SOCIAL AND LINGUISTIC FACTORS AS PREDICTORS OF CONTACT-INDUCED CHANGE

Sarah G. Thomason
University of Michigan

1. Introduction

Two claims made by Thomason & Kaufman (1988) have elicited particularly strong reactions from specialists in language contact: first, that there are no absolute linguistic constraints on the kinds or numbers of features that can be transferred from one language to another; and second, that when social factors and linguistic factors might be expected to produce opposite results in a language contact situation, the social factors will be the primary determinants of the linguistic outcome. Both claims have frequently been challenged in recent years. Some of these challenges are based on a misunderstanding of our arguments; most seriously, some critics argue that we dismiss linguistic predictors as entirely irrelevant to an analysis of contact-induced change.

More interesting objections to our position are based on genuine theoretical and/or empirical disagreements between Thomason & Kaufman (and also Thomason 2001, among other writings) and the critics. This paper explores one set of disagreements in an effort to arrive at a better understanding of the relative importance of social and linguistic predictors in language contact situations. I will argue that there is still good reason to view social factors as primary and linguistic factors as secondary in predicting contact-induced change, and that it is still true that no one has successfully proposed absolute linguistic constraints on contact-induced change. (I should emphasize that throughout this paper phrases like ‘linguistic constraints on contact-induced change’ should be read in the traditional meaning of the term ‘constraint’—i.e., ‘it cannot happen’—rather than in the much more flexible ‘constraint’ of the sort popularized by Optimality Theory.)

The structure of the paper is as follows. After reviewing a few of the critics’ most important criticisms (§2), I will set the stage for my own analysis in §3, by presenting a historical linguist’s starting point in the search for causes of language change, by giving a definition of contact-induced change, and by outlining what is required to support a claim that contact-induced change has occurred. In §4 I will explain why I reject theories that claim, or assume, the existence of absolute linguistic constraints on such changes, and in §5 I will conclude by returning to the question of social vs. linguistic predictors of contact-induced change.
2. What do the critics say?

First, some critics have interpreted the Thomason & Kaufman position as a complete rejection of any role for strictly linguistic factors in processes of contact-induced change. One example is in Gillian Sankoff’s important article ‘Linguistic outcomes of language contact’ (2001:640-641):

Lacking a quantitative perspective...T&K are forced to deny the importance of internal linguistic factors. The burden of T&K’s argument is that, given enough social pressure, anything can happen language internally, and they adduce examples in which suggested internal, structural constraints have been overridden. Sociolinguists have, understandably, been largely approving of the pride of place T&K attribute to social constraints. However, in rejecting the contribution of internal linguistic structure, T&K have thrown out the baby with the bathwater...T&K are very far from the truth in their blanket rejection of internal constraints. In this chapter, I will review literature from a quantitative sociolinguistic perspective, in which internal constraints have been shown to act jointly with external constraints in shaping language contact outcome.

Similarly, Ruth King (2002) says that Thomason & Kaufman consider social predictors alone, claiming that we

argue that social factors alone determine the linguistic outcomes of language contact; linguistic factors are said to play no role. Thus theirs is an “anything goes” perspective according to which elements from any linguistic subsystem may be borrowed.

Parts of these comments are puzzling—for instance, it isn’t clear what the lack of a quantitative perspective has to do with a putative denial of ‘the importance of internal linguistic factors’, since a great many non-quantitative studies have demonstrated the importance of linguistic factors in motivating contact-induced (and other) changes. In any case, Sankoff and King are by no means the only authors who have interpreted Thomason & Kaufman this way. But here’s what we actually say about social and linguistic factors (1988:4; emphasis added):

...the history of a language is a function of the history of its speakers, and not an independent phenomenon that can be thoroughly studied without reference to the social context in which it is embedded. We certainly do not deny the importance of purely linguistic factors such as pattern pressure and markedness considerations for a theory of language change, but the evidence from language contact shows that they are easily overridden when social factors push in another direction.

We actually agree with Sankoff’s view that ‘internal constraints...act jointly with external constraints in shaping language contact outcome’. Sizable parts of Thomason & Kaufman 1988, and of my writings since then (e.g. Thomason 2001), are in fact devoted to exploring the role of linguistic factors—especially typological distance between the source and receiving languages and considerations of universal markedness—in processes of contact-induced language change; our rough borrowing scale, to take just one example, offers a correlation between social and linguistic factors that (we argue) predicts which kinds of linguistic features are likely to be borrowed under which sets of social circumstances.

The issue of absolute constraints on contact-induced change, which is also raised by Sankoff & King, is more interesting, because it’s trickier. In chapter 2 of our 1988 book, Kaufman and I surveyed all the proposed linguistic constraints that we could find in the literature. It was possible, and usually easy, to find counterexamples to everyone of them. Since then, many scholars have reacted skeptically to our conclusion that, in the absence of any successful linguistic constraints in the available proposals, it is reasonable to claim that the burden of proof had shifted from our position (‘there are no absolute linguistic constraints’) to the ‘there are too absolute linguistic constraints’ position. That is, we argued that the default assumption should now be that absolute
constraints do not exist, because so far all efforts to come up with successful linguistic constraints have failed; and anyone who wants to claim that they do exist must offer proof, in the form of a successful absolute constraint.

The authors who have addressed this issue (at least the authors whose writings I’m familiar with) have not done this. Skeptical reactions have mostly been of two different kinds. One seems to be a general expression of distress at the thought that linguistic structure might not be all-powerful in this domain. This viewpoint is expressed, for instance, by Aikhenvald (2002:3):

It is hard to agree wholeheartedly with the basic starting-point of Thomason & Kaufman (1988:35) that “it is the sociolinguistic history of the speakers, and not the structure of their language, that is the primary determinant of the linguistic outcome of language contact.” Typologically different linguistic structures tend to change in different ways, even when the speakers share a great deal of sociolinguistic history.

In part Aikhenvald’s concern echoes Sankoff and King, but Aikhenvald introduces another angle with her comment about about the role of typology in guiding different changes. In this respect she is in tune with other critics who point out the absence of particular types of contact-induced changes in particular contact situations. (Since she talks about tendencies rather than absolutes, it may also be that she doesn’t disagree with us at all, in spite of her assertion that she does: we never claimed that all kinds of structural interference are equally probable; we merely claimed that none can be absolutely ruled out.)

Another prominent criticism that focuses on the absence of particular types of change in particular contact situations is Silva-Corvalán’s argument (1994:134) that

only those [linguistic features] that are compatible...with the structure of the borrowing language...will be adopted, disseminated, and passed on to new generations.

Silva-Corvalán’s conclusion is drawn from her research on Spanish-English contact in Los Angeles, a detailed and sophisticated analysis of a wide range of contact phenomena. The problem with her argument is that there is no reason to expect that EVERY kind of contact-induced change will occur in EVERY contact situation; even in the most intense contact situation, only some kinds will occur (Thomason 2000). As Sankoff suggests, there is a complex interplay between social and linguistic causal factors in any changes in contact situations, and it is wildly improbable that any pair of situations will have the exact same mix.

Since Silva-Corvalán makes no attempt to refute the counterexamples to her proposed universal constraint based on structural compatibility that are adduced (in response to previous authors’ similar claims) in Thomason & Kaufman (1988), her argument is, in effect, that it didn’t happen here, so it can’t happen anywhere. But this can’t be right: nobody, certainly not Thomason & Kaufman and as far as I know not anyone else either, claims that any kind of language change can be absolutely predicted, in any kind of situation. The fact that certain types of contact-induced change are possible in a given contact situation therefore does not mean that we can confidently expect to find them. In internally—as well as externally—motivated language change, even the most natural and common changes often fail to occur. To put it slightly differently, an intense language contact situation is a NECESSARY condition for extensive structural interference to occur, but it is not a SUFFICIENT condition. So the absence of typologically incompatible contact-induced innovations in Silva-Corvalán’s data, while interesting, has no relevance to the general claim that such changes are possible.

This point pertains to Aikhenvald’s argument as well. No matter how similar two (or more) sociolinguistic contexts are, no historical linguist would predict that the languages in that context will inevitably undergo identical changes. They won’t. This is how we get language families: even in the absence of structural interference from other languages, two separated subgroups of a single
speech community will, after about 500-1000 years have passed, turn into separate languages. And language families develop from a single parent language; in a contact situation, typologically disparate languages may be unrelated (as in Aikhenvald’s case study) or only distantly related, so there is even less reason to expect them to undergo similar or identical contact-induced changes.

Another relevant point in this context is that Aikhenvald sets her scene by specifying sociolinguistic homogeneity for the languages in the contact situation. If she’s right about this, then there are no different social factors that could motivate different kinds of contact-induced change in that situation; and in such a situation, of course linguistic factors will be the primary determinants of whatever differences there are in the contact-induced changes that occur. (I do not, however, believe that completely homogeneous social contexts can exist in complex contact situations; the existence of identical social motivators for the linguistic behavior of each language community involved seems very implausible.)

Other scholars appeal to a more abstract conceptualization of contact-induced change, and it is not always clear (at least to me) what the implications of their arguments are for the Thomason & Kaufman claim about the lack of absolute linguistic constraints on interference. King (e.g. 2000) and Myers-Scotton (e.g. 2002) propose lexically-based theories of contact-induced change. Myers-Scotton analyzes contact phenomena ‘in terms of models of language competence and production’ (2002:5), with a focus on ‘how parts of the mental lexicon may be organized’—more specifically, on how the elements of her ‘4-M Model’ interact in contact situations involving code-switching and/or convergence (by her special definition of ‘convergence’—essentially, apparently, any contact-induced change that does not involve transferred morphemes). In discussing examples from Thomason & Kaufman that involve transferred ‘late system morphemes’ (roughly, inflectional morphemes other than number markers), Myers-Scotton argues that ‘they are the result of a Matrix Language turnover that did not go to completion’, rather than (as Thomason & Kaufman would have it) ordinary, if unusually deep, structural borrowing. Her claim here seems dangerously close to circularity: the sole evidence for a partial Matrix Language turnover is the borrowing of ‘late-system morphemes’, so if ‘late-system morphemes’ are transferred, it must be a Matrix Language turnover. Without other (converging) criteria for establishing the existence of a partial Matrix Language turnover, it is difficult to see how such a claim could possibly be tested.

Even so, there are some examples in the literature that appear to be problematic for Myers-Scotton’s position. Two come from studies of the Ethiopian Highlands Sprachbund (Leslau 1945, 1952). The general Semitic causative formation employs a single prefix; but under the influence of Cushitic languages whose speakers shifted to Semitic in Ethiopia, Ethiopic Semitic forms the causative by means of a double prefix—a mirror of the typical Cushitic double-suffix causative formation. In a similar process, verbal negation in Ethiopic Semitic has been remodeled under Cushitic influence, from a prefix-only construction to a prefix-suffix combination; Cushitic languages make use of a suffix only. Both of these morphological developments involve ‘late-system morphemes’—inflectional processes. Although the actual phonological material is native Semitic in each case, the modern Ethiopic Semitic formations are partly inherited from Semitic and partly derived from Cushitic (in this case, via shift-induced interference; see §3.3 below). Crucially, neither is an exact match for either the Semitic or the Cushitic formation. Such partial transfer is a very common phenomenon in contact-induced change; it is one outcome of a process I’ve called ‘negotiation’ (e.g. in Thomason 2001:142-146, and see also Lindstedt 2000). The relevance to MyersScotton’s claim about ‘late-system morphemes’ is that, although both of these examples certainly comprise inflectional interference, they nevertheless constitute only a partial transfer of morphological material from one set of languages to another; so it looks as if analysis under a hypothesis of partial Matrix Language turnover is truly indeterminate. And this in turn does not seem to fit well in a theory that is (as far as I can tell) designed to provide a precise answer for every situation.
Another type of case that may be problematic for Myers-Scotton is highlighted by a comparison between morphological interference (in ‘late system’ morphemes) in Mednyj Aleut and Evenki. Mednyj Aleut is a bilingual mixed language that famously has an entirely Russian inflectional finite verb morphology, while its noun inflection and nonfinite verb inflection remain native Aleut constructions. Myers-Scotton classifies this case as ‘the prototypical example of the scenario proposed by the Matrix Language Turnover Hypothesis that ends in an arrested shift’ (2002:265).

Contrast this with (some varieties of?) Evenki, in which the volitional mood, and only the volitional mood, has both a mood suffix and an entire set of personal endings borrowed from the Turkic language Yakut (Malchukov 2003). The Evenki case provides the first hint in the literature, to the best of my knowledge, of a process that could eventually, through a sequence of subsequent changes, lead to the result we see in Mednyj Aleut: borrowed inflectional morphemes appear first in a marked, relatively rare paradigm; that’s where they remain (so far) in Evenki, but in Mednyj Aleut they have spread throughout the finite verbal system. The point, in the present context, is that this makes excellent sense from a standard historical linguistic viewpoint, where restructuring very often begins in less frequent paradigms and constructions, and only later (if ever) spreads to the system as a whole. Innovations tend to enter the system in maximally inconspicuous places: then they spread to new paradigms and new contexts. It doesn’t look promising for the Matrix Language Turnover Hypothesis, however, unless that hypothesis is to be taken merely as a label for the well-known historical process that also occurs in internally-motivated change, perhaps even more frequently than in contact-induced change. It is irrelevant here whether Mednyj Aleut actually did go through a sequence of changes that began with something like the current Evenki system. What is significant is that a matrix-language turnover, in Myers-Scotton’s terms, seems unlikely to look like current Evenki. If speakers are in fact switching from one matrix language to the other, it seems counterproductive to start with a relatively infrequent paradigm. More importantly, a process that acts essentially the same way in both internally-motivated change and externally-motivated change is of dubious relevance for a theory that focuses solely on externally-motivated change.

King argues for ‘a theory of grammar contact by which contact-induced grammatical change is mediated by the lexicon’ (2002). For her, the existence of structural interference that is not mediated by lexical borrowing is in doubt; if my interpretation of her position is correct, then this would be (in her view) an absolute linguistic constraint on possible structural interference. She dismisses the numerous counterexamples in the literature, saying that ‘lack of detailed evidence and/or analysis make the conclusions drawn less than convincing’ (2002). It’s hard to argue against a position that dismisses all the available counterevidence so confidently. Of course King, like other sociolinguists who have conducted sophisticated and detailed studies of ongoing structural change in current contact situations, has a right to be proud of the quality of her data and the comprehensiveness of the analyses permitted by such a rich database. Nothing comparable can be expected from any past contact situation. Nevertheless, historical linguistics have developed sophisticated methods of their own over the past century and a half for the analysis of linguistic changes, and much of the evidence provided by these methods is very solid indeed. Rejecting all the relevant evidence in this domain so casually surely isn’t a good idea: any successful overall theory of contact-induced change must ultimately account both for the results of past changes and for the evidence provided by modern sociolinguistic analyses. Rejecting one or the other body of evidence will leave us all the poorer.

I will return to some of these issues in §4 below. Before addressing the substantive issues, however, some background is required.

3. A historical linguist’s starting point

Discussions of contact-induced change don’t usually include any specific comparison of change processes in contact situations, on the one hand, with processes of internally-motivated change on the other. But in the context of this paper it is obvious that such a comparison is needed, because in
internally-motivated change, linguistic causal factors are indeed likely to be most prominent (though there are also, arguably, internally-motivated changes in which social factors predominate). And although it is quite easy to find identical changes that have internal motivations in one language but external motivations in another, some externally-motivated changes belong to types that are rare or nonexistent, at least over the short term, in internally-motivated linguistic change. In this section, therefore, I will first review some basics of language change in general, focusing on causes of change, and then present briefly the basics of my view of contact-induced change.

3.1. Causes of linguistic change

In historical linguistics, the three traditional causes of language change are drift (pattern pressures/structural imbalances and universal structural tendencies), dialect borrowing, and foreign interference. Of course no sharp boundary separates the second cause from the third: these should be viewed as belonging to a single continuum. The position of a particular contact situation on the continuum depends primarily on how similar the speech varieties are to each other, both lexically and structurally. Closely related systems will be investigated by dialectological methods, while languages that are very different both typologically and lexically will be studied as languages in contact. There will inevitably be borderline cases, where both dialectological and language-contact methods will prove useful (see Thomason 2003 for further discussion of these two causes).

A fourth, much less traditional, cause of linguistic change is deliberate decision by speech communities (Thomason 1997; 2001:149-152; 2007). The most familiar of deliberate changes are lexical (e.g. teenage slang), but more and more examples have recently come to light in which communities make apparently conscious changes in their language’s structure. The existence of deliberate change is ultimately the main reason why no one-size-fits-all overall theory of language change is likely to succeed; its importance for a comprehensive theory of language change will be discussed in §4 below.

3.2. A definition of contact-induced change

The second preliminary that is required is a definition of contact-induced change. Mine is somewhat idiosyncratic; here it is (and see also Thomason 2001:62-63):

Contact is a source of linguistic change if it is less likely that a given change would have occurred outside a specific contact situation.

Crucially, this definition does not entail that contact is the only source of a given change; multiple causation is always a real possibility, as it is in strictly internally-motivated linguistic changes. The definition of course includes the most obvious category, linguistic interference—the transfer of linguistic features (with or without actual morpheme transfer) from one language to another.

The definition includes at least three other phenomena as well. First, some changes in some dying languages do not make the dying language more similar to the dominant language, the one to which its speakers are shifting; such changes do not involve interference features, but by my definition they are contact-induced changes, because they would be less likely to occur if intense contact of a particular kind weren’t leading to language death. Second, my definition also includes changes that occur in a late stage of a chain-reaction process that was triggered in the first instance by a borrowing. A rather common example is the borrowing of a subordinate conjunction into a language in which native subordination constructions are non-finite: the borrowing of a conjunction often leads eventually to the development of finite subordinate clauses, but the finite subordinate clauses are not themselves borrowed. Nevertheless, because the borrowing of the conjunction set off a series of changes that resulted in a significant change in the morphosyntactic expression of subordination, even the late-stage changes are ultimately contact-induced. Third, my definition of
contact-induced change includes deliberate changes introduced by a speech community in order to make their speech more different from the speech of their neighbors, who speak dialects of the same language or a very closely-related language.

3.3. A major sociolinguistic distinction.

My third background position is the claim that there is a robust sociolinguistic predictor of contact-induced change (first introduced in Thomason & Kaufman 1988 and modified slightly in Thomason 2001 and elsewhere). As we will see in §4, this distinction is relevant to the main argument in this paper.

The vast majority of firmly established contact-induced changes fall into one of two categories. Changes under conditions of full bilingualism—that is, changes in which imperfect learning plays no role—begin with non-basic vocabulary, and only later (if at all) include structural features and perhaps also basic lexical items. Thomason & Kaufman (1988) used the term borrowing in a narrow sense to refer to this type of interference; typically, though by no means always, the agents of borrowing in this sense introduce features from a second language into their first language. The notion of dominant vs. subordinate language is often not applicable to changes in this category.

In sharp contrast, where imperfect learning does play a role in the process, the predominant interference features are phonological and syntactic, although lexicon and morphology may also be transferred from one language to the other. (The relatively few known instances of interference via superstrate shift, most notably the Norman rulers’ shift from French to English in England, comprise a partial exception to this generalization, as they often include a great many loanwords.) This is shift-induced interference, so called because it usually occurs when an entire speech community shifts to another group’s language. The process is complex. First, the shifting speakers may fail to learn certain features of the target language (TL)—often features that are universally marked and therefore relatively difficult to learn—that are lacking in their original language; and second, they may also carryover features from their first language (L1) into their version of the TL. These two kinds of interference features combine to form the shifting group’s version of the TL, the TL\(_2\). In addition, however, if the shifting group is integrated into the TL speech community, original TL speakers may borrow a subset of the TL\(_2\) interference features, thus forming TL\(_3\), a melded version of the TL. Note that in this last situation Van Coetsem’s (1988) concept of ‘imposition’ from a subordinate language to a dominant language breaks down, because the agents of the last phase of the process are original speakers of the dominant variety; and worse, they may not even speak the subordinate variety fluently, because the process of transfer (borrowing, in the narrow sense of Thomason & Kaufman) may come about through passive familiarity (Thomason 2001:139-142).

One point that needs to be emphasized here is that inflectional morphology tends to lag behind other linguistic subsystems in both borrowing and shift-induced interference (Thomason & Kaufman 1988:52). Derivational morphology tends to undergo contact-induced change almost as readily as syntax does, but both the transfer of inflectional categories and the transfer of inflectional morphemes themselves are relatively rare. The reasons presumably lie in the nature of inflection: as the most tightly-structured closed system in any language that has a complex inflectional morphology, inflection coheres in ways that make outside interference harder (harder to learn). There are, to be sure, solid examples of inflectional interference, but except in the case of dialect borrowing, they are rarer than examples of contact-induced change elsewhere.

One other point that requires extra emphasis is that neither borrowing nor shift-induced interference is inevitable in contact situations. Some speech communities (for instance many Native American languages of the United States Pacific Northwest) don’t borrow from neighboring languages even under conditions of intense contact and strong assimilatory pressure; and many
cases of shift-induced interference leave little or no trace of the shifting group’s TL2 in the dominant language (this is true of most immigrant communities in the U.S.).

### 3.4. Possibility vs. probability of linguistic change

The fourth preliminary is a distinction between possible and probable changes. I claim that the question of **possibility** is settled as soon as an innovation appears anywhere, just once, in a single person’s speech, regardless of whether the innovator is an adult or a child. The subsequent fate of that innovation is a matter of linguistic and social **probabilities**. One implication of this claim is that relevant examples include not only completed linguistic changes but also such ephemera as speech errors and the joking introduction of foreign elements into one’s speech. Moreover, because calculating the social and linguistic probabilities is so difficult—indeed impossible, now and for the foreseeable future—we can reasonably talk only about **necessary** conditions for change, not about **sufficient** conditions for change. In other words, we cannot predict when or whether change will occur. This is as true of contact-induced change as it is of internally-motivated change: even the most intense contact situations don’t always lead to significant contact-induced changes, just as the most common internally-motivated changes often fail to occur.

### 3.5. How to prove that contact-induced change has occurred

My fifth and final preliminary is a list of the steps that must be taken before a claim of contact-induced change can be considered to be firmly established (following Thomason 2001:91-95). I will not return to this issue in §4, but I include it here because it is necessary to underscore the point that the changes mentioned in §4 are not based on guesswork. They are established by rigorous methods familiar to historical linguists.

In this catalogue of requisites, I will concentrate on structural rather than lexical interference, because establishing that words have been borrowed is much less likely to present major difficulties for a historical linguist; the debates in the literature almost always center around structural interference. First of all, in order to establish that contact-induced change has occurred, it is necessary to look at the languages as wholes: if structural interference of any kind has occurred, it is highly unlikely to be isolated in the system. The best strategy is to look for a series of independent, unrelated structural changes in the proposed receiving language.

Second, identify a source language (or, in a case of language shift, one or more languages related to the proposed source language). It must be possible to show that the proposed source language is or was in contact with the proposed receiving language, and that the contact situation was intense enough to make structural interference a reasonable prospect. Third, identify shared (or partially shared) structural features in the proposed source and receiving languages. Fourth, prove that the proposed interference features did not exist in the receiving language before it came into contact with the source language: that is, prove that the proposed receiving language has changed with respect to these features. This can be done either by inspection of documented earlier stages of the receiving language or, if that is impossible (as it usually is), by examining languages related to the receiving language in an effort to show that the relevant features cannot be reconstructed for their common parent language. (Considerations of markedness enter in here too, because of the possibility of independent innovation, but I will not go into those details here.) Fifth and finally, prove that the proposed interference features were present in the source language before it came into contact with the receiving language, by the same means as with the fourth requisite.

It is also necessary to keep in mind that, as noted above, establishing an external cause does not exclude an internal cause. Multiple causation is common, and a complete explanation for a given change in a contact situation must take potential internal as well as external motivations into account. The reverse holds too, of course: if there is any evidence of significant language contact,
4. Why there are no absolute linguistic constraints on contact-induced change

In this section I will present an argument for the lack of absolute linguistic constraints on contact-induced change. The theoretical part of the argument has to do, first of all, with the question of just what a linguistic change is. Is it the innovation—the new word or grammatical morpheme or sound or rule or meaning—that is introduced by one or more individual speakers? Or is it the spread of the innovation through the speech community, until it makes sense to say that a change has occurred in the speech community’s language? The answer is closely tied to another question: are the primary agents of language change children or adults? In this paper I will consider only the issue of what a linguistic change is (but see Thomason In press a for discussion of the agent-of-change issue).

For over forty years, generativists have placed the locus of language change in children’s acquisition of their first language (e.g. Halle 1962, Lightfoot 1979, 2002); they claim that children construct their grammars on the basis of input from adults and older children, and that their innovations are included in the internalized grammar and represent language change from one generation to the next. A popular opposing view, espoused by sociohistorical linguists such as Lesley Milroy and James Milroy (p.c. 2003), is that language change is not the innovation, but rather the spread of the innovation through the speech community. Those who hold this view argue that very young children are never the primary agents of change, because they do not engage in the kinds of social interaction that spread an innovation through a community.

I believe that both of these positions are too extreme, because every linguistic change process includes both the original innovation and its spread through the community. It seems quite arbitrary, from a theoretical viewpoint, to focus exclusively on either the one or the other, although methodological considerations will dictate a practical focus (by a given set of scholars) on innovations or the spread of innovations. There can obviously be no language change without an innovation; equally obviously, if an innovation remains confined to the speech of one speaker, or a scattering of isolated speakers, no one would call that phenomenon a change in a language (provided that one defines ‘a language’ as having some kind of reality in a speech community, not just in a single individual’s competence).

Now, the original innovation is unobservable—not in principle, but in practice. You might be the first person to produce or to hear an innovation that later becomes a change in an entire language or dialect, but you would have no way of knowing that you’ve produced or heard the very first utterance of the innovation. This generalization has two types of exception, however. One of them occurs in the kinds of deliberate change that will be the topic of the second part of my argument for the lack of absolute linguistic constraints (see below); the other obtains in rather trivial cases, where someone invents a new word and sees it spread through an entire speech community. In all other instances, the innovation that triggers a linguistic change is unobservable.

A linguistic innovation may or may not have a strictly linguistic motivation: it may arise entirely through processes inside the head of a monolingual or bilingual person, in which case it arguably has only a linguistic trigger, or it may arise from something that a monolingual or bilingual person hears from another speaker (e.g. a misperception of a particular speech sound or word), in which case it results in part from social interaction. Even if the innovation is entirely linguistic in nature, however, the spread of any innovation through a speech community must certainly be social at least in part, because it is governed by such things as social networks, prestige, and various demographic factors. The spread of a structural innovation is also very likely to depend in part on linguistic factors; the likelihood of its successful spread is enhanced if it is less marked (easier to learn), if it
fits well typologically with the language’s pre-existing structural patterns, and if it fits neatly into the language’s internal tendencies of drift having to do with inherited structural imbalances. Nevertheless, the spread of an innovation cannot be due entirely to linguistic factors.

If this is correct, then it follows that the spread of an innovation can’t be relevant to the question of what linguistic changes are possible. Any search for absolute linguistic constraints on language change, whether internally-motivated change or contact-induced change, must focus on the process of innovation by one or more individuals. Even the most powerful linguistic tendencies that favor or disfavor the spread of a given innovation cannot be absolute constraints; if they were, the innovation (ephemeral or permanent) could not have occurred in the first place.

This brings me to the second part of my argument against the existence of absolute linguistic constraints. The problem with proposed absolute linguistic constraints on language change is that there is a growing body of evidence suggesting that speakers can, and sometimes do, introduce deliberate changes into their own speech and often into a language as a whole (Thomason 2007). In fact, I would now claim that speakers can change any part of their language(s) that they can become aware of—that is, including not only the usual sorts of lexical innovations that one sees in (for instance) teen-age slang, but also structural innovations. Ever since I started discussing such changes in public eight years ago, audience members and other colleagues have been offering new examples. Instances of deliberate structural change are not flooding into the literature, but they constitute a steady trickle, and they are accumulating.

In a way this should come as no surprise. The literature of linguistic anthropology and (to a somewhat lesser extent) sociolinguistics is full of analyses of speakers’ manipulation of language in the creation of social identities (see e.g. Gal 1979, Le Page & Tabouret-Keller 1985, Irvine & Gal 2000). Many of the examples in this literature come from bilingual and multilingual contexts, where speakers deploy their multiple repertoires to construct a bilingual, bicultural identity separate from either (or any) of the source cultures. In other words, it is already well known (in some circles) that speakers can do all sorts of things to their languages. In principle, of course, they could do this by manipulating lexicon alone; in practice, however, the linguistic alterations are much more diverse.

Examples of deliberate structural changes can be found in all grammatical subsystems (for discussion of these and other examples, see Thomason 2007). They include such changes as the introduction of borrowed phones and phonemes, e.g. Ma’a/Mbugu bilinguals’ insertion of the voiceless lateral fricative /l/, of Cushitic origin, into Bantu words (Mous 1994:199); inflectional changes, e.g. morphological adjustments in sixteenth-century Portuguese that distanced the language from Spanish (Wright 1998, citing Yakov Malkiel), the lexically-determined reversal of all masculine and feminine gender markers in the Usai dialect of Buin (Laycock 1982, cited by Kulick 1992), and the introduction of inflectional superlatives into Estonian by the language reformer Johannes Aavik (Oksaar 1972:491); and syntactic changes, e.g. certain (possibly ephemeral) adjustments in object marking in Nisg̱a’a under the influence of English (Tarpent 1987:157) and the (ephemeral) restructuring of ditransitive constructions in Montana Salish, also under the influence of English.

The most dramatic examples of deliberate manipulation of bilingual repertoires, however, appear in the structures of those bilingual mixed languages that must have arisen abruptly, not through any kind of gradual, incremental language change. Because they arose abruptly, and because they serve as symbols of in-group identity, it is difficult to imagine a developmental route that did not include an element of deliberation; certainly the identity-creation aspect of their development fits with concepts and instances of language manipulation found in the literature of linguistic anthropology. The best-known examples are Michif, with its French noun phrases combined with Cree verb phrases—with mostly independent phonology, morphology, and syntax in each component (Bakker & Papen 1997); Mednyj Aleut, with Russian finite verb inflection grafted onto Aleut verbs; and...
Media Lengua, with Spanish vocabulary but Quechua grammar (Muysken 1997). Other instances of dramatic bilingual language mixing, some or perhaps even most of them ephemeral, have been turning up with increasing frequency in recent years. In Australia, Carmel O’Shannessy has documented the ongoing creation by adolescents of a mixed language, Light Warlpiri.

A caveat: it is still the case that the great majority of linguistic changes seem to be unconscious, not deliberate. The best reason for believing this is that the standard methods of historical linguistics, in particular the Comparative Method, apply unproblematically and uncontroversially to most aspects of most languages. But my argument here does not require that deliberate changes be frequent, much less predominant, in language history. It requires only that deliberate changes occur, and that they include some changes that can be seen as violations of proposed absolute linguistic constraints on contact-induced change. I submit that changes like the wholesale incorporation of Russian inflection in Mednyj Aleut finite verbs and of French noun phrases in Cree sentence frames, and less spectacular changes like the innovative synthetic superlative construction of Estonian and Ma’a/Mbugu speakers’ introduction of /l/ into their Bantu language, ought to be excellent candidates for impossible changes in the ordinary course of language change.

Some scholars have attempted to define such developments as unnatural and thus to exclude them from theories that pertain to ‘natural’ linguistic changes (see e.g. Greenberg 1999:632). That is, they wish to distinguish between ‘natural’ linguistic change (the kind usually studied by historical linguists) and ‘unnatural’ linguistic change (real or hypothetical examples of changes that don’t fit into the usual picture of language change). If it were possible to make this distinction in a principled way, then perhaps one could still hold onto the belief that certain types of changes are impossible as ‘natural’ changes. But no principled way of making this distinction is available—first because the linguistic behaviors that produce deliberate changes closely resemble linguistic behaviors that do not produce exotic results (Thomason 1995), and second because any usable definition would be circular: ‘unnatural’ changes would be all and only those that violate proposed absolute linguistic constraints on change, and the evidence for the existence of the constraints would be that such changes never occur in ‘natural’ change.

Unless and until proposals are made that can’t be counterexemplified by the actual and potential evidence from processes of deliberate change, I will remain skeptical about the prospects for successful absolute constraints. I should add, finally, that the arguments and examples discussed in this section do not in any way preclude the influence of linguistic tendencies, both in the innovation of deliberate changes and in the propagation of deliberate changes through a speech community. I am talking here about possibilities, not about probabilities.

5. Social factors trump linguistic factors as predictors of contact-induced change

In order to test the claim that social factors are ultimately more influential than linguistic factors in guiding contact-induced change, we need to consider situations in which linguistic factors might be expected to drive change in one direction and social factors might be expected to drive change in another. Perhaps the most obvious example is unfortunately completely trivial: without language contact, there can be no contact-induced linguistic change. So, for instance, even though loanwords are by far the most common linguistic outcome of borrowing situations (in the narrow sense of ‘borrowing’ discussed above in §3.3), we are hardly likely to find any loanwords from (say) Montana Salish in Zulu, because the two languages have never (to the best of anyone’s knowledge) been in contact. Admittedly, this isn’t a very interesting test; but it does show that the social factor—lack of contact—wins out over the linguistic probabilities, namely, lexical borrowing.

More interesting cases can be cited in response to various constraints mentioned in the language-contact literature. One of the best known is found in Weinreich’s pioneering and classic book: ‘in the interference of two grammatical patterns it is ordinarily the one which uses relatively free and
invariant morphemes in its paradigms...which serves as the model for imitation’ (1953:41). But this can’t be right: the social relations between the two speech communities, not the structures of their languages, determine the direction and even the extent of interference. One of Weinreich’s examples is the replacement of possessive constructions comprising a noun plus a pronominal possessive suffix by analytic possessives in Estonian (under the influence of German). But this was a case in which German was a prestige language, and Estonian therefore borrowed structure from German; and it happened that the German possessive construction was analytic while the original Estonian possessive construction employed pronominal possessive suffixes. As we saw in §4, however, the direction of influence was the same even when the Estonian construction was more analytic than the functionally equivalent structure in German: the change in superlative constructions introduced deliberately by Aavik falls into this category. Or, to take an example from possessive constructions, the addition of third-person possessive suffixes to Greek nouns in some Asia Minor Greek dialects under Turkish influence (Dawkins 1916:201) reduced the analyticity of Asia Minor Greek; this happened because Greek was influenced by Turkish for social reasons—Turkish was clearly the dominant language, and most Turkish/Greek bilinguals were Greeks, not Turks. So Weinreich’s prediction does not hold. It is no doubt true that the linguistic tendency he points to here is real and strong, so that under appropriate social circumstances we will find more changes toward analyticity than away from it. But the linguistic tendencies do not decide which language will be the donor and which the receiver of interference features.

In the same way, the robust correlation between presence or absence of imperfect learning, on the one hand, and types of interference features on the other (see §3.3 above) is a social fact with linguistic consequences. The linguistic patterns—lexical borrowing first and most and structural interference later and less in borrowing situations, vs. phonology and syntax first and most in cases of shift-induced interference—are conditioned by the sharp distinction between the two social situations, not vice versa.

In the most extreme situations, most notably bilingual mixed languages that arose at least in part through deliberate change, we find the most striking evidence of the dominance of social factors. Again, it is not that linguistic factors are irrelevant in such cases. In Michif, for instance, it is surely no accident that the noun-phrase/verb-phrase split resembles a common code-switching pattern. But the fossilization and coalescence of such a pattern into a single language, learned by children as a first language, cannot be attributed to linguistic processes alone. It was the social situation, the relations between the bilingual creators of Michif and their social context, that caused Michif to crystallize into a bilingual mixed language.

Mednyj Aleut too was created by fluent bilinguals—that much is evident from the fact that distortion in the Russian and Aleut components of the language is minimal—for some purpose other than ordinary communicative functions. Like the creators of Michif, the people who developed Mednyj Aleut had no communicative need for a new language; they already had two languages at their disposal. Their need, or desire, was for a language of their own (cf. the title of Bakker 1997: ‘A language of our own’). Both Michif and Mednyj Aleut, like Media Lengua, Ma’a, and all the other bilingual mixed languages that have been reported in the literature, are in-group languages. They are not used for communicating with outsiders; indeed, they couldn’t be used for such purposes, because none of them is intelligible to monolingual speakers of either component language. (It is debatable whether they are intelligible to people who are fluent in both component languages but who have not learned the bilingual mixed language. It seems unlikely.) In other words, they were created for social reasons.

Myers-Scotton claims that ‘the grammatical organization [of Mednyj Aleut]...can be explained entirely in terms of structural mechanisms’ (2002:265), and she also analyzes Michif in terms of her Matrix Language Turnover Hypothesis (2002:254-258). But since the structures of Michif and Mednyj Aleut differ sharply, it is difficult to see how the same hypothesis can explain their different
grammatical organizations entirely. Other bilingual mixed languages, such as Media Lengua with its lexicon/grammar split, differ yet again. On the hypothesis that social factors are primary and linguistic factors are secondary, these differences present no such problems. It is certainly true that we lose predictive power if we abandon the effort to discover a deterministic linguistic framework; but if the only available deterministic linguistic frameworks fail to predict the outcomes that actually occur, then we are better off without them. If I am right in my belief that the dominant force in the emergence of a bilingual mixed language is its creators’ creativity—their ability to manipulate their multiple linguistic repertoires for social purposes—then there is no reason to expect uniform linguistic results, even when the social circumstances are in many ways very similar.

As noted above (§2), no historical linguist believes that linguistic change can be predicted. In internally-motivated change, general tendencies can certainly be identified, but specific changes can only be safely “predicted” when they are already under way. The situation is worse in contact-induced change, if only because the availability of at least two distinct and independent linguistic systems greatly expands the immediate possibilities for change. We can make rough predictions about types of change to be expected under different social conditions (for instance with the distinction between borrowing and shift-induced interference), and we can often predict the direction of change; but we can’t even be sure that contact-induced change will occur in any given contact situation, much less predict what form it will take if it does occur.

Some authors have called this position pessimistic. Their underlying assumption, apparently, is that it is ‘giving up’ to acknowledge that contact-induced change cannot be wrestled into a single tidy predictive theory, and that there are no absolute linguistic constraints on what changes can occur as a result of language contact. To me, the lack of neatness and predictability is one of the most interesting things about language change. The search for causes of change, and for areas where probabilistic predictions can be made with due caution, is endlessly fascinating.

References


