	barbed wire
zhúá? xtíla	maize + (<i>Ca</i>) <i>stilla</i> ‘Spanish/foreign’	wheat

Table 15. Loanblend [BF + LF, LM, BA = LA].

Its counterpart in Table 16, a **blend form shift** with a borrower meaning, is assumed to have lost the original borrower form by which this native plant was known, since as mentioned above, sheep are imports.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
ìyyà zhubàrà karnérú	‘flower’ + ‘tail’ + <i>carnero</i> ‘lamb’	lamb’s tail (<i>Sedum spp.</i>)

Table 16. Blend form shift [BF + LF, BM, BA + LA].

The construction of the compound is essentially a *bahuvrihi*, referring to a plant that has ‘a flower [like] the tail of a lamb’, and is thus AZ, as with *ìyyà wellařàré?è* ‘heavenly blue morning glory’ “flower [like] broken pitcher”, but the metaphor upon which it is based is clearly Spanish and thus the mapping is a blend.

³⁰ This may also be translated directly from the Spanish *alambre de pua* ‘wire of spine/thorn’.

Another type of blend with a mixed mapping is one in which the mappings and the meanings are equal, as in Table 17, which we can call a **super blend**:

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
gútè?...mensaje	‘to give’ + <i>mensaje</i> ‘message’	to give a message

Table 17. Super blend [BF + LF, BM = LM, BA = LA].

The next category, in Table 18, is what Weinreich (1953) called loan renditions, i.e., calques in which both the lender mapping, a metaphor, and the borrower mapping, in these cases the syntax and/or morphology, are used to convey the lender meaning. To maintain coherence in the terminology, I call it a **blend rendition**.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
gúni...bwelta	‘to make/do’ + <i>vuelta</i> ‘turn; occasion’	to go for a walk
thú...ttu nesesidad	‘to have’ + ‘a’ + <i>necesidad</i> ‘poverty’	to be in poverty
éttíá lista kí?	‘call out’ + <i>lista</i> ‘list’ + GEN	to call the roll

Table 18. Blend rendition [BF + LF, BM = LM, BA + LA].

The first is based on Spanish *dar vuelta* “to give a turn”, close to English *to go for a spin*. The verb has been changed in this case, which may reflect either imperfect understanding (although see Table 27 below for *attu bwelta*, which offers a different meaning associated in Spanish with *vuelta*), or a more iconic verb to replace the idiomatic ‘to give’. The second may be based on *estar en una necesidad* “to be in a necessity”, and has changed the verb, although it is also possible that *necesidad* was borrowed by itself, and this verb more closely follows the AZ mapping for conveying the meaning. The third is adding a meaning to ‘to call out’ *éttíá* on the basis of Sp. *pasar lista*, ‘to pass [through] the list’, and again is changing the verb to match the AZ mapping.

Table 19 demonstrates the mixed mapping categories that appear in AZ. The assumption that monomorphemic forms indicate the origin of the mapping accounts for the lack of category fillers in the lender and borrower form rows. However, the possibility of their existence, while perhaps unlikely, cannot be ruled out, as was noted above in the description of the lender mapping table.

FORM	MEANING				Borrower
	Lender	BA≠LA	BA=LA		
		(BM=LM)	(BM+LM)		
Lender			Homologue		
Mixed	Loanblend Homophone		Superblend	Analogue	Blend Replacement
Borrower		Blend Rendition			

Table 19. Mixed mapping [BA = LA or BA ≠ LA].

7.3 Borrower mapping

We can now examine the “creation” end of the contact spectrum. As above, we begin with lender forms.

7.3.1 Lender forms

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
lachasuéla	<i>la hacha</i> ‘the axe’ + <i>azueta</i> ‘adze’	adze
lásu kabréstú	<i>laso</i> ‘loop’ + <i>cabresto</i> ‘halter’	halter

Table 20. Loanword creation [LF, LM, BA].

In these examples, the AZ mapping of a superordinate term, ‘axe’ in the first example, with a defining term ‘adze’, is evident, even though the form and the meaning is clearly Spanish, based on the importation of steel tools.

<i>AZ</i>	<i>Spanish form and gloss</i>	<i>Current AZ gloss</i>
duëndé	<i>duende</i> ‘goblin’	malignant
kosku, josku	<i>josco, hosco</i> ‘dark red color of animals’	well, very well; of beautiful color
kwáyú	<i>caballo</i> ‘horse’ [kaβáyo]	colt
lúkkú	<i>loco</i> ‘crazy’	surly, “snooty”
maski?	<i>mas que</i> ‘more than’	don’t do it!; it doesn’t matter
rruínu	<i>?ruín</i> ‘ruin’	affected, simpering
(ka) uxtísia	(PL) + <i>justicia</i> ‘justice’	municipal authorities

Table 21. Shifted loanwords [LF, BM – LM, BA].

Shifted loanwords (Table 21) are based on an AZ mapping. In these words, although the semantic connections in all but the fifth example are clear, there has been a loss of original meaning, which, as noted above, is impossible to pin down chronologically. The phonology of *maski?* is such that the form must be borrowed.³¹ The last example is clearly based on the borrower mapping because the optional use of the plural proclitic *ka* gives the word the same meaning, which can be contrasted with the word for ‘justice’ *la?uxtísia* NOMINALIZER ‘that which is’ + *justicia* ‘justice’.³²

The same diachronic ambiguity holds true for the **loanword extension** in Table 22, although under the circumstances of early contact, ‘foreign’ could refer to anything European. The evidence that this is an early loan, aside from the phonological aspect mentioned above regarding *zhúa xtíla* ‘wheat’, is that the word for ‘strawberry’ *dígá?ekstranjeru* “foreign (black)berry” uses *ekstranjeru* (← Spanish *extranjero*) to designate ‘foreign’.

³¹ There are no native words in AZ that are *m*-initial or that contain the consonant cluster *sk*.

³² The nominalizer *la?*- (*la?go* ‘food’ ← *go* ‘to eat’) is close to Spanish *la* (feminine definite article), but the use of *ka* (PL) ahead of it without the glottal stop shows that there is no constraint against vowel contiguity, and thus that *la?* is probably not a confusion.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
xtílá	Castilla	Spanish, foreign (European)

Table 22. Loanword extension [LF, LM + BM, BA].

7.3.2 Mixed forms

We can turn now to borrower creations with mixed forms. We begin with lender meanings as in Table 23 of **blended loans**.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
kuttsí kwíní	<i>cochi</i> ‘pig’ + ‘small & fat’	pig (small breed)
kuttsí lúlaʔá	<i>cochi</i> ‘pig’ + ‘Oaxaca’	pig (large breed)

Table 23. Blended loans [BF + LF, LM, BA].

In this next category, Table 24, we have a **creation blend shift**, where the lender meaning has been altered. The phonology of these terms ensures that they are loans coupled with suffixes, and the meanings of the loanwords have been lost. One can only speculate on the change in the semantics in the first.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
nékútó ³³	<i>conejo</i> ‘rabbit’ + DIM	‘daddy’ “little rabbit”
sópaní	<i>sopas</i> ‘sops’ + “done/made” suffix	soaked, steeped

Table 24. Creation blend shift [BF + LF, BM – LM, BA].

The next division is that of the **creation blend**. The first category within it, in Table 25, is one in which the meanings are equal and the mapping is one of using the form as a direct object. The first two verbs here are evidence of the AZ propensity to use a general + specific term, as is the case above in the loanword creations. The last three follow a more general (and cross-linguistically common) pattern, using a helping verb combined with the nominal form of the lender verb in order to convey the meaning. AZ natively uses this construction with nouns and deverbalized adjectives.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
éyakka ... arrepentir	‘to compose (oneself)’ + <i>arrepentir</i> ‘to repent’	to regret; to repent
gúdètà?...lístá	‘to make inclined’ + <i>lista</i> ‘ready’	to make ready
gáppá...interés	‘to have’ + <i>interés</i> ‘interest’	to be interested
gúni...prueba	‘to make/do’ + <i>prueba</i> ‘proof’	to test, to prove
gúni...remédíu	‘to make/do’ + <i>remedio</i> ‘remedy’	to cure

Table 25. Creation blend compounds [BF + LF, BM = LM, BA].

³³In *nékú-tóʔ*, the *k* would be geminate after the first root vowel, as would the *p* in *sópaní* if these were AZ forms. Because AZ has no tri-syllabic roots, in borrowing [konexu] ‘rabbit’ *nékú*, the first syllable was dropped. ‘Soup’ in AZ is *indate*.

Table 26 below covers creation blend derivatives. For these words, the derivational morphology is all AZ, and this verb is the only fully incorporated Spanish verb in the lexicon, something telling in its own right.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
fwértèní	<i>fuerte</i> ‘gravity’ + ‘done’	serious
tántuání	(<i>en</i>) <i>tanto(que)</i> ‘inasmuch as’ + ‘done’.	inasmuch as
gú-kodia	POT+CAUS - <i>joder</i> ‘to “screw”’ (fam.)	to be “screwed” up

Table 26. Creation blend derivatives [BF + LF, BM = LM, BA].

The last category of creation blends (Table 27) is what I term here **semantic reduplication**. Each term consists of the Spanish word (in most cases, a conjunction) followed by its AZ synonym. There is no pattern in Spanish for compounds like these, but there are certain AZ words that use a pattern of reduplication to indicate a sort of limit. Two verbs demonstrate a multi-morphemic version of this pattern: *gúduathua* ‘to fill to the point of forming a meniscus’ ← *gú+dua* ‘to put’, and *étze?étze?* ‘to reconcile’ ← *étze?* ‘to meet’ (an ultimate meeting of the minds). Reduplication as an intensive is a common pattern in, e.g., Jamaican Creole³⁴ *blæk* ‘black’, *blæk blæk* ‘very black’, and the most intense would be the limit.

<i>AZ form</i>	<i>Spanish form + AZ morpheme</i>	<i>English gloss</i>
desdebá	<i>desde</i> ‘out of’ + <i>bá</i> ‘from’	out of, from
asta?na	<i>hasta</i> ‘until’ + <i>na</i> ‘until/since’	until, till
áttu buéltá	<i>áttu</i> ‘other/again’ <i>vuelta</i> ‘time/occasion’	again
para ki?ní	<i>para que</i> ‘so that’ + <i>ki?ní</i> ‘because’	so that
porki?ní	<i>porque</i> ‘because’ + <i>ki?ní</i> ‘because’	because
sin ki?ní lă	<i>sin que</i> ‘without’ + <i>ki?ní lă</i> ‘so that NEG’	without

Table 27. Semantic reduplication [BF + LF, BM = LM, BA].

7.3.3 Borrower meanings

We can now look at blends created as signs for borrower meanings. The data in Table 28 follow the head-first pattern of AZ, so the mapping is clearly native. The surprise in this group is the fact of its existence, with (in almost all cases) the subordinate qualifier as a loanword, which constitutes a kind of form shift. From the phonological evidence, most of these are recent, with consonant clusters not found in AZ native words. The only exceptions to having an AZ head are marking reversals (see above), which occur only with fully (literally overwhelmingly) incorporated loanwords. These are the last two items in the group. One is *mizhí ìxxì?* ‘bobcat’ “forest/wild housecat”, which coexists with the fully AZ *betziagá*, and the other is *kuttzí ìxxì?* ‘collared peccary’ “forest/wild pig”. Both pigs and housecats (but not peccaries or bobcats) are imports, and the use of loanwords for native fauna is unexpected.

³⁴ Gooden 2003.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
bedaʔ zhubàná eskalérá	‘fox’ + ‘tail’ + <i>escalera</i> ‘ladder’	civet cat
béera paisán	‘chicken-sized fowl’ + <i>faisán</i> ‘pheasant’	great curassow (bird)
bèllà fléchá	‘snake’ + <i>flecha</i> ‘arrow’	speckled racer (snake)
bèllà kwártá	‘snake’ + <i>cuarta</i> ‘quarter’	whipsnake
bèllà lechèrà	‘snake’ + <i>lechera</i> ‘bullsnake’	bullsnake
betzuʔtíʔ máchú	‘dung beetle’ + <i>macho</i> ‘male’	rhinoceros beetle
bezhitùʔ tájú	‘coatimundi’ + <i>atajo</i> ‘small group of animals (livestock)’	coatimundi (group type)
beʔyá benénú	‘mushroom’ + <i>veneno</i> ‘poison’	type of deadly mushroom
binní bíntú	‘(small) bird’ + <i>pinto</i> ‘painted/spotted’	black-and-white warbler
binní órá	‘(small) bird’ + <i>hora</i> ‘hour’	type of wren
binní xkèʔè kwáyú	‘(small)bird’ + ‘dung’ + <i>caballo</i> ‘horse’	bronzed cowbird
dǎ (xkèʔè) kunéjú	‘bean’ + (‘dung’) + <i>conejo</i> ‘rabbit’	pinto bean
exxubólá	‘avocado’ + <i>bola</i> ‘round’	avocado type
exxumáchi	‘avocado’ + <i>machín</i> ‘spider monkey’	avocado type
ìyyà kampáná	‘flower’ + <i>campana</i> ‘bell’	“bell flower”
ìyyà kartúchú	‘flower’ <i>cartucho</i> ‘cartridge’	arum (flower)
ìyyà kaskabel	‘flower’ + <i>cascabel</i> ‘rattle’	woolly senna (flower)
ìyyà kwarésmá	‘flower’ + <i>cuaresma</i> ‘Easter’	poinsettia
ìyyà nánchí	‘flower’ <i>nanche</i>	pickle tree flower
sópa etta	<i>sopa</i> ‘soup’ + ‘tortilla’	type of soup
tzúkiʔ lèʔè bintu	‘mid-size bird’ + ‘belly’ + <i>pinto</i> ‘spotted’	orange-billed nightingale- thrush
wèlaʔ ya nuésí	‘caterpillar’ + ‘tree’ + <i>nuez</i> ‘nut’	tufted caterpillar
wèlaʔ ya umbrílú	‘caterpillar’ + ‘tree’ + <i>membrillo</i> ‘quince’	type of caterpillar
ya nuésí	‘tree’ + <i>nuez</i>	walnut tree
ya sédrú	‘tree’ + <i>cedro</i>	“cigar-box” tree
zhubàná tizhérá	‘tail’ + <i>tijera</i> ‘scissors’	earwig (insect)
kuttzí ìxxiʔ	<i>cochi</i> ‘pig’ + ‘wild’	collared peccary
mizhí ìxxiʔ	<i>mistón</i> ‘cat’ + ‘wild’	bobcat

Table 28. Form shift blends [BF + LF, BM, BA].

7.4.1 Lender Meaning

The first category in the group of borrower mapping with lender meanings is that of the **creation shift** as seen in Table 29. In this case, a borrower word shifts its meaning to a lender meaning, from generic ‘animal’ to the imported ‘horse’. In this case, the likelihood is that the quintessential animal is the largest and thereby most salient, and on that basis, this category could be considered another form of marking reversal. It is probable

that there are other creation shifts that are historically opaque because of the near-synonymy of meaning.

<i>AZ form</i>	<i>Spanish form</i>	<i>English gloss</i>
biaʔ	‘animal’ ³⁵	horse

Table 29. Creation shift [BF, LM – BM, BA].

The next category is that of the **compound creation**, in which only borrower resources are used to create a sign for a lender significatum, as in Table 30. ‘Mule’ is straightforward in its construction. ‘Fig’ appears to be a description of the shape of the avocado combined with the size and internal distribution of guava seeds. The ox is “that animal which scratches/plows the field”, since plowing is another import to a slash and burn agricultural tradition and is not done without draft animals. There is a Spanish mapping for ‘beasts’, *ganado mayor*, which means ‘large livestock’, but the description used in AZ is clearly a native mapping, consisting of a list of prototypical members as opposed to a description.

<i>AZ form</i>	<i>Morphology</i>	<i>Morphology gloss</i>	<i>Gloss</i>
biaʔwéguʔ	biaʔ + wéguʔ	‘animal/horse’ + ‘fat’	mule
exxuwí	exxu + wí	‘avocado’ + ‘guava’	fig
guʔná	gu- + aʔná	‘animate’ + ‘to plow/scratch’ ³⁶	ox, bull
guʔnábiaʔ	guʔná + biaʔ	‘ox’ + ‘horse’	beasts

Table 30. Compound creation [BF, LM, BA].

7.4.2 Mixed meanings

The next category is that of mixed meanings. The first case, the **creation extension**, is one in which the meaning of a borrower word is extended on the basis of similarity between the new *significatum* and the *significata* covered by the existing word, as in Table 31. As above, the mapping is assumed to remain the same (borrower) in the absence of any evidence to the contrary.

<i>AZ form</i>	<i>Original meaning</i>	<i>Added meaning</i>
gú, gû	sweet potato	potato (<i>papa</i>)
gúxataʔ	to flatten/smash	to iron (<i>planchar</i>)
íthi ... (žítsiʔ)	to wring (breast, teat)	to milk (<i>orderñar</i>)
ìyyà	metal	syringe (<i>jeringa</i>)
kíttśá	to thunder	to shoot (a gun) (<i>disparar</i>)
tсени	indolent, apathetic	scarecrow (<i>espantapájaro</i>)

Table 31. Creation extension [BF, BM + LM, BA].

³⁵ Based on Fernández de Miranda 1995 and Nellis & Nellis 1983.

³⁶ Terrence Kaufman (p.c.).

The potato was imported from South America by the Spaniards, and as such constitutes an item of lender meaning. Regarding *gúxata?* ‘to iron’, AZ has other verbs for smoothing, but they involve cutting or abrading. ‘To milk’, ‘syringe’, and ‘to shoot a gun’ are clearly imported concepts.

7.5 Borrower meanings

This category covers the rest of the native vocabulary.

BORROWER MAPPING RECAPITULATION

Table 32 illustrates the categories involved in the contact phenomena with a borrower mapping.

FORM	MEANING			
	Lender	Mixed (=LM)	Mixed (+LM)	Borrower
Lender	Loanword Creation		Form Shift (-BF)	Shifted Loanword Extended Loanword
Mixed		Creation Blend Compound/ Derivative/Stem	Loan Homonyms (BF=LF)	Form Shift Blends (-BF) Creation Blend Shift
		Semantic Reduplications		Reverse Substitution
Borrower	Compound Creation		Creation Extension	Native Vocabulary
	Creation Shift			

Table 32. Borrower mapping table.

The shaded areas here are those that in some way (in some cases only by implication) are included in the taxonomies offered in the previous literature. As noted above, for AZ/Spanish contact, the chance of loan homonyms and the closely parallel semantic confusion is smaller (= 0) than for most of the languages used as exemplars in Haugen, Weinreich, Romaine, and Hock & Joseph, most of which are genetically related and thus are inclined to having forms from common roots that have undergone semantic drift in different directions since genetic branching occurred. Note that if the mixed forms in Table 33 were to be split into two categories (equal forms and truly mixed forms) as the mixed meanings are, there would be no occurrences of equal meaning and equal form, nor of added meaning and mixed form. This results from the way the mapping is defined. If an added lender morpheme (to account for the mixed form) were to add a meaning to a borrower word based on an indisputably borrower mapping, as opposed to creating a meaning, the mixed form/added meaning category could be filled, but we would not expect to see equal form and equal meaning based on an indisputably borrower mapping

except by the previously stated default rule for monomorphemic classification (and extremely close-to-synchronic genetic branching).

Table 34 below gives the revised version of the featural matrix, with blanks darkened and the previous types of phenomena mentioned in Table 1 shaded.

		MEANING				
		FORM	Lender	Equal	(Added)	Borrower
MAPPING	Lender	Lender	Loanwords	Gap Loan/ Replacement	Loanword Extension	Form Shift Calque
		Mixed		Calque Blend	Loan Blend Extension	Calque Blend Loanblend Shift
		Borrower		Loan Concept	Loan Extension	
	=	Lender		Homologue		
		Mixed	Loanblend Homophone	Superblend	Analogue	Blend Form Shift
	≠	Borrower		Blend Rendition		
	Borrower	Lender	Loanword Creation		Semantic Confusion (-BF)	Shifted Loanword Extended Loanword
		Mixed		Creation Blend Compound/ Derivative/ Stem Semantic Reduplications	Loan Homonyms (BF=LF)	Created Blends (-BF) Creation Blend Shift Reverse Substitution
		Borrower	Compound Creation Creation Shift		Creation Extension	Native Vocabulary

Table 34. Full featural matrix.

8 Feature future

There are several areas of further investigation with regard to featural analysis. As can be seen in Table 34, there are a few empty categories in the matrix. For those words for which the mapping is equal (as with the kinds of compounds Jackendoff (in prep.) referred to as possibly universal) we might well expect to find representative members of these subcategories. However, the distinction in these cases is problematic, since it is much the same as the default assumption above for monomorphemic cases; the mapping is assumed to be from the same source as the form. For the other, more random,

blanks, it will require investigation cross-linguistically to see if there are examples in other contact situations, and if not, to determine why these are not represented in contact phenomena.

The use of the mapping would be explanatory in discussing, e.g., the composition of compounds. An example from the AZ/Spanish data, *lasu kabrestu* ‘halter’, demonstrates a typical native left-headed compound (contrasting with right-headed English *tuna fish*) using loan words for both members. It is not a loanword, but is a creation. The use of a lender mapping in some cases might be the definitive factor in anomalous compositions such as the reversal of the headedness of compounds. These occur in English atypically left-headed *court(s)-martial* and *attorney(s) general*, both borrowed from French, although whether as calques or directly borrowed is difficult to determine, and they contrast with typically right-headed *divorce court* or *general practitioner*. The same occurs in Vietnamese native vs. “Sino-Vietnamese” compounds, which tend to switch from left- to right-headedness, possibly based on lender influence (Stewart 2000).³⁷ These demonstrate that the mapping at the lexical level is an important aspect in more than just calques.

A historical and diachronically sociolinguistic analysis of lexical contact results would be facilitated by the inclusion of the mapping, as well as the other patterns, in the investigation. Perhaps the most ambitious project in terms of scope, with regard to the frequency and type of borrowing/creation, is that of Brown (1994), who looked for the existence of some 77 European lender *significata* (e.g. ‘wheat’, ‘horse’, ‘hour’) in over 200 languages of the Americas. He separated the creations from the borrowings and examined that distinction to see what kinds of correlations existed between languages as, e.g., a possible indicator of the type of contact. Although not explicitly stated, we can assume that most of these were monomorphemic lexemes in the various languages that now include signs for these *significata*. For the remainder, however, examination of the use of mixed native and lender resources (e.g. AZ *zhúá?xtíla* ‘wheat’ “Spanish maize”) could provide an additional means of analysis, in conjunction with knowledge of the duration, of the intimacy of contact at the time of the borrowing, and conversely, when that time was.

The breakdown of features could also make possible a statistical analysis of the types of phenomena found in relexified languages such as Media Lengua (cf. Muysken 1997), mixed languages like Michif (cf. Bakker & Papen 1997), and creoles and pidgins to study how each mapping is used in conjunction with the forms and meanings to give indications of, e.g., the origins of treatments of grammatical categories. This same kind of model may also be useful in looking at shift, at whatever grammatical level.

9 Conclusion

I have demonstrated that the characteristics used by Haugen and others to taxonomize the lexical results of language contact fail in four ways.

³⁷ Stewart (2000) notes, however, that certain compounds in Vietnamese fail to follow the etymologically based reversal pattern.

- a. These characteristics are inconsistently used and are applicable for description only within specific superordinate categories; they do not apply across all data.
- b. They fail to adequately differentiate categories and types, separating phenomena with common aspects and conflating types with different aspects.
- c. The characteristics are at least partially gradient in nature, with the result that they cannot apply evenly across the data, nor can they offer the possibility of quantitative analysis.
- d. In many cases, the characteristics given refer to the stimuli for the phenomena rather than to the sources of the linguistic aspects of the phenomena.

The features chosen for the analysis here are based on a semiotic approach to the lexicon and consist of the meaning, which is referentially fixed, the form, which is not, and the mapping of the relationship between the form and the meaning. While they do not allow for dividing the data into hierarchical categories, they do apply across all data, and because they are emic in nature, they do so evenly. This allows for the arrangement of these phenomena in a three-dimensional matrix. These features eliminate the conflation of types and offer, minimally, an example of clear distinction of non-phonological, and thus emic, evidence of extensive bilingualism. Furthermore, they can be used in a model to represent the process of borrowing, with attribution of each feature to its relevant source(s).

References

- BAKKER, PETER, and ROBERT A. PAPAN. 1997. Michif: A mixed language based on Cree and French. *Contact languages: A wider perspective*, ed. by Sarah G. Thomason, 295–363. Amsterdam: John Benjamins.
- BERLIN, BRENT, and PAUL KAY. 1969. *Basic color terms: Their universality and evolution*. Berkeley: University of California Press.
- BRODY, JILL. 1987. Particles borrowed from Spanish as discourse markers in Mayan languages. *Anthropological Linguistics* 29.4.507–21.
- BROWN, CECIL H. 1994. Lexical acculturation in Native American languages. *Current Anthropology* 35.2.95–117.
- DIEBOLD, A. RICHARD, JR. 1961. Incipient bilingualism. *Language* 37.1.97–112.
- GOODEN, SHELOME. 2003. Reduplication in Jamaican Creole: Semantic functions and prosodic constraints. Twice as meaningful: Reduplication in Pidgins, Creoles and

other contact languages (Westminster creolistics series 8), ed. by Silvia Kouwenberg, 93–103. London: Battlebridge Publications.

- HAUGEN, EINAR. 1950. The analysis of linguistic borrowing. *Language* 26.210–31.
- HAUGEN, EINAR. 1953. *The Norwegian language in America: A study in bilingual behavior*. Philadelphia: University of Pennsylvania Press.
- HOCK, HANS HENRICH. 1991. *Principles of Historical Linguistics*. Berlin & New York: Mouton de Gruyter.
- HOCK, HANS HENRICH, and BRIAN D. JOSEPH. 1996. *Language history, language change, and language relationship: An introduction to historical and comparative linguistics*. Berlin: Mouton de Gruyter.
- JACKENDOFF, RAY. In preparation. The ecology of English noun-noun compounds. Oxford: Oxford University Press.
- LYONS, JOHN. 1977. *Semantics*. Cambridge: Cambridge University Press.
- MUYSKEN, PIETER. 1997. Media Lengua. *Contact languages: A wider perspective*, ed. by Sarah G. Thomason, 375–426. Amsterdam: John Benjamins.
- NELLIS, NEIL, and JANE GOODNER NELLIS. 1983. *Diccionario Zapoteco de Juárez: Ca Titsa' Qui' Riu*. Serie de Vocabularios y Diccionarios Indígenas “Mariano Silva y Aceves” Num. 27. México: Summer Institute of Linguistics.
- ROMAINE, SUZANNE. 1995. *Bilingualism*. Oxford: Blackwell.
- SAXTON, DEAN, and LUCILLE SAXTON. 1969. *Dictionary: Papago and Pima to English; English to Papago and Pima*. Tucson: University of Arizona Press.
- STEWART, THOMAS W., JR. 2000. A hole in the head parameter? Element ordering in compounds. Columbus: The Ohio State University, MS.
- STUBBLEFIELD, MORRIS, and BARBARA DE HOLLENBACH. 1991. *Diccionario Zapoteco de Mitla*. Mexico, D.F.: Instituto Lingüístico de Verano.
- THOMASON, SARAH GREY, and TERRENCE KAUFMAN. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- WEINREICH, URIEL. 1968 [1953]. *Languages in contact: Findings and problems*. Paris: Mouton.
- WITKOWSKI, STANLEY R., and CECIL H. BROWN. 1983. Marking reversals and cultural importance. *Language* 59.3.569–82.

AN EVALUATION OF GERMAN-CROATIAN CONTACT

Mark E. Nuckols

Abstract

This paper is a study of the influence of German on Croatian. It attempts to provide a historical background and to summarize and evaluate the linguistic findings of some scholars in the field. The study focuses mainly on the period 1526–1918, when the Kingdom of Croatia, Slavonia, and Dalmatia was under the political control of the Habsburg Empire, and it is also limited to the contact in those areas of the Croatian-speaking world that were under Habsburg rule, i.e. Croatia and Slavonia, not Dalmatia. I consider the socio-historical context of the contact and the history of the Croatian literary language before examining specifically the results of contact which are visible in the Croatian language of today. In evaluating the results of contact, I draw largely on the criteria developed by Thomason and Kaufman (1988), as well as on the work of other scholars and my own observations. Although the influence of German on Croatian is almost exclusively lexical, calquing from German is extensive and points to a higher degree of contact than might be expected: The large number of loanshifts and loanblends indicates a higher degree of bilingualism than pure loanwords would suggest.

1 Introduction

Historically, both South and West Slavic languages have been involved in language contact of one sort or another with neighboring non-Slavic languages. Under the Byzantine and Ottoman Empires in the Balkans, the predecessors of modern Bulgarian,

Macedonian, and Serbian felt the influence of Greek from the time of the conversion of the South Slavs to Christianity in the ninth century until the conquest of Constantinople by the Ottoman Turks in 1453; then Turkish influence on the South Slavs became felt as the Ottoman Empire expanded northward and westward up the Balkan peninsula until roughly 1683, when it suffered a major defeat in its failed siege of Vienna, after which Turkish military domination of the region waned until the First World War. This military and political domination placed several nationalities under the auspices of the same rulers, not to mention the fact that many of the contiguous peoples lived in ethnically mixed areas, intermarried, and many of their members lived a nomadic shepherding lifestyle that brought them into frequent contact with speakers of various neighboring languages. This contact among speakers of the Balkan languages was sustained and intense enough to produce results typical of a Sprachbund. There have been several detailed studies (Schaller 1975, Joseph 1983) and several less detailed studies (Décsy 1973:105–23, Birnbaum 1965) of a Balkan Sprachbund involving Bulgarian, Macedonian, Albanian, Rumanian, and Greek, as well as the (at least limited) participation of Serbo-Croatian (namely the southernmost Torlak dialects), Arumanian, and other languages or dialects.

There has, however, been fairly limited discussion of the possible existence of a central European Sprachbund. Gyula Décsy (1973:87–105) writes of a “Danubian league” (*Donau-Bund*) involving the West Slavic languages Czech and Slovak, the western South Slavic languages Slovenian and Croatian (but not Serbian), and the neighboring Finno-Ugric language Hungarian, citing as causes of the development of this league geo- and socio-political conditions present in the Habsburg Empire from roughly the sixteenth century to 1918. Décsy also proposes that there are numerous other Sprachbünde on the European continent, although he does not suggest that any of them overlap. Eric Hamp (1979, 1989), however, does suggest that some European Sprachbünde overlap or coincide and that all the languages of the Balkans have been involved in one type or another of contact, though often with different neighbors, thus creating a “cluster” or “crossroads of Sprachbünde”, as suggested in the titles of his 1979 and 1989 articles respectively. Given the recent political fragmentation of the Balkans, which has produced the burning question of whether the language once known widely as “Serbo-Croatian” is really one language, two languages (Croatian and Serbian), or four (add Bosnian and Montenegrin), it is worth considering the different spheres of linguistic, cultural, and political influence or dominance that these two varieties of Serbo-Croatian underwent.

Croatian provides an interesting subject of study as a language involved in contact for another reason: it and its various dialects have been influenced and enriched by languages as diverse as Italian, German, Hungarian, and to a lesser extent, in its Bosnian dialects (many ethnic Croats in Bosnia still identify their language as Croatian), Turkish. The present study is limited to the influence that German has exercised on Croatian, which is indeed a large undertaking in itself, and this examination is by no means exhaustive; instead, this paper attempts to provide historical background and to summarize and evaluate the linguistic findings of some scholars in the field. The study focuses mainly on the period 1526–1918, when the Kingdom of Croatia, Slavonia, and Dalmatia was under the political control of the Habsburg Empire, which in differing periods within this time

frame could be called variously the Holy Roman Empire (until 1806), the Austrian Empire (1806–1867) and Austria-Hungary (1867–1918). It is also limited to the contact in those areas of the Croatian-speaking world that were under Habsburg rule, i.e., Croatia and Slavonia, but not Dalmatia. It also excludes the Croats of Austria proper (namely Carinthia) who were a minority living among German speakers, which is quite a different situation from Croatia and Slavonia where Croats were the majority. I consider the socio-historical context of the contact and the history of the Croatian literary language, before examining specifically the results of contact that are visible in the Croatian language of today.

In evaluating the results of contact, I draw largely on the criteria developed by Thomason and Kaufman (1988:74–76), but I also make some suggestions and additions to the factors they consider, drawing on the work of other scholars and, to a lesser extent, my own observations. The influence of German on Croatian, it turns out, is almost exclusively limited to the lexicon, but the phenomenon of calquing from German is extensive and points to a higher degree of contact than might be realized by simply considering “lexical influence” without regard to the type of influence. As I argue later, the large extent of loanshifts (which include semantic extensions based on a foreign model as well as loan translations or calques) and loanblends (which contain an admixture of native and imported elements)—including many at the phrasal level—tends to indicate a higher degree of bilingualism than pure loanwords¹ would suggest. Some of the calquing between German and Croatian is syntactic in nature, which suggests that some grammatical changes in the structure of Croatian began developing under the influence of a limited number of semantically related lexical items. Joseph (1983:191–93) has suggested that finite complementation, one of the most notorious features of the Balkan Sprachbund, may well have become diffused in a like manner. Since this paper is limited to examining contact between two languages only, it cannot make a conclusion either about the existence of a central European Sprachbund or about whether Croatian should be considered a member (either core or peripheral), but it can make some determination of whether German-Croatian contact was sufficiently intimate in nature and intense in degree and whether it involved enough mutual bilingualism to be considered Sprachbund-like. Additional research considering the level of influence between Croatian and Hungarian would have to be completed before coming to any conclusions about the presence of Croatian in a central European Sprachbund, but it is hoped that the present examination will be a worthwhile endeavor in that direction.

2 Political history

The Triune Kingdom of Croatia, Slavonia, and Dalmatia became part of the Kingdom of Hungary shortly after its native dynasty, the Trpimirovići, died out in 1091. In August 1526, the Croats again found themselves without a king when the Hungarian king Louis II Jagellon died on the battlefield at the Hungarian town of Mohács in a historically decisive contest with the Ottoman Turks. The Ottomans then pushed northwards to Buda and Pest, and their conquests led to the 150-year Turkish occupation

¹ Here, I am using the terminology of lexical contact phenomena in accordance with the classification of Haugen (1953) as found, with slight modifications, in Winford 2003.

of a large central swath of Hungary. The Jagellons had several years earlier made a pact with the Habsburgs that if the male line of one of the families should die out, the other family would take over their dominions. According to this treaty, the Habsburgs should have automatically become Hungarian sovereigns, but the matter still awaited a vote by the Hungarian nobility. The nobility elected a rival but was eventually forced to recognize the Habsburg dynasty; the Croatian diet, however, had supported the Habsburgs all along, and, in any event, most of them already considered the union with the Hungarian Crown of St. Stephen to have been dissolved by virtue of the fact that the Hungarian rulers had died out. The question of whether Croatia fell under the immediate rule of Vienna or of Budapest was never fully resolved before the dissolution of the Habsburg monarchy in the aftermath of World War I.

Croatia and Slavonia were under the administrative auspices of Vienna (Dalmatia had come under Venetian control in some areas and Turkish in others), and this administration was very centralized at first, since the Imperial government in Vienna wanted efficient coordination in terms of military efforts against the Turks. This centralization eased in the eighteenth century, following definitive military victories over the Ottomans in the late seventeenth century, although Croats, according to Banac (1984:231) were subject to another round of “absolutism and harsh Germanization that followed the defeat of the 1848 revolutionary wave”. The *Ausgleich* or “Compromise” of 1867 created the “Dual Monarchy” in which all matters pertaining to both Austria and Hungary were handled at the “imperial” level, and all matters affecting only Hungary were now handled at the “royal” level; i.e., Hungary remained a quasi-independent kingdom within the empire of Austria-Hungary. Hungary now enjoyed the right to determine her own internal politics in matters such as economy and education, while matters of defense and foreign relations were dealt with from the imperial capital of Vienna. The Croats, however, were not as fortunate as the Hungarians in terms of recognition of the historical rights, and many felt that, after the *Ausgleich*, they were exposed to Hungarian political domination and a strong dose of “Magyarization” or assimilation to Hungarian language and culture.

3 Socio-historical background

Décsy (1973:87–89) lists Serbo-Croatian as a member of what he considers to be a *Donau-Bund* or “Danubian League” which additionally includes Czech, Slovak, Hungarian, and Slovenian. Among the features of this area, Décsy claims, are initial stress, vowel quantity, minor role of diphthongs, lack of vowel reduction, and a strong tendency toward prefixation (“*grosse Präfixfreudigkeit*”). Although it is beyond the scope of this paper to examine in detail the possibility of Croatian’s membership in any proposed Sprachbünde, it is worth noting that Décsy mentions the influence of German in central Europe without including it in his *Donau-Bund*. Décsy recognizes the status of Latin and German as supranational languages and the deeply felt influence of German speakers on Czechs and Slovenes, yet he maintains that German influence on the Slovak and Croatian was relatively “negligible” (“*geringfügig*”—1973:88). The word “negligible” would seem to understate the degree of German influence on Slovak and Croatian, since, as he notes (89), German, alongside French, was the language of the Hungarian high nobility in the eighteenth century. The fact that the lower nobility tended to speak the local

language—Slovak, Magyar, or Serbo-Croatian—does not detract much from the influence of German (or other languages) on the local languages in the Kingdom of Hungary. If Décsy's observation on the lower nobility is correct, then it likely means that they preferred to speak the local language in their households and neighborhoods. No doubt these members of the lower nobility were by and large educated, and by the eighteenth century that education was acquired largely through German as a language of instruction. The lower nobility, then, seems to have represented a transitional group between the highest classes, which had a strong preference for German, and the lowest classes, which, in our area of focus, which would have displayed a strong tendency to speak Croatian.

In the whole area of the Donau-Bund, the middle class citizens of the towns and cities had been of German origin (*deutschstämmig*) since the Middle Ages, and from the eighteenth century onwards there were islands of rural German population. Furthermore, German served as the “common second language” of the Habsburg Monarchy from the sixteenth century to 1918. Décsy writes further that there was an “Austrian commercial dialect” which counted as a standard language, although it contained “observable deviations from the High German of non-Austrian lands”. As we shall see later, there are lexical items in colloquial Croatian that indeed likely came from colloquial Austrian. The presence of German-speakers, especially of economically and politically powerful ones, certainly would have provided a strong motivation for the population of Croatia and Slavonia to learn German, but these facts do not suffice to establish how many Croats knew German or how well.

Kessler (1981:159) finds, much like Décsy, that German was the common language of the middle and upper classes of the entire Habsburg monarchy (at least those who thought themselves “better”), and it was not just in Croatian-speaking areas that such people thought of the local language as a “lingua exotica”, a language to be used with the servants or a language of the “vulgar class”. “The German language”, he writes, “was in this case a status language, not a mother tongue in the emotionally loaded, nationalistic sense”. Kessler also notes that one of the leaders of the Illyrian movement responsible for codifying a united Serbo-Croatian language in the middle of the nineteenth century, Ljudevit Vukotinović, complained that “There are people who speak our [Croatian/local] language only when they are forced to, for instance with peasants or one's own servant”, but there were even some “especially in higher society who are ashamed to have a servant who is so primitive that one cannot” give him orders in a foreign language (i.e., not “our” language, Croatian). If Kessler paints an accurate picture here, then there clearly was a substantial amount of bilingualism in the Kingdom of Croatia and Slavonia by the nineteenth century, among Croats first and foremost, and to a lesser extent among German-speakers. It would appear that the higher the class (indeed, the more breeding a servant had), the greater the preference for German, though at some levels, the upper classes did not altogether disdain speaking the local language.

If Guldescu (1970) is correct in many of his observations on the state of Croatian-Austrian relations from the fourteenth to the sixteenth centuries, then there probably was even earlier language contact that occurred through the military. He notes: “Since the

fourteenth century the Habsburgs had ruled the partially Croatian province of Istria, although Venice had taken over the western shore of this peninsula. Croatian soldiers had long served, too, under the Counts of Görz (Gorica) who were Habsburg vassals ... ” (1970:15). There likely was a significant amount of language contact through Austrian-Croatian intermarriage. Guldescu further speaks of “inter-marriage and land inheritance on the part of the two nobilities”, remarking that “Military service under the Habsburg banners, the settlement of Austrian artisans and peasants in Croatia, and Croatian migrations to Austria effected a blending of the two populations also. Indubitably there is a heavy admixture of Austrian blood in the veins of many Croats today.”

Piškorec (1997:35) notes the participation of German-speaking population groups from the thirteenth century in the founding of Croatian towns and points to a special role for the German language that came with the establishment of a military border zone for defense against the Ottomans in the first half of the sixteenth century. His study considers this role in the town of Đurđevac, in the northeastern-most part of today’s Central Croatia, a few miles from the present-day border with Slavonia—and historically belonging to Slavonia—and just a little farther from the border with Hungary, but considerably distant from the modern borders of the German-speaking world. He observes that the military importance of the town shrank significantly after the border with the Ottoman Empire was “pushed back” by Austrian military successes late in the seventeenth century (1997:36). German nevertheless remained as an administrative language until 1871. Piškorec makes the noteworthy observation that “unlike other Croatian areas and towns, where an observable proportion of the population can be shown to have spoken German as a mother language, most of the inhabitants of Đurđevac were Croats of long-standing Croatian origin who first took up the German language on their educational and career paths” (1997:37). The first German-language school in Đurđevac was built between 1756 and 1759, and German was first used in religious instruction in the 1780s in “Slavic-German” schools. Priests appointed by the Zagreb bishop were, “as a rule”, Croats who had studied abroad and were competent in Latin and German (39). Additionally, knowledge of both spoken and written German was necessary for promotion in the military (41). Piškorec also notes that there were probably a relatively small number of people (apparently he means in the second half of the eighteenth century, but the context is not entirely clear), 16–60 individuals, who were fully bilingual and thus served as the most important vehicles of the German-Croatian contact in Đurđevac.

Piškorec’s observations are enlightening but, of course, do not give an entirely clear picture of the contact situation in the rest of Croatia and Slavonia. The fact that the town in his study was Slavonian and, as he notes, not originally under terribly strong German influence, would tend to coincide with Hamp’s (1989) observation (discussed below) that Slavonia experienced some German influence, but that it was not as heavy as that experienced by those Croatian areas farther west, i.e., closer to German-speaking Austria.

While the evidence cited above does point to a German dominance over Croatian, as opposed to a contact situation of equal bilingualism, it would be wrong to discount the German-Croatian interaction described here as not of a Sprachbund nature on that basis

alone.² As Thomason (2001:107–8) has noted, the Ottoman Turks exercised political and military domination in the Balkans, and their language contributed hundreds, if not thousands, of lexical items to the general Balkan vocabulary, yet Turkish does not display any of the typical Balkan features. Greek is traditionally considered a member of the Balkan Bund, although often treated as peripheral, since it has fewer structural Balkanisms than the core members Albanian, Bulgarian, Macedonian, and Rumanian. Nevertheless, Greek, especially in the Byzantine period when the Hellenic world exerted a very strong cultural and political influence over the Balkans, contributed large numbers of words to the Balkan lexicon. Perhaps German, in the eighteenth and nineteenth centuries, played a role in central Europe analogous to that of Greek in the Byzantine period and Turkish in the period of Ottoman domination of the Balkans?

In any event, in the socio-linguistic situation in Croatian-speaking areas, there is no evidence of language **shift**. Certainly neither of the groups abandoned its native language in favor of the other language. However, bilingualism was clearly present in Croatia and Slavonia, although the degree of bilingualism is a more complicated matter. We are left, then, to consider how this bilingualism affected Croatian (or German), but first, we should take up the status of Croatian as a literary language in order to avoid the misconception that German was so dominant in all spheres of (at least public) life, that it was such a prestige language, that among ethnic Germans and Croats, only the latter bothered to learn or speak the other's language.

4 History of the Croatian literary language

The fact that German enjoyed the status of a prestige language and an administrative language after 1526 should in no way be taken to mean that Croatian was excluded from usage in literature and administration. A consideration of the history of literary Croatian is in order here.

To begin understanding the use of Croatian, or dialects thereof, as a literary language requires us to digress a bit historically from the time of Habsburg rule in Croatia and Slavonia, to “begin at the beginning”, so to speak. The role of the Cyrillo-Methodian mission is crucial to this sociolinguistic history. The year 864 is traditionally cited as the date when two Slavic-speaking brothers, Saints Cyril and Methodius, were summoned from their hometown of Thesaloniki in Greece (then part of the Byzantine Empire) to Christianize the West Slavs of the Great Moravian Empire, situated in what is now the Czech Republic and Slovakia. Part of their mission involved the translation of biblical and liturgical works into the Slavic vernacular, which for the most part still enjoyed relative unity in the Slavic-speaking world in the ninth century. The language employed by the brothers, which had some South Slavic features, has come to be known as Old Church Slavonic (OCS) or Church Slavic and was even accepted by the Roman pontiff for use as a liturgical language, though its status as a liturgical language has wavered over the

² Of course a Sprachbund situation requires three participating languages at the least. This study merely considers the intensity of the contact between these two languages. Hungarian-Croatian contact could easily involve another paper of this length. Note my comment on the number of languages involved in my discussion of criteria for evaluating contact phenomena early in §5.

centuries, especially after the Great Schism which divided Eastern and Western Christianity in 1066.

Additionally, St. Cyril devised an alphabet that was deemed to be better suited to the inventory of Slavic speech sounds than either Latin or Greek letters. That alphabet, as most (but not all) scholars agree, was the Glagolitic alphabet, which differs radically from either Latin or Greek. The alphabet that now bears his name, Cyrillic, is generally—but not entirely—agreed to have been invented by his disciples, and relies largely on Greek prototypes. Glagolitic has enjoyed a long-standing tradition in Croatia, one that has continued even to the present day, though with decreasing usage over the centuries and increasingly limited to the monasteries of the Istrian peninsula and the Dalmatian islands. Banac (1984:200–201) delineates the “Golden Age of Croatian Glagolitic” as roughly 1075–1475. After the Great Schism, however, OCS was no longer as sacred, and features of the local vernacular soon began creeping into the church language of Croatia, while at the same time, Croatia’s thorough identification with the Church of Rome meant increasing competition with Latin in ecclesiastical usage.

Even in Croatian secular society in the Middle Ages, many writers were clearly “Latinists”, meaning they preferred to write in Latin, although most of them did not altogether disdain the Croatian vernacular. Some writers of the late Middle Ages wrote mainly or exclusively in Croatian, while others attempted to write prose and poetry in both Latin and Croatian. By the fifteenth century, secular writing was almost exclusively in Latin letters, and the coastal town of Dubrovnik had become a literary center, so its local dialect became very influential in the development of the Croatian standard language. At the same time, a strong Italian influence entered into the literary culture of Dubrovnik following the acquisition of the Dalmatian communes by the Venetians in the late fifteenth century (Banac 1984:203–4).

Additionally, the Venetian prelates began resisting the use of the Slavic liturgy in the churches. Although the continued use of the Slavic liturgy was permitted, the Council of Trent (1545–47, 1551–52, 1562–63), convened in response to the Protestant Reformation to solidify Catholic dogma and practice, determined that liturgical services should be held in the ancient languages, with as little acquiescence to vernacular usage as possible, which had as a consequence a greater disparity between the OCS used in the traditional Slavic liturgy and the contemporary Croatian vernacular.

Early attempts at standardization of a Croatian literary language can be said to have begun roughly in the sixteenth century, though these attempts were neither as deliberate nor as political as those of the nineteenth century. The main obstacle to the formation of a single literary language was the existence of three main dialect groups for Croatian: the Čakavian, originally spoken in Istria and many Dalmatian islands; Kajkavian, indigenous mainly to Slavonia, including Zagreb; and Štokavian, originally spoken south of Kajkavian and east of Čakavian. (The names for these dialects come from the words for “what” as reflected in each of the dialects, although this is not necessarily the best criterion for dividing them.) The dialect question was further complicated by migrations, most notably those precipitated by the Ottoman conquests in the Balkans.

The sixteenth to nineteenth centuries also saw competition between the three scripts—Glagolitic, Latin, and Cyrillic—with Glagolitic gradually pushed back into Istria and restricted to Church usage, Latin winning out almost everywhere, and Cyrillic being used almost exclusively by Orthodox Slavs.

The nineteenth century saw the birth of the Illyrian movement (1835–48) which sought as its primary goal to unite all Slovenes, Croats, and Serbs. This involved the endeavour to devise a single literary language for the three groups. Although the attempt to include Slovene in this single, unifying language failed, a Serbo-Croatian literary standard was finally agreed upon. This standard was the result of numerous compromises, not simply that Serbs would use Cyrillic script and Croats Latin characters, but also that certain regional variations would be permitted. Štokavian had already gained tremendous ground in the eighteenth century, and in the nineteenth century it became the literary norm, due partly to the numerical superiority of its speakers in Croatian dialects alone, not to mention the fact that Serbs are overwhelmingly Štokavian speakers; thus the use of Štokavian was a very natural compromise.

Clearly, then, there was a continuous Croatian literature, as works were written and published in Croatian from the Middle Ages and with increasing frequency to the present day. In 1604 the first Croatian grammar, *Institutionum linguae illyricae*, was published by a Jesuit named Bartol Kasić. The seventeenth century also saw the translation of the Bible into modern Croatian idiom, and Jesuits worked at translating texts like the catechism into Croatian. Later, in the eighteenth century, works such as the sermons of Stefan Zagrebec in 1715–34, and Hilarion Gasparoti's *Czvet Szveteh* ('Blossom of the Saints', a hagiographical work) appeared (Hadrovics 1985:143). Katicić (1984:274) observes that by the middle of the eighteenth century

the already existing and largely unified Štokavian literary language in the east and south began to cover a wider range; work on the normalization of orthography and terminology was intensified, grammars were written and served as textbooks in schools. The penal code of Joseph II was translated into this language and was used in the courts for some time.

So there was a substantial literature published even before the Croats's national awakening in the nineteenth century, and as we can see from the above quote, there was significant use of Croatian in the schools and courts in the eighteenth century. We thus are faced with a complex situation in which it was advantageous, as we saw above, for Croats seeking positions in the military, the clergy, or trade to know German, yet German was not so dominant that no one in the upper classes bothered to learn Croatian, and Croatian was even beginning to enjoy some status in both educational and legal systems. These facts, of course, still say little about the spoken use of the two languages in more intimate settings (a point to which we shall turn later, citing the nature of certain lexical borrowings as evidence), although they do indicate that there was present on Croatian territory more than a mere "one-way bilingualism", a term used by Thomason and Kaufman (1988:95) to contrast with the "mutual bilingualism and multilingualism" of Sprachbund situations. The bilingualism in the area under study, however, was probably never fully

mutual, though it most likely did involve rather intimate contact, as we shall see from examples of the types of lexical borrowings found in the following section.

5 Degree of influence: An analysis of grammatical and lexical borrowing

In evaluating the contact situation involving Croatian and German, it is necessary to devise some criteria by which to judge certain results of contact. I outline in brief here the scale of five levels of intensity of contact devised by Thomason and Kaufman (1988: 74–76) before suggesting some important additional considerations. The levels or categories are listed as follows:

1. Casual contact: lexical borrowing only
2. Slightly more intense contact: slight structural borrowing
3. More intense contact: slightly more structural borrowing
4. Strong cultural pressure: moderate structural borrowing
5. Very strong cultural pressure: heavy structural borrowing

Thomason and Kaufman (1988:95) indirectly place Sprachbund phenomena in the range of categories 4 and 5 when they claim that the linguistic results of the Balkan Sprachbund “include features characteristic of moderate to heavy structural borrowing”. However, they do not place Sprachbund phenomena directly within the rubric of this scale, because the scale is meant to apply to two-language situations and, as they note, “Sprachbund situations are notoriously messy” (95), i.e., they typically involve multidirectional borrowing, and one cannot place all the features of a Sprachbund within one neat isogloss. Let us consider, then, how some of Thomason and Kaufman’s criteria might be applied to Sprachbund situations.

On their level 4, we find “major structural features that cause relatively little typological change”. On the level of phonology, this may include “new distinctive features in contrastive sets represented in native vocabulary ... new syllable structure constraints”. Grammatically, it may involve “fairly extensive word order changes” and “borrowed inflectional affixes and categories”. Level 5 involves even more intense contact, and, in the description provided by Thomason and Kaufman, it involves “major structural features that cause significant typological disruption” such as “added morphophonemic rules”, “changes in word structure rules (e.g. adding prefixes in a language that was exclusively suffixing or a change from flexional toward agglutinative morphology)”, “extensive ordering changes in morphosyntax (e.g. development of ergative morphosyntax)”, and so forth. Level 3, just a step below the Sprachbund level, may involve the phonemicization of previously allophonic variations, “easily borrowed prosodic and syllable-structure features”, and some changes in word order.

In terms of the lexicon, Thomason and Kaufman distinguish between what may be termed “cultural borrowing” or “need borrowing” of lexical items at level 1 and the borrowing of function words in categories 2 and 3. No lexical borrowings are listed in their borrowing scale (74–76) at levels 4 and 5; however the authors discuss in their examples following the scale numerous instances where lexical borrowing has acted as a

vehicle for structural change. They mention, for instance, that in Ossetic (Iranian), “lexical borrowing from Caucasian languages ... is heavy; through these loanwords Ossetic has acquired a series of glottalized stop phonemes that have also spread to native Iranian words ...” (84). This phonemic borrowing and spread to native vocabulary falls into their level 3. They evaluate contact between Chinese and the Mongolian language Baonan as lying between categories 4 and 5, with structural interference including “lexical semantics, e.g. new functions for the native Baonan verb meaning ‘hit’, to match the functions of the corresponding verb in Chinese” (90).

Despite this last example, which is, of course, a case of loanshift, there is no explicit discussion of calquing of compounds (simple calques) as opposed to calquing of more complex phrases (complex calquing) in Thomason and Kaufman’s scale, even if the existence of such a difference is implied. Birnbaum (1965) and Schaller (1975) both suggest phraseological similarities as a Balkan Sprachbund phenomenon, and calquing has been proposed as a Sprachbund feature by Campbell and colleagues (1986:553–55), who list fifty-five examples of “semantic calques or loan translations” (in their terminology) in Meso-America. Certainly calquing should be given higher priority than “pure loanwords” (again Haugen’s terminology), at least those of the cultural- or need- borrowing type, in the scale of contact-induced change, for calquing provides better evidence of bilingualism, since it involves the translation of words and phrases directly from one language into another and therefore requires rather refined L2 knowledge.

This translating of phrases sometimes even compromises “normal” patterns in the borrowing languages. Therefore, I would like to suggest that calques, especially when extensive, may begin to promote grammatical changes (this is suggested by Thomason and Kaufman, but they only mention borrowing affixes); that is, certain new patterns from the lending language may begin to “creep in” to the borrowing language at the lexical level (see discussion with exemplification below). More extensive grammatical change is possible in situations of heavy calquing, as opposed to the less extensive level three example of borrowed affixes, for example. Let me suggest then, that extensive calquing should be included as a Sprachbund phenomenon, especially where it

- a. produces observable changes in word-formation patterns
- b. affects large classes of lexical items in a manner that at least approaches structural diffusion.

We should also consider the differences among types of lexical borrowing: pure loanwords, where borrowed items appear in the borrowing language essentially in the same form in which they appeared in the lending language; simple calques, by which is understood lexical items such as compound words in which the borrowing language translates the two components of the word more or less directly and literally; and complex calques. In the last category are phrasal calques, in which we find common expressions, figures of speech, collocations, and turns of phrases translated word for word, as nearly as possible; and morphosyntactic calques, which may involve the borrowing of case government (or rules of assignment), the function of certain verbs as auxiliary verbs, or similar borrow-

ings. Using these more precisely defined criteria then, let us consider the nature of German lexical and grammatical influence on Croatian.

5.1 Morphological and phonological adaptation of loanwords

We find that German loanwords are adapted phonologically and morphologically (for the most part) to Croatian. Phonologically, for instance, front rounded vowels are rendered as front unrounded vowels (in German loanwords) as in Gm. *Bühne* > ĐCr *bine* ‘stage’. Reduced or schwa sounds, which also do not exist in Croatian, are usually rendered as /a/, which is “the only central vowel in Croatian”. Sometimes, however, schwa sounds may be rendered /e/ under the influence of German orthography, as in Gm. *Besteck* > *beštek* ‘bribe’, as noted by Piškorec (1997:72). Not only are standard German language and orthography influential; Piškorec also mentions that colloquial pronunciation of the German models often plays a role as well, as in Gm. *zurück* > ĐCr. *curuk* ‘back (return)’, with the second vowel being rendered unrounded under the apparent influence of Bavarian models. Both Magner (1966) and Piškorec (1997) find that borrowed verbs tend to be formed by adding Croatian infinitive endings, such as *–noti* or *–uvati*, as in Gm. *drücken* > *druknoti* ‘push, press’ and *danken* > *dekuvati* ‘thank’. The verbs can then be conjugated easily in accordance with ordinary rules of Croatian grammar. Nouns borrowed from German are typically left as they are, though they sometimes cause difficulty in inflection in Croatian; in some instances, German nouns that do not have an ending typical for the Croatian singular are made into pluralia tantum, as in Gm. *Spielhose* (sg.) > ĐCr *špilhoze* ‘playpants’. Some adjectives borrowed from German are uninflectable in Croatian for either case, number, or degree because their endings do not work well with either derivational patterns or agreement patterns in Croatian; three examples given by Piškorec (1997:72) are *fraj*, *fro*, and *šik* from Gm. *frei*, *froh*, and *schick* ‘free’, ‘early’, and ‘chic’, respectively. In these examples, it would seem strange to add the vowels *–y*, *–a*, and *–o* to the ending in order to form the definite nominative singular for the masculine, feminine, and neuter genders respectively. A few uninflectable items aside, then, there is no real evidence that these borrowings are leading to a “breakdown” of the Croatian inflectional system.

Lexical borrowing, however, is not limited to loanwords—there are hundreds of calques in Croatian based on German models. This, in my opinion, makes Thomason and Kaufman’s scale too limiting a gauge of bilingualism in a contact situation, since it does not take into account the **type** of lexical borrowing, and calquing typically requires a much more refined bilingualism than does borrowing.

5.2 Loanwords, calques, and “intimate” borrowings

Rammelmeyer (1975:128) insists that “The Lexicon has changed more than any other area of the Serbo-Croatian language within the last two centuries”, and concludes that German served as the most influential model for this change. He finds large numbers of German loanwords and loan translations in Serbo-Croatian from dictionaries compiled at the beginning of the nineteenth century and determines that these words must have been well-established in the lexicon in the eighteenth century (130–31). As to the means

of transmission, Rammelmeyer points to the fact that many calques were created out of the need to express new ideas in culture and science, adding that some literary loan translations may have been produced rather *ad hoc* during the process of translation from German. Such loan coinages by lexicographers count for little in the discussion of contact between the spoken languages, since they occur entirely in the quiet of scholars' studies and have nothing to do with the kind of immediate, intimate, and spontaneous contact among speakers that one seeks in a hypothetical Sprachbund setting.

However, not all or even most of the lexical influence came about by lexicographers deliberately coining new words on the basis of German models (or models from Latin and Greek). Magner (1966) finds numerous German loans that seem to be the sort that would have been transmitted through colloquial speech, such as *ganz* 'completely, entirely', a word that is used with very high frequency in German, and *ziher* (< Gm. *Sicher*) 'surely', again, a very high frequency word in German. Magner finds adjectives, nouns, adverbs, and verbs borrowed from German or based directly on German roots. Rammelmeyer notes that most German loanwords in SCr are of a material nature, i.e., they refer to concrete, everyday objects from "material culture" and are used in colloquial speech (1975:129). He adds the important comment that it is actually the loan translation, not the loanword, that is typically considered the "more educated" variant. He lists doublets such as the following, given with their German source words:

	German source word	Colloquial loanword	Calque, standard Croatian	Meaning
(1)	Weck-er wake-AGV	veker-ica waker-DIM	budi-lica wake-AGV/DIM	'alarm clock'
(2)	Kell-ner cellar-AGV	kelner	konob-ar cellar-AGV	'waiter'
(3)	Schnitz-el cut-DIM	snicla	od-rez-ak off-cut-DIM	'cutlet'

Rammelmeyer lists hundreds of loan translations in his glossary, though the vast majority of them are literary or scientific in nature and would not have been the result of spoken contact between the languages. Still, there are significant numbers of items that would likely be the result of spoken contact. Below are listed a few examples for illustration.

(4)	–maćuh-ica stepmother-DIM	Stiefmutter-chen stepmother-DIM	'pansy'
-----	------------------------------	------------------------------------	---------

As this is neither a scientific name nor a likely a need borrowing, this term was probably not calqued by botanists; therefore it was probably transmitted by ordinary contact between speakers, and it shows a subtlety of knowledge of the source language's use of the

diminutive for ‘mother-in-law’. Furthermore, the association of “stepmother” with “pansy” is so arbitrary that it must be a calque, i.e., it cannot be mere coincidence.

- | | | | |
|-----|---------------------------|-------------------------|---------------|
| (5) | -pred-soblje
fore-room | Vor-zimmer
fore-room | ‘antechamber’ |
|-----|---------------------------|-------------------------|---------------|

This would seem to be the type of word one would expect to be transmitted in the social setting described above, that is, in the communications between upper classes (the only people who would likely have such a feature in their home) and servants.

- | | | | |
|-----|--------------------------------|--------------------------------|-----------------------|
| (6) | -op-hod-ljiv
around-go-like | um-gäng-lich
around-go-like | ‘friendly, easygoing’ |
|-----|--------------------------------|--------------------------------|-----------------------|

This would seem to be the type of word that would be transmitted in colloquial speech, rather than a high style.

For a look at the more general influence of German in central Europe and its relation to the language of Croats, let us consider the following quote from Hamp (1989):

We may illustrate the differential penetration of “Habsburgisms” with an interesting series: In Serbia we hear *hvala lepo*, a visible calque on *danke schön* [literally “thanks prettily”]; the first element is simply replaced without regard for the fact that the source was a verb with its subject deleted, but the second element conserves (anomalously) the adverbial syntax. Further west, in Slavonija, we hear *hvala l(ij)epa*, with the adjectival concord required by the NP *schönen Dank*; this implies, for an earlier time, a more sensitive bilingualism. Moving yet further west, we have *najlepšia hvala = schönsten Dank* [where Croatian uses the superlative form of the adjective as in German], which required earlier a yet more refined grammatical bilingualism.

Hamp’s claim that German influence is felt more clearly the farther west one goes stands to reason: not only are the more westward areas of Serbia and Croatia closer to the German-speaking world, but they also have a longer history of German settlement and the Habsburg Dynasty ruled these areas since earlier times. Although the extension of a single group of related phrases provides only a very limited example, it is quite clear that calques from German do become more common as one moves westward, and such loan translations are, as Unbegaun (1932:28–32) observes, more common in Slovene than in Croatian, and more common in Croatian than in Serbian.

For another look at “Habsburgisms”, Magner’s 1966 study of Zagreb dialect provides a few interesting polite expressions. The greeting *zdravobóg*, a compound which adds the word ‘god’ to a salutation meaning literally ‘greetings’ (Magner 1966:80), represents a calque of the notoriously “Austrian” (but still found elsewhere in the southern German-speaking world) expression *Grüß Gott*, meaning literally ‘May God greet [you]’. *Servus*, from Lat. ‘servant/slave’, originally suggesting “I am at your service”, a meaning

likely now obscure to most speakers, is another notorious “Austrianism”, variations of which are found in Hungary and Slovakia and at least in this Zagreb dialect of Croatian. Another likely “Habsburgism” is the clearly German expression *kistijánt/kistihánt* < Gm. *Ich küsse die Hand* meaning literally “I kiss the [your] hand”, a greeting which is used still frequently in Hungary (*Kezét csokolom*), but which has apparently fallen out of use elsewhere in the former Habsburg Empire. It also has a calqued counterpart in an expression realized variously in Zagreb Croatian as *merúke/mruke/ljubimrúke*, the last of which means clearly “I kiss [your] hand”, although the first two are clearly epenthetic versions which have eliminated all but the conjugational ending of the verb.

Once more, we can find large numbers of calques from various areas of life which would not so likely have been deliberate coinages created by lexicographers, as the following examples from Rammelmeyer (1975) illustrate (hyphens here indicate morpheme boundaries and do not reflect orthographical convention):

(7)	bodljivo svinjce prickly pig	<	Gm.	Stachels-chwein prickly pig	‘porcupine’
(8)	brako-lom marriage-break	<	Gm.	Ehe-bruch marriage-break	‘adultery’
(9)	dugo-prst-ic long-finger-AGV	<	Gm.	Lang-finger long-finger	‘thief’
(10)	dzep-arac pocket-money	<	Gm.	Taschen-geld pocket-money	‘allowance’
(11a)	is-puh out-puff	<	Gm.	Aus-puff out-puff	‘exhaust’
(11b)	auspuh	—variant of (11a), a loanblend and, according to Rammelmeyer, a folk etymology.			
(12)	iz-luft-ati out-air-INF	<	Gm.	aus-luft-en out-air-INF	‘to air out’ (loanblend)
(13)	iz-nos out-carry (noun)	<	Gm.	Be-trag CAUS-carry (noun)	‘amount, sum (e.g. total on a bill)’
(14)	kameno-lom stone-break (noun)	<	Gm.	Stein-bruch stone-break (noun)	‘quarry’
(15)	kreditno sposoban credit able	<	Gm.	kredit-fähig credit-able	‘credit-worthy’

- (16) leden-jak < Gm. Eis-vogel 'kingfisher'
 ice-AGV ice-bird
 (here the association with ice seems to be rather arbitrary)
- (17a) u-pad-ljiv < Gm. auf-fäll-ig 'conspicuous'
 at-fall-ADJ at-fall-ADJ
- (17b) u oci padajuc < Gm. in die Auge fallen 'catch someone's eye'
 'fall in the eye' 'fall in the eye'

Example (17b) approaches the type of borrowings I have referred to as “complex calques”. More such complex calques are demonstrated by the following examples:

- (18) Cr. iz-davati se za koga (za sto)
 out-give Refl.Pron. for someone-ACC (for something-ACC)
- Gm. sich für jemanden (für etwas) ausgeben
 Refl.Pron. for someone-ACC (for something-ACC) out-give
 'pretend to be (i.e., imitate, impersonate) someone (something)'

- (19a) iz-loziti (dijete) < (ein Kind) aus-setzen
 out-set-INF (child-ACC) (a child-ACC) out-set
 'to expose' (as in a child to something like the weather)

There is likewise a reflexive expression:

- (19b) iz-loziti se (opasnosti) < sich (einer Gefahr) aussetzen
 out-set-INF Refl.Pron. danger-DAT Refl.Pron. danger-DAT out-set-INF
 'to expose oneself (to a danger)'
- (20) o-dugo-vlačiti < in die Länge ziehen
 PF-long-pull/draw in the length pull
 'drag out, hesitate'

Although Rammelmeyer lists over three hundred calques, I have tried to pick out a number that are not likely to be “learned calques”, i.e., the type that would be created rather deliberately and artificially by lexicographers. The words in this list come from various walks of life: finance, nature, industry, domestic life, and the streets. Of course, there are large numbers of a military, scientific, legal, academic, or technical nature, and these are much more numerous. But I think the sampling above at least begins to demonstrate that there must have been considerable spoken contact between Germans and Croats during the period in question, especially in the larger towns such as Zagreb, and that the Croats often did have a good command of German, not only in terms of their knowledge of root words, but also in their knowledge of German derivational patterns.

5.3 Syntactic borrowing through the lexicon

I begin this section by pointing out that one commonly cited Balkan areal “structural” feature might just as well be classified as a lexical feature; that is, it is somewhere between lexical and grammatical in nature. The “one-on-ten” construction for the numbers 11 to 19 is cited by Schaller (1975:150–51) as a Balkanism, with due attention to the fact that it was present in OCS (Birnbäum (1965:21) also notes its presence in all of Slavic), and to the fact that this formation pattern is also found in Hungarian. One should also note that this pattern has been extended in Hungarian to the numbers 21 to 29. In addition to its presence in Hungarian, and in Balkan and other Slavic languages, this feature is also found in Albanian and Rumanian, so it is probably correct to refer to it as a Balkan Sprachbund feature, although it is not exclusively Balkan. I am in partial agreement over its classification as a grammatical feature, however, since it is perhaps more accurately described as a word-formation pattern, and one that affects only a very limited number of lexical items. So it does not have the extensiveness in terms of the lexicon that one looks for in a Sprachbund feature.

One feature occurring in Zagreb dialects of Serbo-Croatian that points to calquing as a source of structural change is an example of syntactic calquing noticed by Thomas Magner. Magner (1966:48) points to the tendency in this dialect, following a likely model of German, to use adverbs rather than the usual Slavic means of prefixation to express direction in verbal expression. Thus, we have (standard) SC *silazim* ‘I go down/descend’ vs. Zagreb *ja idem dole* after the Gm. *ich gehe hinunter*, and SC *izlazim* ‘I go out’ vs. Zagreb *ja idem van* < Gm. *ich gehe hinaus*. Magner lists several such examples, claiming that this usage is “extensive” and “preferred” in this dialect.

I believe that the important point to be gleaned from this example is that, if certain parallel turns of phrases are calqued, and there are enough of them affecting a large enough class of lexical items, such as verbs of motion as in this example, then we have the beginnings of structural diffusion. This, essentially, is the observation that has been made by King (2000). In other words, new grammatical patterns can “creep in”, so to speak, at the grammatical level. I do not bring up this instance of calquing in one limited dialect of SC in order to suggest it as a Sprachbund feature (though it probably should be investigated further), but rather to illustrate further the notion of “structural diffusion through the lexicon” that was discussed above for numeral formations. More attention needs to be devoted to the type and extent of calquing observable in contact situations, something that Unbegaun no doubt saw the need for when writing his comparison of the Balkan and central European areas in the quote above.

6 Conclusion

While Thomason and Kaufman’s scale provides largely adequate criteria for determining the closeness of contact between languages, it seems to me that the scale could be refined somewhat in terms of its use of lexical transfer as a measure. Their consideration of lexical transfer does not distinguish between borrowing of single lexical items and calquing, even though the latter sometimes may require a more sophisticated know-

ledge of the L2 than the former does. In other words, as we have seen from many of the foregoing examples, rather precise word-formation patterns may often be invoked in translating not only root words, but also less obvious linguistic structures such as diminutive suffixes. To be certain, the transfer of morphological and syntactic features from L2 to L1 is an indicator of a closer contact situation than mere lexical transfer generally is. But perhaps in some situations, a more refined knowledge of the L2 may result in a greater ease in switching between L1 and L2, in switching between disparate structures or linguistic systems without interference occurring. In such a situation of more highly developed bilingualism, might not the greatest sign of close contact actually be lexical transfer in the form of calquing, rather than grammatical transfer?

Unbegaun (1932:47) writes of several German-influenced Slavic languages under consideration that “the four languages in question, Croatian, Slovene, Czech and Sorbian, all offer in effect the common trait of being spoken by populations more or less bilingual and accustomed to making use of calques to some extent”. He concludes his article with the following words:

We have spoken for a long time of the linguistic community of the Balkan world, a community which affirms itself by general traits of vocabulary and syntax, and indeed of morphology. But, if we someday take to determining a similar community of central Europe, it is the calque which will be its most characteristic indicator. (1932:48)

Unbegaun is probably correct to observe that Croats’ bilingualism is “less developed” (28) than that of the Czechs (where something approaching structural influence can be observed in the development of a quasi-definite article and a quasi-compound perfect tense with ‘have’ + past participle). He is also right, in my opinion, to include the Croats in a central European “linguistic community” or Sprachbund, even if their participation in it tends to be somewhat less pronounced, somewhat more at the peripheries of the linguistic area.

References:

- BANAC, IVO. 1984. Main trends in the Croatian language question. *Aspects of the Slavic language question. Volume I: Church Slavonic—South Slavic—West Slavic*, ed. by Riccardo Picchio and Harvey Goldblatt, 189–259. New Haven: Yale Concilium on International and Area Studies.
- BIRNBAUM, HENRIK. 1965. Balkanslavisch und Südslavisch: Zur Rechtweite der Balkanismen im südslavischen Sprachraum. *Zeitschrift für Balkanologie* 3.12–63.
- CAMPBELL, LYLE; TERRENCE KAUFMAN; and THOMAS C. SMITH-STARK. 1986. Meso-America as a linguistic area. *Language* 26.34.530–70.

- DÉCSY, GYULA. 1973. *Die linguistische Struktur Europas*. Wiesbaden: Otto Harrassowitz.
- GULDESCU, STANKO. 1970. *The Croatian-Slavonian Kingdom: 1526–1792*. The Hague: Mouton.
- HADROVICS, L. 1985. The status of the Croatian regional languages immediately before Gaj's Reforms. *The formation of the Slavonic literary languages*, ed. by Gerald Stone and Dean Worth, 133–46. Columbus, OH: Slavica.
- HAMP, ERIC. 1979. Linguistic areas or clusters? *Quatrième Congrès International des Études du Sud-est Européen: Abrégés des communications*, 282–83. Ankara: Türk Tarih Kurumu.
- HAMP, ERIC. 1989. Yugoslavia – a crossroads of Sprachbünde. *Zeitschrift für Balkanologie* 25. Wiesbaden: Otto Harrassowitz.
- HAUGEN, EINAR. 1953. *The Norwegian language in America: A study in bilingual behavior. Vol. I: The bilingual community. Vol. II: The American dialects of Norwegian*. Bloomington: Indiana University Press. (Reprinted 1969.)
- JOSEPH, BRIAN D. 1983. *The synchrony and diachrony of the Balkan infinitive: A study in areal, general, and historical linguistics*. (Cambridge studies in linguistics, Supplementary volume 1.) Cambridge & New York: Cambridge University Press.
- KATICIĆ, RADOSLAV. 1984. The making of standard Serbo-Croatian. *Aspects of the Slavic language question. Volume I: Church Slavonic—South Slavic—West Slavic*, ed. by Riccardo Picchio and Harvey Goldblatt. Columbus, OH: Slavica.
- KESSLER, WOLFGANG. 1981. *Politik, Kultur und Gesellschaft in Kroatien und Slawonien in der ersten Hälfte des 19. Jahrhunderts: Historiographie und Grundlagen*. München: R. Oldenbourg Verlag.
- KING, RUTH. 2000. *The lexical basis of grammatical borrowing: A Prince Edward Island French case study*. Amsterdam: John Benjamins.
- MAGNER, THOMAS F. 1966. *A Zagreb Kajkavian dialect*. (The Pennsylvania State University studies 18.) University Park: The Pennsylvania State University.
- PIŠKOREC, VELIMER. 1997. *Deutsches Lehngut in der kajkavischen Mundart von Durđevac in Kroatien. Schriften zur deutschen Sprache in Österreich, Band 22*. Frankfurt: Peter Lang.
- RAMMELMEYER, MATTHIAS. 1975. *Die deutschen Lehnübersetzungen im Serbokroatischen: Beiträge zur Lexikologie und Wortbildung*. Franz Wiesbaden: Steiner Verlag.

SCHALLER, HELMUT W. 1975. *Die Balkansprachen: I Eine Einführung in die Balkanphilologie*. Heidelberg: Carl Winter.

THOMASON, SARAH. 2001. *Language contact*. Washington: Georgetown University Press.

THOMASON, SARAH GREY, and TERRENCE KAUFMAN. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

UNBEGAUN, BORIS. 1932. Le calque dans les langues slaves littéraires. *Revue des études slaves* 12.19–48.

WINFORD, DONALD. 2003. *An introduction to contact linguistics*. (Language in society 33.) Oxford: Blackwell.

LOOKING FOR ROOTS IN THE SUBSTRATE: THE CASES OF EBONICS AND ANGLO-IRISH

Terence Odlin

Abstract

Despite many differences in the sociolinguistic setting of Hiberno-English in Ireland and African-American Vernacular English in the USA, arguments about substrate influence have been invoked in both cases to promote the notion of separate linguistic identities. In the case of Ireland, Henry (1958, 1977) has insisted that the proper term to describe the vernacular now used by many in rural Ireland is “Anglo-Irish”, as opposed to “Hiberno-English” or “Irish English”, and he argues that “a new language” was created as a result of the substrate influence that became especially prominent in the nineteenth century. There have likewise been strong claims about the significance of substrate influence in African American Vernacular English, or to use the term advocated by the Oakland School Board, “Ebonics”. In 1996 the Board declared this variety to be “not a dialect of English” but instead an instance of “African Language Systems”. The arguments of Henry and of the Oakland School Board may not convince linguists that Anglo-Irish and Ebonics are indeed distinct languages, but these claims do warrant reconsidering the question of where English begins and ends.

1 Introduction

Substrate influence can be defined as the role that one's native language (most typically) can play in the acquisition of (most typically) a second language. The hedges just used, "most typically", acknowledge the fact that substrates are sometimes involved in the acquisition of a third language, but the second language situation is likely the most common and is the focus in this paper (cf. Odlin 1989, Ceñoz & Jessner 2000). Ever since the nineteenth century the study of substrate influence has received increasing attention. Much of the early interest focused on historical questions, and the burgeoning research literature on language contact shows the ongoing scrutiny of the diachronic dimensions of substrate influence (e.g., Thomason & Kaufman 1988). Yet another research community has also focused on such influence—namely, those linguists who investigate second language acquisition—and that field rapidly grew in the second half of the twentieth century (e.g., Ringbom 1987). Although acquisition studies and historical research do not overlap a great deal, work on cross-linguistic influence in one can benefit work in the other, as Salikoko Mufwene (1990) and others have pointed out.

In both historical linguistics and second language acquisition, there have been skeptics questioning the empirical reality of cross-linguistic influence, as in the arguments of Roger Lass (1990) questioning the influence of Irish on Hiberno-English. However, the increasing detail in the evidence for such influence not only in Ireland but also in many other language contact situations has made the skeptics a small band—and not an especially well-informed one. In contrast to the skeptics, there are others who accept as a matter of course the reality of cross-linguistic influence, and they invoke it to affirm the value of both the speakers of the substrate languages and of the speakers of the varieties showing cross-linguistic influence. In this paper I compare the stances toward substrate influence in two cases, the first in Ireland, as seen in some of the writing of P. L. Henry, and the second in the United States, as seen in a 1996 resolution of the school board in Oakland, California regarding what the board termed Ebonics. My comparison will, I hope, cast light on the question of where dialects end and languages begin.

2 Anglo-Irish

P. L. Henry is probably best known for his detailed description of the vernacular titled *An Anglo-Irish Dialect of North Roscommon*, which was published in 1957. Most of the monograph addresses specific points where the vernacular diverges from the dialects of Britain and America and especially where there are parallels with the indigenous Celtic language of Ireland known both as Gaelic and Irish. In the introduction Henry distinguishes "common AI" and "rural AI", with the former capable of serving as "a link between the rural dialect and StE [Standard English]". In a lexical survey published in 1958, he discusses the ambiguities of the term *Anglo-Irish* (AI), which can have three senses: emphasis on the original settlers, emphasis on the native Irish, and also a neutral sense. The first sense is probably that which is used most by scholars and includes some treatments of language as found in a book by Loreto Todd (1989). With regard to the second sense, Henry notes that the stress falls on the second element of the compound. Commenting on the rich Anglo-Irish literary tradition that developed in the nineteenth

century, Henry implies the importance of Irish substrate influence, as seen in the following passage:

The reality underlying later literary developments was that Ireland was forging a new language on the pattern of the old. English, transformed in the mouths of an Irish-speaking people, was fraught with potentiality. (1958:56)

As seen in the quotation, Henry goes so far as to term Anglo-Irish a new language. Although these observations from the 1950s give a sense of Henry's thinking, it is in an address published in 1977 that we get an elaborated discussion of the overall significance of substrate influence. As in his earlier observations, he views Anglo-Irish as a "new language" (1977:24), but here he amplifies on the reasons. For Henry, substrate influence endowed Anglo-Irish with a wellspring of creativity. The speakers of earlier times were not engaged in mere "learning", which for Henry is no more than imitation:

Whereas in nineteenth century Ireland the colonist stock held the reins, the tie with England was close and there was no scope for the rise of Anglo-Irish as a national speech norm fashioned by the people and therefore adapted for their own needs, educational, social and political. The situation is symbolised by an Education Machine which could not understand the *creation* but only the *imitation* and learning of language. (1977:25, emphases in the original)

Also evident in the quote is Henry's conviction that Irish schools in the days before independence had little understanding or sympathy for what made the speech of bilingual children distinctive.

Although Henry stresses the uniqueness of Anglo-Irish and its creative potential, he does concede that in structural terms it owes much to English, as seen in the following: "Generally speaking, the material basis of the language, that is, words and grammatical forms, were very largely from English" (1977:34). On the other hand, he adopts a quasi-generative stance, as seen in the next quotation, to argue for a distinctive linguistic identity: the English-derived material parts are "set in motion by a deeper structure, namely, that of meaning" (1977:35). (On the final page of the article, moreover, he actually uses the term *generative*.) In the same paper, he gives numerous examples of idioms that have an Irish parallel; it appears that his examples come from a manuscript in the archives of the Department of Irish Folklore at University College Dublin. Although some of the examples he gives may not necessarily reflect Irish substrate influence, I have checked several of the UCD idioms to see if they also appear in British English, and they do not (Odlin 1991).

Meaning, then, plays a key role in Henry's arguments for the distinctiveness of Anglo-Irish. Not surprisingly, he invokes assumptions common in linguistic relativism:

You will be obliged to concede on reflection that languages are distinctive in the first instance because they embody contrasting views of life, and secondly because they employ their own favoured linguistic modes to mediate their particular world view. (1977:27)

Such thinking brings to mind not only modern work on relativism but also the Romanticism of thinkers such as Wilhelm von Humboldt (1836/1988). For Henry, the vernacular of modern Ireland embodies not only a distinctive history but also a distinctive world view.

3 Ebonics

Turning now to Ebonics, we see similarities and differences regarding the significance of substrate influence. Some background on the Oakland resolution may be useful. In December of 1996, the school board issued a document that the media considered highly newsworthy, and the controversy generated by the original resolution lasted well into the following year, even after the board revised the resolution in January of 1997. I have written a detailed analysis of the two resolutions and will not attempt to summarize all the points made in that article (cf. Odlin 1999). However, the most striking feature of the December resolution was the board's declaration that Ebonics is a separate language, not a variety of English:

...numerous validated scholarly studies demonstrate that African American students as part of their culture and history as African people possess and utilize a language described in various scholarly approaches as "Ebonics" (literally Black sounds) or Pan African Communication or African Language Systems; ... these studies have also demonstrated that African Language Systems are genetically-based and not a dialect of English.

While terms such as "African Language Systems" are not widely used by most linguists to describe either the languages of Africa or anything in the Americas, the wording of the Oakland resolution indicates the school board's interest in substrate influence. In fact the board invokes an African language family in relation to Ebonics: "these studies demonstrate that such West and Niger-Congo African languages have been officially recognized and addressed in the mainstream public educational community as worthy of study". In the revised resolution, moreover, the board makes its historical argument somewhat clearer: "these studies have also demonstrated that African Language Systems have origins in West and Niger-Congo languages". As with Henry's analysis of Anglo-Irish, then, the Ebonics resolution invokes substrate influence to posit the existence of a separate language—although the board retreated from this position somewhat in the revised resolution (Odlin 1999).

Also similar to Henry's position is the desire of the Oakland board for the local schools to affirm the unique identity of pupils who speak a distinct linguistic variety. The board states that the aims of "the Oakland Unified School District in providing equal

opportunities for all of its students dictate limited English proficient educational programs [*sic*]. Although the sentiments expressed are similar to Henry's, the wording of the resolution indicates a special agenda. The phrase "limited English proficiency" is customary in American bilingual education programs, and other wording in the Oakland resolution makes it clear that the school board was pursuing a novel strategy to obtain Federal funds available through bilingual education. The board could not expect to obtain such funding unless Ebonics was deemed to be a distinct language, and the first part of the December resolution lays the ground for obtaining such support. The furor in the national media led the board to claim that it never had any interest in obtaining bilingual education funds, but the language of the December resolution makes their denials hard to believe (Odlin 1999).

Although the interest in obtaining Federal funding was probably the main reason for declaring Ebonics to be a separate language, the school board continued to emphasize the role of substrate influence even after the bilingual education argument was dropped. In the revised resolution, the board expresses its intent to

implement the best possible academic program for the combined purposes of facilitating the acquisition and mastery of English language skills, while respecting and embracing the legitimacy and richness of the language patterns whether they are known as "Ebonics," "African Language Systems," or "Pan African Communication Behaviors" ...

Although the wording is not quite as explicit as Henry's, the resolution suggests that the "legitimacy and richness" of Ebonics can be traced at least partly to the African substrate. Whether or not funding opportunities could improve from a new name, the nomenclature might persuade some members of the public that the speech ways of African American children have as much history as any other and deserve as much respect.

4 Substrate influence

What the arguments of P. L. Henry and the Oakland school board have in common, then, is primarily the foregrounding of substrate influence. In both cases, such influence is seen not only as a rich source for innovation but also as the basis of a new language. Yet in both cases the substratist positions leave important questions unanswered. Henry seems to believe that Anglo-Irish is "a kind of Irish" (1977:36) but it remains unclear whether or not he would consider Anglo-Irish to be a Celtic instead of a Germanic language. Likewise the Ebonics resolution invokes the taxonomic term *Niger-Congo*, but the authors do not make clear whether or not they would maintain that Ebonics belongs to the Niger-Congo family. Classifying the one as Celtic and the other as Niger-Congo would at least provide some clarity as to just how strongly the authors believe Anglo-Irish and Ebonics to be separate languages. On the other hand, those classifications would be extremely dubious in view of the structural evidence. Anglo-Irish resembles dialects of English far more than it does any Celtic language, and the same goes for Ebonics as compared with any Niger-Congo language. The Appendix includes versions of the Lord's Prayer in Irish and in Yoruba so that readers can get a sense of

how different Celtic and Niger-Congo languages are from an English-related creole, Gullah, which is spoken in South Carolina and Georgia.

The Gullah text clearly resembles English and what the school board terms Ebonics, and the resemblance in both cases is obviously far greater than any resemblance to Yoruba. Even so, the fact that most linguists do consider Gullah to be a separate language might seem to provide a promising argument in the case of Anglo-Irish and Ebonics: one might claim that both are also creoles. However, there are compelling counter-arguments in both cases (Odlin 1997, Winford 1997, 1998).

5 Languages vs. dialects

The cases of Anglo-Irish and Ebonics raise the old question of where a dialect ends and a language begins. Cross-linguistic distances are obviously relevant to that question, and if language distance were the only factor, neither Henry's argument nor that of the Oakland board would deserve much attention at all. On the other hand, both cases involve the use of historical arguments to pursue another aim: namely, that of raising the estimation of Anglo-Irish and Ebonics. A classic article by William Stewart (1968) discusses four factors relevant to the language/dialect distinction: standardization, autonomy, historicity, and vitality. The stance taken by Henry and also by the Oakland board addresses two of those factors: historicity and vitality. That is, both Anglo-Irish and Ebonics have distinctive histories (even if substrate influence could not be invoked), and both varieties have no shortage of speakers who confer vitality on them. Even so, neither variety has really been standardized. Moreover, speakers and writers of these varieties have not established a high degree of autonomy. The Oakland board declared Ebonics to be a separate language in December of 1996, but a month later it no longer saw fit to follow through on the logic of such claim, that is, to seek funding for a bilingual education program. Furthermore, neither Henry nor the Oakland board chose to write their polemics in actual Anglo-Irish or Ebonics.

Elsewhere I have argued that the process of a dialect becoming a language is similar to the way a new country may obtain diplomatic recognition (Odlin 1999). P. L. Henry and the Oakland board have consciously adopted terminology to affirm the linguistic independence of Anglo-Irish and Ebonics, largely from a concern that both the speech ways and the speakers have not gotten the respect they deserve. All the same, it is possible to agree with the sentiments underlying such declarations of independence but at the same time to predict that neither variety will ever be considered anything other than English by most people. Here it will help to think of some of the competing terms: *Hiberno-English* and *Irish English*, on the one hand, and *Black English* and *African American Vernacular English* on the other. The fact that no term is universally used in either case suggests that the status of these varieties remains contested, and this terminological limbo seems likely to continue for many years to come.

6 Conclusion

The cases of Anglo-Irish and Ebonics are, I think, interesting in their own right, but they also point to a larger concern: namely, the notion of the unity of the English-speaking world. As noted above, the declaration of the Oakland school board in 1996 got considerable attention in the media, yet there has been much less publicity of a recent decision of the British government that may have very serious ramifications for language policy. In 1998 negotiators concluded the terms of the Good Friday Agreement, the official framework for a peace settlement in Northern Ireland, and some of the terms in the agreement involved language. It is beyond the scope of this paper to provide a detailed discussion of all the linguistic issues that complicated the peace negotiations, but one outcome should be noted: the various negotiating parties, including the British government, agreed to recognize Ulster Scots as a language of Northern Ireland and, by implication, as an official language of part of the United Kingdom. Ulster Scots resembles the vernacular English of southern Scotland (which is often called Scots). The Ulster variety certainly has distinctive characteristics, and now language activists in Belfast and elsewhere are trying to put the variety to novel uses including, for example, official job notices written in the newly declared language (cf. Görlach 2000). The distinctive identity of Ulster Scots does not depend mainly on substrate influence, but it does show one more example of how declarations such as those of P. L. Henry and the Oakland school board could, for better or worse, call into question the unity of the English language.

Appendix. The Lord's Prayer in Gullah, Irish, and Yoruba.

Gullah, a creole considered by some linguists to be the most archaic form of African American Vernacular English, has a great deal of vocabulary identical with English, even though it diverges in some areas of grammar (especially verb phrase construction). Irish, a Celtic language, shares several Indo-European affinities with English, but most are quite opaque without some background in Indo-European linguistics (e.g., *ríocht* 'kingdom'). As a Niger-Congo language Yoruba is, not surprisingly, a language quite different from the other two in the appendix.

The Lord's Prayer in Gullah

(Source: <http://www.gullahtours.com/prayers.html>)

Our Fadduh awt'n Hebb'n, all-duh-weh be dy holy 'n uh rightschus name. Dy kingdom com.' Oh lawd leh yo' holy 'n rightschus woud be done, on dis ert' as-'e tis dun een yo' grayt Hebb'n. 'N ghee we oh Lawd dis day our day-ly bread. 'N f'gib we oh Lawd our trus-passes, as we also f'gib doohs who com' sin 'n trusspass uhghens us. 'N need-us-snot oh konkuhrin' King een tuh no moh ting like uh sin 'n eeb'l. Fuh dyne oh dyne is duh kingdom, 'n duh kingdom prommus fuh be we ebbuh las'n glory. Amen.

The Lord's Prayer in Irish

(Source: <http://www.stmarysdrogheda.ie/Patrick%27s%20day/MassPrint.htm>)

Ár n-athair, atá ar neamh: go naofar d'ainm. Go dtaga do riocht. Go ndéantar do thoil ar an talamh, mar dhéantar ar neamh. Ár n-arán laethiúl tabhair dúinn inniu, agus maith dúinn ár bhfiacha, mar mhaithimid dár bhféichiúnaithe féin. Agus ná lig sinn i gcathú, ach saor sinn ó olc. Óir is leatsa an Ríocht agus an Chumhacht agus an Ghlóir, tré shaol na saol.

The Lord's Prayer in Yoruba

(Source: <http://www.christusrex.org/www1/pater/JPN-yoruba.html>)

Baba wa ti mbe li orun.Ki a bowo fun oruko re ki ijoba re de.Ife tire ni ki ase li aiye
Bi won ti nse li orun Fun wa ni onje ojo wa loni Dari ese wa ji wa Bi ati ndari ese ji awon
ti ose wa.Mafa wa sinu idanwo.Sugbon gba wa lowo bilísi. Amin.

References

- CEÑOZ, JASONE, and ULRIKE JESSNER (eds.) 2000. *English in Europe: Acquisition of a third language*. Clevedon, UK: Multilingual Matters.
- FILPPULA, MARKKU. 1999. *The grammar of Irish English*. London: Routledge.
- GÖRLACH, MANFRED. 2000. Ulster Scots: A language? *Language and politics*, ed. by John M. Kirk and Dónall P. Ó Baoill, 13–31. (Belfast studies in language, culture, and politics 1.) Belfast: Queen's University Belfast.
- HENRY, P. L. 1957. An Anglo-Irish dialect of North Roscommon. Dublin: University College, Department of English.
- HENRY, P. L. 1958. A linguistic survey of Ireland: Preliminary report. *Lochlann: A Review of Celtic Studies* 1.49–208.
- HENRY, P. L. 1977. Anglo-Irish and its Irish background. *The English language in Ireland*, ed. by Diarmaid Ó Muirthe, 20–36. Cork: Mercier Press.
- LASS, ROGER. 1990. Early mainland residues in southern Hiberno-English. *Irish University Review* 20.137–48.
- MUFWENE, SALIKOKO. 1990. Transfer and the substrate hypothesis in creolistics. *Studies in Second Language Acquisition* 12.1–23.
- ODLIN, TERENCE. 1989 *Language transfer*. Cambridge: Cambridge University Press.

- ODLIN, TERENCE. 1991. Irish English idioms and language transfer. *English World-Wide* 12.175–93.
- ODLIN, TERENCE. 1997. *Hiberno-English: Pidgin, creole, or neither?* (CLCS occasional paper 49.) Dublin: Centre for Language and Communication Studies, Trinity College Dublin.
- ODLIN, TERENCE. 1999. An army, a navy, and Ebonics. *Language contact, variation, and change* (Studies in languages 32), ed. by Jussi Niemi, Terence Odlin, and Janne Heikkinen, 195–233. Joensuu, Finland: University of Joensuu. Online: www.english.ohio-state.edu/people/odlin.1/papers/.
- RINGBOM, HAKAN. 1987. *The role of the first language in foreign language learning*. Clevedon, U.K.: Multilingual Matters.
- STEWART, WILLIAM. 1968. A sociolinguistic typology for describing national multilingualism. *Readings in the sociology of language*, ed. by Joshua Fishman, 531–45. The Hague: Mouton.
- THOMASON, SARAH, and TERENCE KAUFMAN. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- TODD, LORETO. 1989. *The language of Irish literature*. New York : St. Martin's Press.
- VON HUMBOLDT, WILHELM. 1836/1988. *On language*. Cambridge: Cambridge University Press.
- WINFORD, DONALD. 1997. On the origins of African American Vernacular English: A creolist perspective. I: The sociohistorical background. *Diachronica* 14.305–44.
- WINFORD, DONALD. 1998. On the origins of African American Vernacular English: A creolist perspective. II: Linguistic features. *Diachronica* 15.99–154.

CONTACT-INDUCED CHANGES – CLASSIFICATION AND PROCESSES

Donald Winford

Abstract

Traditionally, contact-induced changes in languages have been classified into two broad categories: those due to “borrowing” and those due to “interference” by an L1 or other primary language on an L2 in the course of second language acquisition (SLA). Other terms used for “interference” include “substratum influence” and “transfer”. Labels like these, unfortunately, have been used to refer both to the outcomes of language contact and to the “mechanisms” or processes that lead to such results. This imprecision in the use of key terms poses serious problems for our understanding of what is actually involved in the two types of cross-linguistic influence. Moreover, it has led to pervasive inaccuracy in our assignment of changes to one or the other category. The aim of this paper is to re-assess the conventional wisdom on the distinction between borrowing and “interference” and to clarify the processes as well as the outcomes characteristic of each. My approach is based on van Coetsem’s (1988) distinction between the mechanisms of borrowing under RL agentivity and imposition under SL agentivity, with their shared but differently implemented processes of imitation and adaptation. Crucially, this approach recognizes that the same agents may employ either kind of agentivity, and hence different psycholinguistic processes, in the same contact situation. It is the failure to recognize this that has sometimes led to inaccuracy in accounts of the nature and origins of contact-induced changes, as

well as to conflicting classifications of the outcomes of contact. The present paper proposes a more rigorous and consistent classification, based on the kinds of agentivity involved.

1 Introduction

Traditionally, contact-induced changes in languages have been classified into two broad categories: those due to “borrowing” and those due to “interference” by an L1 or other primary language on an L2 in the course of second language acquisition (SLA), particularly language shift. The second type of change, interference via shift, has also been referred to as substratum influence, especially in the context of creole formation, and as transfer, in the context of SLA. Labels like these, unfortunately, have been used to refer both to the outcomes of language contact and to the “mechanisms” or processes that lead to such results. Statements like the following, from Thomason and Kaufman (1988: 69), are typical of what we find in the literature:

If we know that contact was intimate enough to make shift as well as borrowing possible, then there is no reason to suppose that one process operated to the exclusion of the other, barring established social or numerical asymmetry that would enable us to rule out one of the mechanisms.

Here, “borrowing” and “shift” are treated as “mechanisms” or “processes” without any clear explanation of what these terms mean. This imprecision in the use of key terms poses serious problems for our understanding of what is actually involved in the two types of cross-linguistic influence. Moreover, it has led to pervasive inaccuracy in our assignment of changes to one or the other category.

The aim of this paper is to re-assess the conventional wisdom on the distinction between borrowing and interference and to clarify the processes as well as the outcomes characteristic of each. Students of language contact have sometimes pointed to the indeterminacy of these terms. For example, Haugen (1950:213) points out that “borrowing as here defined is strictly a process and not a state, yet most of the terms used in discussing it are ordinarily descriptive of its results rather than of the process itself”. He further notes that the classifications of borrowings into loanwords, loan translations, and the like “are merely tags that various writers have applied to the observed results of borrowing” (ibid.). Hammarberg (1997:162) makes a similar point about the different ways in which the term “transfer” has been used and interpreted, namely

(a) at the level of *strategy*, with regard to the learner’s plan of action to solve a particular problem; (b) at the level of *execution*, with regard to the event or process of carrying out the strategy; and (c) at the level of *solution*, with regard to the product (as manifested in the learner’s L2 performance) of the applied strategy.

Classifications of the outcomes of language contact are of course useful and necessary. But their focus on results often obscures the nature of the mechanisms and

psycholinguistic processes that lie behind them. By reifying terms like “borrowing” and “transfer” we have tended to commit ourselves to pre-determined classifications of contact phenomena, and even to misapply the labels in some cases. Moreover, in doing so, we have tended to overlook some of the similarities in process between the two types of cross-linguistic influence—similarities that sometimes make the boundary between the two fuzzier than might first appear.

Perhaps the most comprehensive (and least appreciated) attempt to sort out the terminological mess in discussions of contact phenomena was made by van Coetsem (1988). He makes a broad distinction between borrowing and what he calls imposition, and defines them in terms of two transfer types, which he labels recipient language (RL) agentivity and source language (SL) agentivity. It will become clear below how these two transfer types are related to the actual mechanisms or processes involved in contact-induced change. Note that transfer in this context is used in a neutral sense, to refer to any kind of cross-linguistic influence, not just L1 influence in SLA.

In this approach, borrowing and imposition are epiphenomena or cover terms for the actual mechanisms involved in the two types of cross-linguistic influence. Each involves a particular kind of agentivity on the part of speakers, as well as a particular direction of change. In borrowing, materials from an external source language are imported into an RL via the agency of speakers for whom the latter is the dominant or primary language, i.e., RL agentivity. In imposition (which corresponds to what SLA researchers call transfer) the source language is the dominant (usually the first) language of the speaker, from which materials are transferred into an RL in which the speaker is less proficient, i.e., SL agentivity. Each type of cross-linguistic influence is associated with particular psycholinguistic processes via which materials are transferred from one language to another. As will become clearer, we need to distinguish the agents of change from the kinds of agentivity they employ in introducing changes. This is so because the same agents may employ either kind of agentivity, and hence different psycholinguistic processes, in the same contact situation. As our discussion will make clear, failure to recognize this has sometimes led to inaccuracy in accounts of the nature and origins of contact-induced changes. Let us now turn our attention to these.

2 Agentivity in borrowing

Thomason and Kaufman (1988:37) define borrowing as “the incorporation of foreign features into a group’s native language by speakers of that language: the native language is maintained but is changed by the addition of the incorporated features”. This appears to coincide broadly with van Coetsem’s (1988:3) definition in terms of RL agentivity:

If the recipient language speaker is the agent, as in the case of an English speaker using French words while speaking English, the transfer of material (and this naturally includes structure) from the source language to the recipient language is *borrowing (recipient language agentivity)*. (italics in original)

In the light of this, several aspects of Thomason and Kaufman's definition appear somewhat vague. In the first place, the term "dominant" or "primary" language seems more suitable than "native" language, since the latter is often in doubt (for example, in some cases of bilingualism among children) or often yields to another primary language in the course of socialization (Weinreich 1953:14). In addition, Thomason and Kaufman's definition does not make it clear whether the agents of borrowing are monolinguals or bilinguals, though elsewhere they mention the latter as possible agents. In fact, as van Coetsem (1988:10) points out, both RL monolinguals and RL-dominant bilinguals can be agents of borrowing. More seriously, it is insufficient to define borrowing only in terms of the agents and direction of change, important though these are. What matters, crucially, is the type of agentivity or transfer mode that is involved. Henceforth I will use the term "RL-dominant" to refer to both RL-monolinguals and RL-dominant bilinguals. Similarly, the term "SL-dominant" will refer to both monolingual and bilingual speakers for whom the source language is the primary language. There are, of course, different degrees of dominance and bilingualism, which may have consequences for the kind of contact-induced change that occurs (see below).

Finally, we must not confuse language dominance with language maintenance. Many languages are maintained over long periods of time, even when large numbers of their speakers have adopted another language as their primary language. Such speakers may be agents of significant structural changes in the maintained language. I argue that such cases generally involve SL agentivity, by which speakers of the dominant language impose its features on their version of the maintained ancestral language. The resulting changes may eventually be adopted by other speakers for whom the maintained language is still dominant (as Thomason and Kaufman point out). Hence we find a combination of the two transfer types in such situations. Crucially, though, the original means by which the changes are first introduced is SL agentivity. Thus it is dubious at best to ascribe such changes to (a "process" of) borrowing.

Distinguishing borrowing from imposition in this way allows us to identify and compare more precisely the mechanisms or processes that lead to each outcome. Haugen (1953:383) points out the difficulty associated with the latter:

Unfortunately, we are unable to watch the mental processes directly, and can only guess at them by observing their results and comparing those results with what the speakers themselves report about their own mental experiences.

Haugen suggests that every lexical borrowing involves two such processes: importation and substitution. The former is typically partial, since it isn't necessary "to take over a word with all its sounds, forms and meanings intact" (*ibid.*). Instead, borrowing language speakers tend to "substitute some of the habits of their own language for those in the source language" (*ibid.*). Van Coetsem suggests instead a distinction between "imitation" (roughly corresponding to Haugen's "importation") and "adaptation" (corresponding to "substitution"). The latter involves the use of L1 habits in modifying features imported from an sl. Henceforth, I follow van Coetsem's terminology, which appears more trans-

parent and applicable. This is not to claim, of course, that these terms represent the actual mental processes that speakers employ.

The twin mechanisms explain much about the types of lexical contact phenomena that have been classified as borrowings. A simple classification is shown in Table 1, adapted from Haugen (1950, 1953).

LEXICAL CONTACT PHENOMENA	EXAMPLES
A. Lexical borrowings	
1. Loan words	
a. "Pure" loanwords	French <i>rendezvous</i> in English.
b. Loan blends	Pennsylvania German. <i>bassig</i> (E. <i>boss</i> + G. <i>-ig</i>)
2. Loan shifts (loan meanings).	
a. Semantic extensions	American Portuguese <i>frio</i> 'cold infection' (on model of Eng. <i>cold</i>).
b. Loan translations	Germ. <i>Wolkenkratzer</i> (cf. Eng. <i>skyscraper</i>)
B. Creations	
1. Purely native creations	Pima 'wrinkled buttocks' for 'elephant'
2. Hybrid creations	Yaqui <i>lios-nóoka</i> (Lit. 'god-speak') 'pray'
3. Creations using only foreign morphemes.	Japanese <i>wan-man-ka</i> 'bus with no conductor' English <i>one + man + car</i> .

Table 1. A simplified classification of lexical borrowings.

The lexical phenomena shown in Table 1 are not exact imitations, but rather the products of various creative processes applied to SL forms or patterns. Some of them, for instance, loanwords and loan blends, illustrate the processes of importation and adaptation that are associated with prototypical lexical borrowing under RL agentivity. In this transfer type, as van Coetsem shows, imitation comes first, and then adaptation alters the imported item so that it conforms fully to RL phonology, morphology, and syntax. In other words, lexical borrowing typically adds new lexical items to the RL without affecting its structure. Most of the categories of lexical borrowing shown in Table 1 conform to this pattern.

However, other lexical contact phenomena such as loan translations appear to involve the transfer of structural patterns from the SL to the RL. Heath (1984:367) refers to this as "pattern transfer" and distinguishes it from borrowing. The question then is whether phenomena like calquing are true borrowings, in the sense in which van Coetsem uses the term. In other words, is imitation of a foreign structural pattern similar in kind to imitation of a foreign lexical item? What kinds of structure can be imitated (or borrowed) under RL agentivity? There seems to be consensus that patterns of the type involved in calquing, as well as derivational morphology, can be imitated in this way. This kind of borrowing, though, is primarily lexical in nature, though it involves the transfer of structural patterns (see discussion below). But what limits are there on borrowing of this type?

The larger issue here is whether, and by what criteria, the transfer of structural patterns from an SL to an RL can always legitimately be viewed as borrowing.

3 The issue of structural borrowing

It has long been a matter of debate whether, and under what conditions, languages can borrow structural features. The answers to these questions are vital to our understanding of contact-induced structural change, as well as to our classification of its products. Thomason and Kaufman argue that there is a scale of borrowing, with slight lexical borrowing at one extreme and extensive grammatical replacement at the other, with varying degrees of structural borrowing in between. This clearly implies that structure can be borrowed in its own right, and in significant degrees.

In fact, it is arguable that many instances of so-called structural borrowing are not the result of direct importation or imitation of the kind associated here with lexical borrowing. As we will see, certain structural innovations in an RL appear to be mediated by lexical borrowing, and are therefore not clear cases of “pure” structural borrowing. In other cases where direct borrowing of structural elements occurs, as it seems to in some situations, it typically involves free morphemes such as prepositions and conjunctions. Bound morphemes appear to be borrowed only in cases where they substitute for RL morphemes that are semantically and structurally congruent with them. Moreover, such borrowing requires a high degree of bilingualism among individual speakers.

The question then is whether other structural features, for example, word order, morphosyntactic categories, argument structures, and the like, can be transferred through the mechanism of borrowing. Before we consider this, let us examine structural innovations that do appear to involve true borrowing.

3.1 Cases of structural borrowing

There is ample evidence that heavy lexical borrowing can introduce new structural features into a language. A well-known example is the extensive borrowing of French lexicon into Middle English in the fourteenth to fifteenth centuries. The introduction of French loans with initial [v ð z] led to the phonemicizing of OE allophonic variants such as [f] and [v], [θ] and [ð], and [s] and [z]. The respective pairs of fricatives were originally allophones, voiced in intervocalic position, but voiceless elsewhere—e.g., [wi:f] ‘woman’ vs. [wi:vas] ‘women’. The introduction of French words like *veal*, *zeal*, etc. led to the development of contrasts, e.g., between *feel* and *veal*, *seal* and *zeal*, leading to a phonemic opposition between the voiced and voiceless fricatives. Similarly, lexical borrowing led to the phonemicizing of /č/ vs /ǰ/ and [š] vs [ž]. On the whole, however, phonological changes were few, confined to the pairs above. No new sounds were introduced into English. Moreover, the tendency toward phonemicization of certain allophonic pairs may have existed even before French influence intervened. For example, Kurath (1956) argues that the loss of geminate consonants in words like [pyfan] (< pyffan) may have created a contrast between intervocalic [f] and the [v] in words like [dri:van] ‘drive’. Also, internal developments such as the loss or reduction of endings and lexical borrow-

ing from Old Norse may have contributed to these changes. At any rate, English phonology changed rather little under direct French influence. Yet neither English phonology nor morphology changed much under French influence. Sounds like [f] and [v] already existed in Middle English as non-distinctive allophones, so only minor adjustment was involved.

Lexical borrowing from French also had some influence on English morphology, particularly on derivational processes. It introduced several derivational affixes such as the prefixes in *dis-connect*, *de-flee*, *en-rich*, *em-bolden*, etc. Similarly, items like *cert-ify*, *charit-able*, *declar-acioun*, *statu-ette*, etc., yielded various suffixes, some of which became relatively productive as early as the Middle English period itself. For instance, the adjective-forming suffix *-able* was soon employed with native stems to yield words like *speakeable*, *knowable*, etc. (Dalton-Puffer 1996). In general, however, relatively few of the many French affixes that had been imported became productive, and the vast majority of French loans underwent adaptation to English morphological processes.

The important point, for our purposes, is that both the phonological and morphological innovations were introduced indirectly through lexical borrowing. Middle English speakers clearly did not isolate morphemes like *-able* in the relevant French words and import them independently of the stems to which they were attached. Thomason and Kaufman (1988:106) discuss a similar situation in Kormakiti Arabic, where lexical borrowing was the source of various structural innovations. While such innovations are clearly borrowings, they were not directly imported in either of these cases. In fact, there seems to be much support for the traditional view that direct structural borrowing is subject to very strong constraints, as has long been argued by linguists such as Meillet, Sapir, and others.

As noted above, direct borrowing of structural elements can occur only when the languages involved are typologically very similar, allowing for the substitution of an RL morpheme by a close counterpart in the SL. We leave aside, for the moment, the direct borrowing of function words, especially conjunctions and prepositions, which appears to occur quite frequently. For example, many indigenous languages in the Americas have borrowed conjunctions like *pero* 'but' and *como* 'as, like' from Spanish. This kind of borrowing is more akin to lexical than to structural borrowing, and like the former, it tends to have little or no impact on the structure of the RL.

A well-known case of structural borrowing is the contact between Ritharngu and Ngandi, two Aboriginal languages spoken in Arnhem Land, Australia (Heath 1981). The Ritharngu (Ri) group is the larger one, and speakers of Ngandi (Ng) are gradually shifting to Ri, so much so that their language is on the path to extinction, being now restricted to a few fluent speakers. This asymmetrical relationship has favored massive lexical borrowing across the two languages, but more so by Ng, which has also adopted several bound morphemes from Ri. Many of these structural borrowings were facilitated either by heavy lexical borrowing or by a close typological fit between the languages in the relevant subparts of their grammars. Once more, however, the overall structure of Ng has not been seriously affected.

This contrasts with the marked structural influence that Ng has had on Ri, as speakers of the former shift to the latter. This influence can be found in phonology (e.g., the distribution of glottal stops); in morphology (e.g., the transfer of enclitic pronouns marking subjects and objects); in morphosyntax (e.g., the emergence of new TMA categories); and in syntax (e.g., the strategy of creating relative clauses by attaching a subordinating suffix to a clause). (See Heath 1978:126ff.) These innovations are clear instances of impositions, in that Ng speakers partially adapted Ri to their own native grammar. The contact between these two languages demonstrates clearly the distinction between borrowing under RL agentivity and imposition under SL agentivity. It also demonstrates that structural borrowing is subject to much stricter constraints than structural imposition, and has much less impact on the grammar of the RL than the latter.

If it is true that direct borrowing (imitation) of structural features is so constrained, how can we explain the sometimes extensive changes that have occurred in maintained languages under influence from external source languages? The answer lies in two factors, the degree of bilingualism involved and the extent to which bilinguals are dominant in one or the other language. It is well known that situations in which a maintained language has undergone significant contact-induced change invariably involve extensive bilingualism. In these cases, the distinction we referred to earlier between the agents of change and the types of agentivity becomes especially important, since it helps us better understand the mechanisms by which structural change has occurred. In fact, both kinds of agentivity may be involved in such situations and can be implemented by the same agents.

3.2 Intertwined languages

When the agents of change are RL dominant, the changes they introduce from the SL are more likely to involve mostly lexical borrowing under RL agentivity. This process can be carried to an extreme, resulting in the creation of mixed or intertwined languages such as Media Lengua, Michif, and others to be discussed below. In simple terms, Media Lengua is a blend of Quechua grammar and Spanish-derived stems (mostly nouns, verbs, and adjectives) to which Quechua grammatical affixes are added. Borrowings from Spanish also include function or closed-class items like prepositions, conjunctions, and personal pronouns. But all of these, like the stems referred to above, have been adapted to Quechua morphology and syntax. (See Muysken 1981 and 1997 for details.) The following example from Muysken (1981:68) will serve as illustration (Spanish items are in italics):

(1) ML: *No* *sabi*-ni-chu Xwan *bini*-skda-da
 NEG know-1SG-NEG John come-NOM-ACC
 ‘I don’t know that John has come’

Q: Mana *yacha*-ni-chu Xwan *shamu*-shka-da
 NEG know-1SG-NEG John come-NOM-ACC

Sp: No sé que Juan ha venido
 NEG I-know that John has come

Note that the Spanish forms simply substitute for the Quechua forms without changing the underlying system. In general, the grammatical features imported from Spanish into ML were relatively few, despite the massive incorporation of free forms. Note also that practically no bound morphology was incorporated into ML from Spanish. The few exceptions include the diminutive suffix *-itu* (< *-ito/-ita* as in *muchachito/a* < *muchacho/a* ‘boy/girl’), and the past participle *-do* as in *cansado* ‘tired’ < *cansar* ‘to tire’. Both features also occur in Quechua, where they are clearly borrowings, and it is clear that the derivational suffixes were not incorporated directly, but only as parts of words borrowed as wholes, as we saw in the case of Middle English.

In short, the patterns of incorporation of free forms into a maintained structural frame, and the adaptation of such forms to Quechua grammar (including phonology) are exactly what we would expect in cases of (mostly lexical) borrowing under RL agentivity. These characteristics are clear evidence that ML was created by Quechua-dominant bilinguals. The strategies we find here are also found in cases of “classic” code-switching of the type that involves insertion of embedded-language content morphemes into the morphosyntactic frame of a matrix language (Myers-Scotton 2002:105). Heath (1978) and others have also compared this type of code-switching with borrowing.

ML is a good example of contact situations in which a maintained ancestral language is the dominant language as well as the recipient language. But what about situations in which the grammatical structure of the resulting contact language comes, not from the ancestral language, but from an external SL? Such situations are of two types. The first involves intertwined languages very similar to Media Lengua, such as Anglo-Romani and Ma’a. The second involves ancestral languages that have undergone massive structural change under external influence, for instance, Asia Minor Greek. I argue that the mechanisms and processes by which Ma’a and Anglo-Romani emerged were the same as those that gave rise to Media Lengua. In other words, they are all akin to cases of borrowing under RL agentivity. On the other hand, I argue that languages like Asia Minor Greek arose primarily through processes of imposition via SL agentivity. Let us consider each case in turn.

4 The case of Anglo-Romani and similar intertwined languages

Scholars have offered different explanations with regard to how intertwined languages were formed, and in some cases the same scholars have taken contradictory positions on the same language. For instance, Anglo-Romani has been characterized as a case of shift to English with consequent incorporation of lexicon from Romani. On the other hand, it has been claimed that Ma’a arose via a process of gradual grammatical replacement, that is, structural borrowing. The same has been argued for languages like Asia Minor Greek, whose grammars have changed dramatically under sustained external influence. Let us consider each of these types of situation in the light of the distinctions between borrowing and imposition discussed above.

Anglo-Romani is spoken by Roma or Gypsy groups in the British Isles. Its grammatical frame is English, but most of its lexicon comes from Romani, the ancestral language. In this respect, it is the converse of *Media Lengua*. Thomason and Kaufman (1988) offer somewhat conflicting claims concerning the origin of the former language. At one point, they argue that “a case like Anglo-Romani apparently represents actual language shift with maintenance of Romani vocabulary” (1988:49). This would imply that English was the matrix language into which Romani lexicon was incorporated, though they do not say this explicitly. Elsewhere, however, they characterize the language as a case of “complete grammatical replacement” due to “extensive borrowing” (1988:103). This presumably means extensive structural borrowing from English into Romani. It is not clear how to reconcile these two statements, or how to interpret their description of the actual processes involved in the creation of Anglo-Romani. For instance, they describe the language as the result of “two entirely distinct historical processes [sic]: inherited vocabulary, borrowed grammar” (ibid.). It’s not clear what kind of processes they have in mind here, and how they relate to the actual mechanisms by which Anglo-Romani was created. In other words, it is not clear whether they equate historical processes with psycholinguistic ones. Moreover, the implication of their statements seems to be that language shift can be equated with extreme grammatical borrowing, which I argue is dubious at best.

Thomason (1995:23) considers the suggestion, made by Boretzky (1985), that Romani lexicon was incorporated into an English frame. This suggests that Anglo-Romani arose after the Roma had shifted to English, and that English was the dominant language into which lexical items from Romani were incorporated. In our terms, this would be a case of massive lexical borrowing, under RL agentivity, similar in kind to *Media Lengua*, except that the RL in this case is not the ancestral language, but the one shifted to. This appears to be the generally accepted view among scholars, though Thomason still seems to maintain that Anglo-Romani, “is the end product of massive structural borrowing” (1995:24).

Given the mostly unanimous consensus on Anglo-Romani, it is surprising that other intertwined languages that arose under similar circumstances have been explained in quite different ways by some researchers. For instance, Thomason (1995:24) unequivocally attributes the formation of Ma’a and Caló (an intertwined language with Spanish grammar and Romani vocabulary) to “massive structural borrowing”. Thomason and Kaufman (1988:50) place situations like this at the outer limits of their continuum of “borrowing” situations, where extreme structural borrowing has occurred. They draw a sharp distinction between Anglo-Romani and Ma’a, arguing that the former represents “actual language shift with maintenance of Romani vocabulary”, while, in the case of Ma’a, “no shift has occurred, but almost all of the original Cushitic grammar and at least half ... of the Cushitic vocabulary have been replaced by Bantu grammar and lexicon” (1988:49).

Again, it is somewhat confusing that the term they use here to explain the origins of Ma’a, that is, “grammatical replacement”, is the same one they used with respect to Anglo-Romani. “Replacement” can come about in different ways. However, it is quite

clear that in the case of Ma'a they intend this term to mean massive structural borrowing. But to assign such extreme changes in grammar to borrowing flies in the face of all we know about the strong constraints on structural borrowing under RL agentivity. Moreover, given the close similarity in make-up between Anglo-Romani and Ma'a, it seems counterintuitive and uneconomical to ascribe the former to shift accompanied by lexical retention and the latter to lexical retention accompanied by massive structural borrowing. Economy would suggest that Ma'a arose in the same way as Anglo-Romani, that is, after the Ma'a shifted to a Bantu language. Under this scenario, their newly acquired language then served as the matrix language into which they incorporated lexical items from their original ancestral language. In other words, the creation of Ma'a involved the same RL agentivity that we found in the case of Media Lengua and Anglo-Romani. This is similar to the position taken by scholars such as Bakker (1997), Brenzinger (1992), and Sasse (1992). It finds support in the fact that the Ma'a also speak a variety of Bantu (Mbugu) whose grammar is closely similar to that of Pare, the language of the surrounding group (Mous 1994:176). This variety differs from Ma'a (also referred to as "inner Mbugu") only in vocabulary and minor structural features. Adopting the above scenario would mean that we have a unified explanation that allows us to classify these contact languages as a single type, as well as to recognize the similar psycholinguistic processes (as distinct from the historical circumstances) by which they came into being.

5 Ongoing language shift and types of agentivity

The cases we considered in the previous sections all involve situations where the RL is clearly dominant, and RL agentivity is the primary factor in the changes that occur in it. In most cases, the RL is a maintained language, or the group's primary language. But what are we to make of situations, such as Asia Minor Greek, where it is clear that extreme structural changes have occurred in an ancestral language under the influence of a politically dominant external language, while the ancestral language is still maintained?

As noted earlier, the tendency is for scholars to assume that any change in a maintained language must be due to borrowing in the first instance. This, presumably, is why Thomason and Kaufman (1988:215) assign languages like Asia Minor Greek and Wutun to level 5 of their borrowing scale, arguing that they arose via massive grammatical borrowing. But a close examination of the structural features in question casts doubt on this claim. Let us consider the changes that occurred in Asia Minor Greek under Turkish influence.

The Cappadocian variety of Greek in particular was influenced far more heavily than those in areas like Silli and Phárasa, while varieties in Pontus in the west of Turkey displayed even less influence. Turkish influence on Cappadocian Greek was pervasive, encompassing the lexicon, phonology, morphology, and syntax. In phonology, for instance, we find reduction of Greek phonological contrasts via elimination of sounds such as /ð/ and /θ/, sounds not found in Turkish. Vowel harmony on the Turkish pattern is found on Greek suffixes attached to Turkish words, while a number of new phonemes derived from Turkish, e.g., /ö, ü, č, ĵ/ entered the Cappadocian variety. Various Turkish-derived morphophonemic rules also appear in the dialect.

In morphology, we find loss of inflections on adjectives (Turkish lacks these) and the use of partly agglutinative strategies of noun and verb inflection, which again follows the Turkish pattern. Similarly, Turkish-derived innovations are found in the tense/aspect system, and in various aspects of syntax, including word order.

Thomason and Kaufman (1988:218) argue that “if Turks did not shift to Greek, all of the interference must be due to borrowing”. This once more illustrates the tendency to equate changes in maintained languages only with borrowing. Moreover, the implication is that changes must have been introduced by speakers who were monolingual or more proficient in Greek, that is, via RL agentivity. This overlooks the strong probability that bilinguals, especially those that were Turkish-dominant, played a key role in introducing these changes. Here again, then, the distinction between agents and types of agentivity becomes crucial.

The nature of the changes that occurred in Asia Minor Greek would seem to indicate that both types of agentivity acted in concert, with Greek-dominant bilinguals implementing RL agentivity, and Turkish-dominant bilinguals (especially children, perhaps) implementing SL agentivity. And some bilinguals may have implemented both types simultaneously. At any rate, the notion of borrowing, as we have defined it here, seems quite inappropriate to explain most of the deep and pervasive changes that occurred throughout the grammar of Cappadocian Greek. Given the strong constraints that apply to borrowing, especially of structural features, such changes could only have come about through the mechanism of imposition, involving adaptation of Greek to Turkish, rather than the other way around.

This scenario is in keeping with van Coetsem’s (1988:83) observation that “the linguistic dominance relation between the RL and the SL ... determines whether RL or SL agentivity will result from the contact”. It follows that reversals in this dominance relationship will lead to changes in types of agentivity. We see this especially in cases where speakers gradually lose competence in their ancestral language as they become linguistically dominant in a language they acquire later (*ibid.*). An approach like this allows for a unified treatment of languages similar to Asia Minor Greek that have been accounted for in terms of “interference due to shift”. The latter include Ethiopic Semitic, Shina, Irish English, and others that Thomason and Kaufman (1988:128–39) treat as unambiguous cases of shift with substratum influence, or, in our terms, as cases of imposition under SL agentivity. Once more, it seems uneconomical to argue for borrowing in cases like Asia Minor Greek and shift-induced interference in others, when the structural changes involved are so similar. It seems more likely that such similarities must be due to the same mechanisms of change.

We can cite a variety of other cases where bilinguals who have become dominant in a newly-acquired second language promote structural changes in their ancestral language via SL agentivity. For instance, Silva-Corvalán (1994) discusses several changes in Los Angeles (LA) Spanish that can be attributed to influence from English, which is the socially dominant language, and has become, for many speakers, the linguistically dom-

inant language as well. One example is Spanish *atrás* ‘behind’, which has acquired the sense of English *back* (Silva-Corvalán 2000:14), as in the following example:

- (2) Se lo dió p’atrás.
to-him it she-gave back
‘She gave it back to him.’

The counterpart of this in general Spanish would be as follows:

- (3) Gen Span. se lo volvió
to-him it she-returned

Dar atrás is clearly a calque on English *give back*, replacing the use of *volver* ‘return’. Changes like these are common, even in the speech of persons quite competent in Spanish. In speakers with reduced competence in Spanish, we find even more extreme cases of calquing on English, such as the following (Silvia-Corvalán 1998:233):

- (4) LA Span. Yo gusto eso.
I like-1s that
Gen Span. A mi me gusta eso.
to me pro please-3s that
‘I like that.’

Here, *gustar*, which has a theme or patient subject and an indirect experiencer object in general Spanish, is reanalyzed as a transitive verb with an experiencer subject and an accusative theme, on the model of English *like*.

Similar changes can be observed in Prince Edward Island French, spoken in Eastern Canada. This Acadian variety has been subjected to strong influence from English, mostly involving lexical borrowing (King 2000). For instance, English prepositions and phrasal verbs have been incorporated into the French variety, yielding forms such as *ender up*, *finder out*, etc. It is also common to find various kinds of mixture of French verbs with English prepositions (*faire up* ‘make up’), while French prepositions occur with English-derived verbs (*picker su* ‘pick on’). One consequence of this is that preposition stranding, found in English but not in French, is now a common feature of Prince Edward Island French. King (2000:136) argues that this change is not the result of direct structural borrowing, but rather “lexical borrowing has triggered reanalysis of the PEI French prepositional system”. It seems likely that this reanalysis was the result of imposition, the agents of which were fluent bilinguals who practiced code-switching. English-dominant bilinguals would have been particularly likely to impose this structural change on their French. King notes that several other structural changes have also been introduced into PEI French, presumably by the same mechanism.

but in borrowing, imitation comes into play before adaptation, while the reverse obtains in impositions (*ibid.*). If there is close imitation of an SL feature in borrowing, it may lead to a deviation from the RL pattern, as when English speakers pronounce Bach as /baɣ/, using a phone not found in their L1. Such deviations are rare, and do not typically affect the RL. In the vast majority of cases, imitated SL items are adapted to RL structure, as we have seen.

Adaptation can produce quite similar results in both borrowing and imposition (van Coetsem 1988:12). By way of illustration, let us consider how English-derived words are adapted by Hindi speakers in both RL and SL agentivity. Hock (1991:393) discusses how English stops and fricatives are substituted by perceived equivalents in Hindi when borrowed into the latter. For instance, English aspirated stops (/p, t, k/) are replaced by Hindi unaspirated stops ([p, t, k]), while English fricatives (/f, θ/) are replaced by Hindi aspirated stops ([p^h, t^h]). (See Hock 1991:394 for an explanation of these substitutions.) When speakers of Hindi speak English, they adapt English sounds in precisely the same way; this is a well-known feature of Indian English. The similarity in outcomes may explain the tendency to confuse the two major mechanisms and their associated types of agentivity. In both cases, the agents of change are adapting materials from an external language to fit the structure of their dominant language. In borrowing, they preserve this structure, particularly the more stable domains of grammar, such as phonology, morphology, and most, if not all, aspects of morphosyntax. In imposition, they transfer varying degrees of their L1 structure to an external recipient language. In many cases, the results of these distinct mechanisms do not, by themselves, indicate which mechanism was involved, in the absence of sound socio-historical evidence.

In some cases the effects of imposition gradually disappear as speakers achieve greater proficiency in the RL, that is, the target language (TL). In other cases, however, these effects multiply and are reinforced by social factors, e.g., lack of access to native varieties of the TL, or the cumulative influence of similar source languages (L1s). Many creoles and extended pidgins arose in this way. The kinds of adaptation that take place in these cases have been described in a variety of ways, as “transfer”, “substratum influence”, “relexification”, “reanalysis”, “convergence”, and so on. But such terms are sometimes used without a clear determination of the types of agentivity involved.

For instance, the term “relexification” has been used to describe the reinterpretation or relabeling of superstrate lexical forms in terms of substrate semantic and morpho-syntactic categories, as found in creole formation (Lefebvre 1996, 1998). The same term was used by Muysken (1981) to refer to the importation and adaptation of SL lexical forms into the unchanged structural frame of an RL, as in the formation of *Media Lengua*. In the present approach, the latter is a case of RL agentivity, identical with what goes on in classic code-switching. To apply the same term to creole formation would imply that the latter involved importation of superstrate forms into a substrate structure that was maintained (that is, RL agentivity). If that were true, creoles would be indistinguishable from bilingual mixed languages, or cases of classic code-switching. By contrast, the position adopted here is that the processes by which creoles were formed involved imposition of varying degrees under SL agentivity, as well as other processes

such as reduction, simplification, and internal innovations also found in the more usual cases of second language acquisition. Such processes operated to varying degrees in different creoles, over varying periods of time, yielding outcomes that were quite different. This means, of course, that the term “creole” must be viewed as a convenient label for languages that share a certain socio-historical background, rather than as a typological designation.

8 Abstract lexical structure and contact-induced change

Another way of approaching a classification of contact-induced changes and their outcomes is to recognize that they all involve processes by which different aspects of RL and SL lexical structures are re-combined to form new lexical entries. However, the nature of the recombination can differ in significant ways, yielding very different kinds of contact phenomena. Myers-Scotton’s work on language contact has been particularly instructive about the ways lexical entries may be re-constituted in bilingual contact situations. Her approach is based on psycholinguistic models of language production, which distinguish three levels or stages of the language production process roughly represented here as follows. (See Levelt 1989:9 for a more detailed representation.)

- The Conceptual level: The messages the speaker intends to convey.
(the Conceptualizer)
- The Functional level Lemmas (abstract entries in a speaker’s mental lexicon) are
(the Formulator): accessed. Lemmas activate morpho-syntactic procedures (e.g.,
argument structure and morphological realization patterns)
- The Positional level: Phonological representations and surface structure are
realized.

A lexical entry consists of a word form or phonological shape, which I simply call a lexical item, its various morphological shapes, and a lemma associated with it. The latter contains information about the semantic, morphological, syntactic, and other properties of the item. In monolingual language production, once a lemma is accessed by the Formulator, it activates the morphosyntactic procedures associated with the relevant lexical items. In bilingual language production, differences arise in the way lemmas are accessed and associated with SL and RL lexical items. The reconstituted lexical entries may differ depending on which aspects of the original lexical entries are involved.

This approach allows us to explain, to some extent, the similarities and the differences between adaptation in borrowing and adaptation in imposition. In lexical borrowing, a new phonological form is introduced to an RL, with its own (often modified) semantic content. In most cases, such items assume all of the formal and structural properties (including the phonological structure) of similar RL items. This is the case in most instances of lexical borrowing, as well as in classic code-switching involving single content morphemes, as described earlier. In the latter case, of course, the imported items may substitute for items with similar meaning in the RL, whereas lexical borrowing often introduces items that have no counterparts in the RL. The point is that, in these cases, only the phonological shapes (and some of the semantics) are new to the RL.

This, in fact, is the process that Muysken (1981) described as relexification, which he defined as “the process of vocabulary substitution in which the only information adopted from the target language in the lexical entry is the phonological representation” (1981:61). As we saw earlier, Muysken proposed that Media Lengua arose via this process. Figure 1 illustrates. It is clear that there is no real difference between what happened in Media Lengua and what happens in classic code-switching and other types of lexical borrowing under RL agentivity.

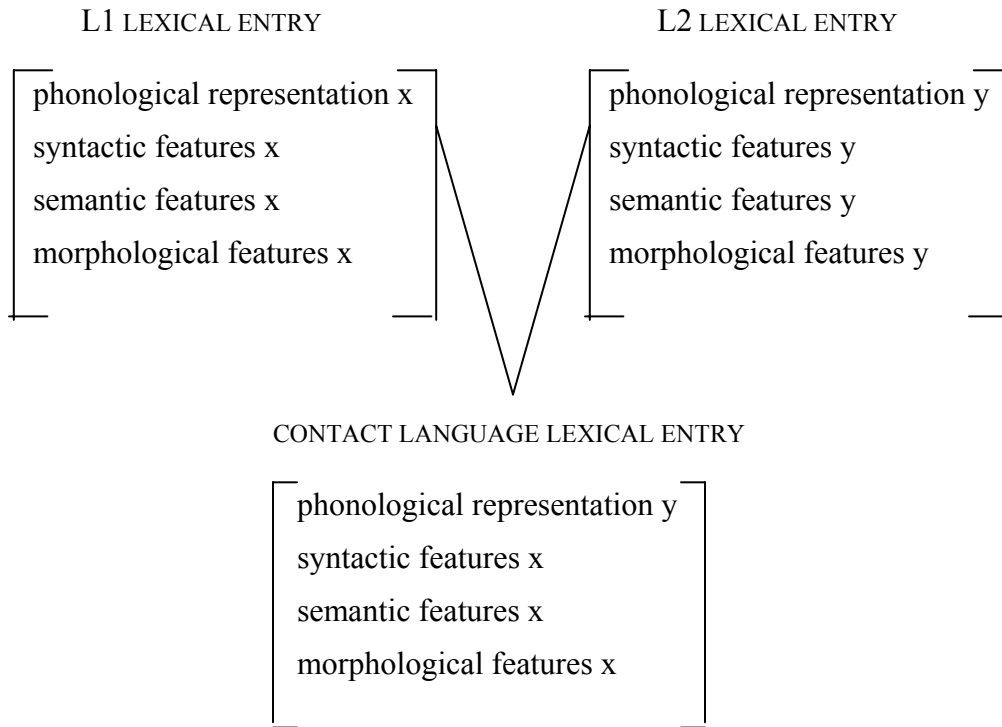


Figure 1. The process of relexification.

In imposition, by contrast, an RL item is adapted so that part of its abstract lexical structure (usually its phonological representation) derives from the SL, and only part, if any, of the rest of its original lexical structure is preserved. Imposition also involves the reconstitution of lexical entries (among other things), in which phonological forms derived from an external RL (usually a target language) are adapted in varying degrees to the properties of perceived equivalents in the L1 (as SL). Unlike lexical borrowing, however, this kind of adaptation allows for various types of combination of RL and SL lexical entries, in ways peculiar to imposition.

The “Abstract Level Model” of codeswitching introduced by Myers-Scotton and Jake (1995) offers a useful way to explain these kinds of contact phenomena, which are quite distinct from classic codeswitching, as well as other kinds of contact phenomena associated with RL agentivity. As Myers-Scotton (2002:19) points out,

In these phenomena, we cannot argue that all the abstract structure is derived from the grammar of one of the participating languages; rather it is clear that more than one language is the source of structure.

The model is based on the assumption that “all lemmas in the mental lexicon include three levels of abstract lexical structure”, namely:

Lexical-conceptual structure;
 Predicate-argument structure;
 Morphological realization patterns. (Myers-Scotton 2002:194)

Myers-Scotton argues that one or more of these levels from a lexical entry in one language can be split and recombined with levels in another language (2002:99). Like the present approach, Myers-Scotton distinguishes the processes involved in lexical borrowing and classic code-switching from those involved in convergence, which she treats as both a process and an outcome (101). She describes the process as “largely a one-way phenomenon ... [that] involves the grammar and lexicon of a source language, generally one that has more socioeconomic prestige, impinging on another language” (2002:172). It is clear she has in mind a process similar to that referred to here as imposition. Moreover, like the present approach, she identifies this “convergence” as “a mechanism in the progressive outcomes of attrition, language shift, language death and creole formation” (101). In all these cases, the abstract lexical structure of items derived from one language can change significantly due to imposition of lexical structure at different levels, from another language. Examples of this would include some of the structural changes described earlier in the English-influenced Spanish of bilinguals in LA, for example, the reinterpretation of *gustar* as a transitive verb with the argument structure of English *like*. This process may become pervasive in certain cases of contact, leading to significant degrees of structural convergence between languages.

Extreme cases of this reconfiguration can be found in creole formation. An example is the reanalysis of English preposition *there* as the locative/existential copula *de*, which in turn was reanalyzed as the Progressive/Imperfective marker in several Caribbean English-lexicon creoles. The model for this was the fact that principal substrate languages such as Gbe employed the same item as both a locative copula and a marker of Progressive aspect. When substrate speakers were confronted with English sentences such as *John there (in the yard)*, they established an interlingual identification between this *there* (pronounced /de/) and their L1 locative/existential copulas, leading to the reanalysis just described (Migge 2002, Winford 2003). This process occurred to varying extents in different creoles, and was carried to an extreme in the more “radical” creoles, such as those in Suriname. The extreme cases of adaptation, in which only a phonological representation derives from the superstrate, are indistinguishable in some respects from the phenomena associated with relexification in the case of Media Lengua. The difference is that neither the morphological realization patterns nor the full argument structure of the substrate languages were preserved, even in the most radical cases. Differences like these have to be accounted for in terms of other processes, such as simplification, leveling, and internal restructuring, which were characteristic of creole formation.

9 Conclusion

This paper has discussed two broad mechanisms by which languages in contact influence each other. Despite differences in approach, there seems to be general consensus on the role and nature of these two mechanisms. However, the variety of terms used to describe the actual mechanisms and their attendant processes has led to some degree of confusion, as witness the conflicting uses of terms like “relexification”, “convergence”, “transfer”, and the like. All of these terms have been used, for instance, to explain the process of creole formation. Unfortunately also, lack of precision and consistency in the use of such terms has led to conflicting classifications of the outcomes of contact. Cases of language shift involving structural assimilation of an RL to an SL, such as Asia Minor Greek, have been described as instances of “structural borrowing” by some, “convergence” by others. Perhaps most importantly of all, we have tended to ignore or overlook the similarities in the processes associated with lexical borrowing, classic code-switching and language intertwining on the one hand, and the similarities in the processes associated with second language acquisition, language shift and attrition, and creole formation on the other. The approach suggested here, based on van Coetsem’s distinction between the mechanisms of borrowing under RL agentivity, and imposition under SL agentivity, with their shared but differently implemented processes of imitation and adaptation, seeks to provide a more consistent framework in which to investigate the outcomes of contact.

References

- BAKKER, PETER. 1997. *A language of our own: The genesis of Michif, the mixed Cree-French language of the Canadian Métis*. Oxford: Oxford University Press.
- BREZINGER, MATTHIAS. 1992a. Lexical retention in language shift: Yaaku/Mukogodo-Maasai and Elmolo/Elmolo-Samburu. In Brenzinger 1992b, 213–54.
- BREZINGER, MATTHIAS (ed.) 1992b. *Language death: Factual and theoretical explorations with special reference to East Africa*. Berlin: Mouton de Gruyter.
- DALTON-PUFFER, CHRISTIANE. 1996. *The French influence on Middle English morphology*. Berlin: Mouton de Gruyter.
- HAMMARBERG, BJÖRN. 1997. Conditions on transfer in phonology. *Second-language speech: Structure and process*, ed. by Allan James and Jonathan Leather, 161–80. Berlin: Mouton de Gruyter.
- HAUGEN, EINAR. 1950. The analysis of linguistic borrowing. *Language* 26.210–31.
- HAUGEN, EINAR. 1953. *The Norwegian language in America: A study in bilingual behavior. Vol. I: The Bilingual Community; Vol. II: The American dialects of Norwegian*. Bloomington: Indiana University Press. (Reprinted 1969.)

- HEATH, JEFFREY. 1981. A case of intensive lexical diffusion. *Language* 57.335–67.
- HEATH, JEFFREY. 1984. Language contact and language change. *Annual Review of Anthropology* 13.367–84.
- HOCK, HANS. 1991. Principles of historical linguistics. 2nd edn. Berlin: Mouton de Gruyter.
- KING, RUTH. 2000. *The lexical basis of grammatical borrowing: A Prince Edward Island case study*. Amsterdam: John Benjamins.
- KURATH, HANS. 1956. The loss of long consonants and the rise of voiced fricatives in Middle English. *Language* 32.435–45.
- LEFEBVRE, CLAIRE. 1996. The tense, mood, and aspect system of Haitian Creole and the problem of transmission of grammar in creole genesis. *Journal of Pidgin Creole Languages* 11.2.231–311.
- LEFEBVRE, CLAIRE. 1998. *Creole genesis and the acquisition of grammar*. Cambridge: Cambridge University Press.
- LEVELT, WILLEM. 1989. *Speaking, from intention to articulation*. Cambridge, MA: MIT Press.
- MIGGE, BETTINA. 2002. The origin of the copulas (*d/n*)*a* and *de* in the Eastern Maroon Creole. *Diachronica* 19.83–136.
- MOUS, MAARTEN. 1994. Ma'a or Mbugu. *Mixed languages: 15 case studies in language intertwining*, ed. by Peter Bakker and Maarten Mous, 175–200. Amsterdam: Institute for Functional Research into Language and Language Use (IFOTT).
- MUYSKEN, PIETER. 1981. Halfway between Quechua and Spanish: The case for relexification. *Historicity and variation in creole studies*, ed. by Arnold Highfield and Albert Valdman, 52–78. Ann Arbor: Karoma.
- MUYSKEN, PIETER. 1997. Media Lengua. *Contact languages: A wider perspective*, ed. by Sarah G. Thomason, 365–426. Amsterdam: John Benjamins.
- MYERS-SCOTTON, CAROL. 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- MYERS-SCOTTON, CAROL, and JANICE JAKE. 1995. Matching lemmas in a bilingual competence and production model: Evidence from intra-sentential code switching. *Linguistics* 33.981–1024.

- NISHIMURA, MIWA. 1986. Intrasentential code switching: The case of language assignment. *Bilinguals: Psycholinguistic and neuropsychological perspectives*, ed. by J. Vaid, 123–43. Hillsdale, NJ: Erlbaum.
- NISHIMURA, MIWA. 1997. *Japanese/English code switching: Syntax and pragmatics*. New York: Peter Lang.
- SASSE, HANS-JÜRGEN. 1992. Theory of language death. In Brenzinger 1992b, 7–30.
- SILVA-CORVALÁN, CARMEN. 1994. *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon.
- SILVA-CORVALÁN, CARMEN. 1998. On borrowing as a mechanism of syntactic change. *Romance linguistics: Theoretical perspectives*, ed. by Armin Schwegler, Bernard Tranel, and Myriam Uribe-Etxebarria, 225–46. Amsterdam: John Benjamins.
- SILVA-CORVALÁN, CARMEN. 2000. El contacto Español-Inglés en Los Angeles: Aspectos demográficos y sociolingüísticos. Paper presented at UC Davis.
- THOMASON, SARAH GREY. 1995. Language mixture: Ordinary processes, extraordinary results. *Spanish in four continents: Studies in language contact and bilingualism*, ed. by Carmen Silva-Corvalán, 15–33. Washington, DC: Georgetown University Press.
- THOMASON, SARAH GREY, and TERRENCE KAUFMAN. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
- VAN COETSEM, FRANS. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- WEINREICH, URIEL. 1953 [1986]. *Languages in contact: Findings and problems*. New York: Linguistic Circle of New York. [Reprinted 1986, The Hague: Mouton.]
- WINFORD, DONALD. 2003. Sranan TMA and substrate influence. Paper presented at the workshop Transatlantic Sprachbund, NIAS, Holland.