

How to Establish Substratum Interference

著者(英)	Sarah G. Thomason
journal or publication title	Senri Ethnological Studies
volume	75
page range	319-328
year	2009-09-18
URL	http://doi.org/10.15021/00002571

How to Establish Substratum Interference

Sarah G. Thomason

University of Michigan

1. Introduction
2. Requisites for Establishing a History of Contact-Induced Change
3. Substrate Interference in Tibeto-Burman and/or its Neighbors?
4. Conclusion

1. Introduction

The main purpose of this paper is to review methodological criteria for establishing the effects of a substrate language on the structure of a target language, and to offer some comments about the applicability of these criteria to the study of contacts between Tibeto-Burman languages and their neighbors. The main conclusion is that, while the possibilities for substrate influence in this area are certainly promising, relatively few concrete proposals can be considered established in the current state of research on the topic. The reason is that, in most cases, not all the requisites for making a strong case for substrate influence can be met. The historical picture of the history of Tibeto-Burman languages, including both a subgrouping model and a substantial amount of reconstruction, is still at a preliminary stage, and the same is true of most of the neighboring language families. Without this information, it is difficult to show that any of the languages, Tibeto-Burman or others, have changed.

The structure of the paper is as follows. This section continues with a definition of contact-induced change and a characterization of substrate interference. Section 2 surveys requisites for establishing that substrate interference has occurred, and section 3 discusses a few examples from the Tibeto-Burman area. Section 4 is a brief conclusion.

First, a definition: contact is a cause of linguistic change if it is less likely that a given change would have occurred outside a specific contact situation (Thomason 2001: 62). Crucially, this definition requires identification of a specific contact situation in which a proposed contact-induced change occurred; it is not possible to argue successfully for contact-induced change by appealing to an unidentified long-vanished substrate language.

Second, the complex nature of substrate interference needs to be recognized. The term ‘substrate interference’ is somewhat problematic, because it implies not just an underlying historical stratum—a language whose speakers died out and/or shifted to another language—

but also a socioeconomically subordinate speech community. The first implication is valid, but the second is not. As we know from a few documented historical cases, such as the shift of Norman French speakers to the language of their conquered English territory, it isn't safe to assume that every speech community that shifted to another language was subordinate. True, the shift from Norman French to English is called a 'superstrate shift'; but my point here is that we can't always know what the social, political, and economic relations between an ancient shifting group and the target-language speech community were. To avoid possible confusion, I will henceforth use the clumsy term 'shift-induced interference'—that is, interference that comes about as part of a process of language shift—instead of 'substrate interference'.

The most important fact about shift-induced interference in this context is that it involves imperfect learning of a target language by a group of shifting speakers. The processes are essentially the same regardless of whether actual shift occurs, and very similar processes also characterize second-language acquisition by individuals; but only cases of language shift are likely to cause difficulties for the historical interpretation of linguistic changes. The reason the fact of shift (or second-language learning by a group of people) is so important is that it carries implications about the types of expected interference features: when contact-induced change occurs in shift situations, the dominant interference features are phonological and syntactic; there may or may not be substantial lexical interference as well. In sharp contrast, contact-induced change that does not involve imperfect learning almost always starts with extensive lexical borrowing and only later includes structural features, primarily in the phonology and syntax. Morphology, especially inflectional morphology, lags behind phonological and syntactic interference in both of these basic types of contact-induced change. (The distinction between shift-induced interference and borrowing in the narrow sense of interference without imperfect learning is discussed in more detail in Thomason and Kaufman 1988 and Thomason 2001: 66ff.)

In some current contact situations the contrast is illustrated by comparing a bilingual group's L_1 , which has borrowed from their second language, with their L_2 , which has undergone shift-induced interference (Thomason 2001: 76). In Peru, for instance, certain Quechua-Spanish bilinguals shows heavier lexical interference and less structural interference in their Quechua (their L_1), but heavier structural interference and fewer loanwords in their Spanish (their L_2) (Godenzzi 1988). Similarly, in Qinhai, China (formerly northern Tibet), bilingual speakers of the Turkic language Salar and Chinese have much lexical borrowing but only moderate morphosyntactic and phonological borrowing from Chinese in Salar, their L_1 , while they have weak lexical interference, strong morphosyntactic interference, and moderate phonological shift-induced interference in their Chinese, their L_2 (Dwyer 1995).

At least two different processes lead to shift-induced interference in a target language. The first is the failure of second-language (L_2) learners to learn certain features of the target language (the TL). This is especially likely to happen with marked features—those features that are harder to learn, as measured cross-linguistically by typological distribution and age of learning for first-language acquirers. An example might be efforts by young native English-speaking tribal members to learn the glottalized and pharyngeal consonants of Montana Salish, the heritage language of the Bitterroot Salish and Pend d'Oreille tribes. Glottalized stops

and affricates in many younger learners' speech are replaced by plain stops, e.g. [t] instead of [tʰ], and the pharyngeal consonants are simply omitted, occasionally with lengthening of a preceding vowel. These learners' L₁, English, lacks both glottalized and pharyngeal consonants; the former are (relatively) hard to pronounce and the latter are hard to hear and also to pronounce, so the learners often fail to learn and use them.

It is worth noting here that the common claim that shift-induced structural interference inevitably simplifies the TL grammar is demonstrably false. Only the first process described here, failure to learn TL structures, is at all likely to simplify TL structure, even locally; and it doesn't always do so. To be sure, one may view the merger of glottalized and nonglottalized consonants in Montana Salish as a simplification, and at first glance the loss of pharyngeal consonants also looks like a clearly simplificatory change. But since the contextual phonological effects of the pharyngeals remain—notably lowering and backing of neighboring vowels, together with compensatory lengthening of vowels that were once followed by pharyngeals—the loss of the pharyngeals complicates the vowel system, and arguably the phonology as a whole.

The second process is carryover of features from the learners' L₁ into their version of the TL. So, for instance, learners of Montana Salish are likely to prefer SVO word order, as in *T Pyel wíčis Čuspín* 'Pete saw Josephine', literally 'OBLIQUE Pete [the name comes from French *Pierre*] see.s/he.him/her/it Josephine'. This word order is perfectly grammatical in Montana Salish, but it is also rare, occurring only when the agent, in this case Pete, is emphasized. The normal, unmarked word order in Salish is VOS, so this Salish sentence would normally be *Wíčis Čuspín t Pyel*. The greater frequency of SVO sentences must be due to the influence of learners' English syntax. The two processes together—failure to learn certain TL features and carryover of L₁ features into the learners' L₂—comprise a shifting group's version of the TL, which we may call TL₂.

A third process may also be involved, depending on whether or not the shifting group merges with the original TL speakers (the speakers of TL₁) into a single speech community. In that case, speakers of TL₁ are likely to borrow a subset of features from TL₂, forming a unified TL₃ that eventually becomes the language of the entire speech community. A standard example is Irish English in Ireland, which is spoken both by descendants of Irish Gaelic speakers and by descendants of English invaders and conquerors. By contrast, the vast majority of English speakers in India are ethnically Indian, not descendants of former British rulers; as a result, the speech variety known as Indian English is in origin a TL₂, not a TL₃. (This picture of Irish English and Indian English is of course oversimplified in many respects—for one thing, no speech community is as homogeneous as this account suggests. But the basic contrast is real.) The existence of TL₃'s in some speech communities underscores the complexity of shift-induced interference as a set of processes, and there is often still another layer of complexity: in some contact situations both shift-induced interference and borrowing may occur more or less simultaneously, with speakers of TL₁ borrowing words and perhaps even a few structural features from the shifting group's original language even as shift-induced interference is changing the TL.

With this general background on shift-induced interference, we are ready to consider the criteria for establishing it. The criteria in §2 are equally valid for borrowing in my narrow

sense and for shift-induced interference, but for reasons that will become clear, it is much easier to satisfy the requirements for borrowing situations than for shift situations.

2. Requisites for Establishing a History of Contact-Induced Change

Four methodological requisites must be satisfied before we can claim to have established interference from one language—let's call it language A—in another language, B. (The discussion in this section is based in part on Thomason 2001: 91-95.) First, we must prove the existence of contact between A and B, and the contact situation must have been intense enough to make the transfer of structural features possible. Casual, ephemeral contacts between languages do not promote structural transfer. The existence of contact is easily demonstrated if there are loanwords from A in B, and if there are many loanwords, then it's reasonable to assume that structure could in principle also have diffused from A into B. In borrowing situations, therefore, establishing the fact of earlier contact is easy, because loanwords are the first and most numerous interference features. But if, in a long-past contact situation, A speakers shifted to B, then there might be few or no loanwords; and if ALL A speakers shifted to B, then it might well be impossible to identify any source language for proposed interference features in B, the proposed receiving language.

This is a, or perhaps the, major problem with recurrent hypotheses of Semitic and/or Berber substrate influence on Insular Celtic languages in the British Isles. There are several intriguing structural parallels between the two groups of languages, and the features in question set Insular Celtic apart from other Indo-European languages; moreover, ancient contacts could have occurred between the two groups, because seafaring Semitic or Berber peoples could have reached the British Isles. But the contacts remain a matter of conjecture—there is no solid evidence—so that hypotheses of shift-induced interference in Insular Celtic remain unconvincing.

The second requisite is to identify diverse shared features in A and B. These can be lexical, phonological, morphological, syntactic, semantic, and/or discourse features; but a convincing case for structural interference must include shared features from at least two different structural subsystems. The reason is that structural interference is never completely isolated: if contact is intense enough to make structural diffusion possible, that diffusion will not be confined to a single interference feature. It is conceivable that a very ancient shift situation left several traces in the form of structural interference, and that the receiving language (B) has lost all but one of those features in the intervening centuries; but in such a case, especially with minimal lexical transfer, the evidence for contact, much less contact-induced change, is all too likely to be insufficient to permit a convincing case to be made. It is therefore necessary to consider the source and receiving languages as wholes, rather than focusing on a single feature at a time, in a search for contact-induced structural changes.

Strong cases for shift-induced interference display several to numerous shared features. A prime example is the Ethiopian linguistic area, in which many speakers of indigenous Cushitic languages shifted to the Semitic languages spoken by relative newcomers to the area. In this case there is clear evidence of the contact and its nature—both Cushitic and Semitic languages are still in contact in Ethiopia, and large-scale shift from Cushitic is well

established—and the interference features in Ethiopic Semitic include many loanwords as well as innovative structural features, e.g. labialized dorsal consonants such as /k^w/ and /x^w/, a new future tense, new word-order patterns, and novel means of expressing certain inflectional categories. The identification of shared structural features is not in itself more difficult for shift situations than for borrowing situations, so this requisite tends to be as easily satisfied for the one type of situation as for the other.

Third, we must be able to prove that the shared features—the proposed interference features—were present in language A before A came into contact with B. That is, we must prove that the features are not innovations in A (maybe as a result of interference from B). Obviously, A cannot be the source of the shared features in B unless the features were present in A during the relevant time period. Demonstrating that the features are old in A will only be possible if the history of A is well understood, at least with respect to this set of features. If, for example, A and all its known relatives have all the features, and if only A (unlike its relatives) was in contact with B and/or its close relatives, then the features are likely to be reconstructible for pre-A. Cushitic languages of Ethiopia arguably have all the features, sometimes in a modified form, that have been proposed as shift-induced interference features in Ethiopic Semitic, so the whole set of features is demonstrably old in Cushitic.

Alternatively, A may be one of the rare cases of a language with a well-documented history stretching back hundreds of years or even one or two millennia; in that case, provided that the contact with B is more recent than the earliest attested variety of A, this requisite will be satisfied if the earliest variety of A had the features in question. There is no doubt, for instance, about the age of the proposed Indic interference features in Indian English, because Indic languages have been attested for over two millennia, and the relevant features are present in the oldest recorded Indic material. (In this case the fact that the features are very widely shared among Indic languages would also provide solid evidence for interference in Indian English.)

But circumstances are not always so favorable, especially in more ancient contact situations. The Pacific Northwest region of the United States and Canada (primarily Washington, Oregon, and neighboring parts of British Columbia) is a famous linguistic area, with numerous area-wide structural features shared by the three core language families, Salishan, Wakashan, and Chimakuan. Unfortunately, however, the most widespread areal features must be reconstructed for all three proto-languages. It is therefore impossible to prove that ANY diffusion has caused these features to spread in the area, though it is hardly likely to be coincidental that all three families have all these features. It is probable that the linguistic area dates from (at least) the time of the proto-languages, and that diffusion occurred at that early period; but another possibility is that all three families are genetically related and inherited the shared features. Proposals of genetic relatedness have certainly been made for these three families, but specialists are virtually unanimous in their belief that probative evidence for relatedness is completely lacking; and given the significant time depths of two of the three families, it may well be that any putative relationship is too distant to be within the reach of the Comparative Method, the only tried and true methodology we possess for tracing genetic relationships among languages.

Fourth, we must prove that the features shared by A and B were NOT present in B be-

fore it came into contact with A. That is, we must prove that B has changed. In the Ethiopic Semitic case, we know that Proto-Semitic did not have the proposed interference features, which means that Ethiopic Semitic has certainly changed. We also know that English did not have the set of proposed interference features from Indic when English speakers first arrived in numbers in India, so it is easy to prove that Indian English reflects change from earlier English. But we do not know whether Chipaya, a Bolivian language with no established relatives (Campbell 1997: 189), has changed structurally under the influence of Aymara, from which it has borrowed many words and a plural suffix. The existence of Chipaya-Aymara contact is therefore certain, and Chipaya also shares numerous typological features with Aymara; but because Chipaya is an isolate, we have no evidence that it has undergone any structural change at all.

If we can satisfy all four requisites listed in this section, we should take one more step before our historical interpretation of the contact situation is complete: we need to search in B for internal sources of the shared features. If there are plausible internal sources for all the features shared by A and B, then the case for interference may be weakened; but if the features are truly diverse, situated in different structural subsystems of B, then interference may be the best historical explanation for their presence in B even if there are plausible internal motivations for all the innovations. The reason is that, all other things being equal, a single unified external explanation for a set of innovations is better than a large number of separate internal explanations. Nevertheless, multiple causation of linguistic change is a common phenomenon, so that external and internal motivations may combine to produce a single change. A search for internal causes of innovations can therefore lead to better explanations even when the case for contact-induced change is rock-solid.

Before concluding this section, I will add a few caveats. The literature contains arguments of various kinds for and against shift-induced interference, and some of these arguments are invalid. First, the mere absence of a plausible internal cause of a linguistic change does not justify a claim that the change must be due to contact; historical linguists have no explanations for the vast majority of known linguistic changes, so it is quite ordinary, not extraordinary, if no motivation for a given change can be found. Second, the often-repeated claim that the absence of loanwords means the absence of contact-induced change is invalid, because shift-induced interference often occurs without significant amounts of lexical transfer. Third, the claim that a given innovation cannot have been due to contact because a similar or identical change has happened elsewhere through internal causation is invalid: some of the most common contact-induced changes are also common internally-motivated changes. And finally, arguments of the form 'Feature *x* must be due to contact because it is typologically rare' are also invalid; they merely shift the locus of the historical mystery from one place to another.

In sum, if you cannot satisfy the four requisites discussed in this section, then you cannot make your case for contact-induced change, including shift-induced interference. This of course does not mean that no interference happened; it does mean that if it happened, we can't prove it.

3. Substrate Interference in Tibeto-Burman and/or its Neighbors?

The specialist papers presented at the Symposium “Linguistic Substrata in the Tibeto-Burman Area” painted a fascinating picture of multiple intersecting contacts among Tibeto-Burman languages and between Tibeto-Burman languages and their non-Tibeto-Burman neighbors. Evidence of movements of populations, assimilation of various ethnic groups, and cultural pressures of various kinds indicated that in many parts of the area the stage was set for considerable contact-induced change, including structural changes.

This rosy picture offers possibilities that apparently cannot all be realized in full at the present stage of research, however. There is currently no generally agreed-upon subgrouping model for Tibeto-Burman as a whole, and even within subbranches of Tibeto-Burman there seem to be numerous indeterminacies and even controversies over subgrouping. Without a firmly-established subgrouping model, it is often impossible to tell whether structural features found in a given language are inherited or innovative; and unless and until that distinction can be made with confidence, requisites 3 and 4—proof that structural features shared by a genetically diverse group of languages are old in the proposed source language (A) and innovative in the proposed receiving language (B)—cannot be satisfied. The only exception would be features that are universal in Tibeto-Burman languages (and therefore reconstructible for Proto-Tibeto-Burman); in such a case, if the features are shared with neighboring non-Tibeto-Burman languages that can be shown to have innovated them, then a case for contact-induced innovation in the non-Tibeto-Burman languages can be made. But no one at the Symposium described such a situation. Therefore, although two of the four requisites were satisfied in several instances—the fact of contact was established, and diverse arrays of shared features were identified—a strong case for contact-induced change was not yet feasible. An added feature of much of the research reported on at the Symposium was that many of the relevant languages are poorly documented or, in a few cases, hardly documented at all. Historical investigation of structural features cannot proceed usefully without a good understanding of the structures of the relevant languages.

Not surprisingly, then, many symposium participants who were most closely concerned with the historical interpretation of shared features in groups of Tibeto-Burman and other languages of the region were focusing their attention on the documentation and analysis of undescribed languages, including Zhangzhung, and on working out the subgrouping of Tibeto-Burman languages and the reconstruction of Proto-Tibeto-Burman (e.g. the papers by Nagano, Hayashi, Evans, and Takahashi). Many of the broader questions of contact phenomena, including shift-induced interference, will be solved in the future, when the languages themselves are better understood.

Even at this necessarily preliminary stage of the historical research, however, a number of very interesting results have emerged. In discussing directional prefixes in languages of the West Sichuan Ethnic Corridor, for instance, Shirai showed that the category and its organization are widely shared by languages of the region—even by one language, Baima, which seems not to be closely related to most of the other languages—and that no such system is found in Tibeto-Burman languages outside the region. In this case, a contact explanation for the areal feature is appealing, and inheritance from Proto-Tibeto-Burman is unlikely.

Innovation in one part of the area and diffusion throughout the region is the most likely explanation for this shared feature, as Shirai concludes; but if contact was indeed responsible for the spread of the directional prefixes, it remains to be seen whether the process was one of borrowing or of shift-induced interference. In any case, a clearer historical picture, and identification of other widely-shared features characteristic of this region, will be necessary to establish this hypothesis firmly.

Another example was LaPolla's discussion of shift-induced interference from Mon on Burmese. The histories of the two peoples are intertwined in Southeast Asia; The Mon were alternately independent (and apparently culturally prestigious) and conquered, and they eventually shifted in large numbers to Burmese. LaPolla cites Bradley (1980), who argues that an array of Burmese features—among them sesquisyllabic structure and the loss of certain contrasts between series of fricatives and affricates—points to Mon substrate influence on Burmese.

A particular focus of the symposium was the nature and development of Zhangzhung, a language classified by Matisoff (2001) and others as belonging to the West Himalayish branch of Tibeto-Burman (though alternative classifications of Zhangzhung within Tibeto-Burman have also been proposed). The history of Zhangzhung is discussed as part of the general issue of substrata in Tibet—specifically, the question of whether the people who were in Tibet when the Tibetans arrived, and who subsequently vanished, influenced Tibetan linguistically as they shifted to the newcomers' language. There is no doubt that many language contacts existed at that early period; there is some question about just what languages were displaced or absorbed by Tibetan. The fact that the putative substrate languages no longer exist is not necessarily an insurmountable barrier to establishing influence from them on Tibetan, because it is entirely possible that modern Tibetan speakers have neighbors whose languages are related to the original languages of the region. One thing that seemed clear at the symposium, however, was that it is premature to make any strong claims about shift-induced interference in Tibetan territory.

4. Conclusion

This paper has reviewed methodological requisites for establishing a history of contact-induced language change, with a focus on the special problems that often arise in the case of old contact-induced changes that occurred as a result of language shift. Then a very brief survey of some proposals of borrowing and (especially) shift-induced interference in Tibeto-Burman territory indicated that, while there is still too little agreement on the subgrouping and reconstruction of Tibeto-Burman and on the descriptions of some of the relevant languages, the known contact situations in the region make it likely that future research will provide solid support for proposals of shift-induced interference and perhaps also structural borrowing in some of the languages. The strategy most likely to lead to convincing results is the one being pursued by Tibeto-Burman specialists, among them those who attended the 2008 Osaka symposium. The first step is to provide the necessary foundation for an in-depth investigation of contact phenomena in the region by documenting and analyzing the languages and by proceeding with efforts to clarify the family tree of Tibeto-Burman. The methods used for these

tasks are standard descriptive approaches for the former and the Comparative Method for the latter. Once these approaches have advanced significantly, the ‘unanalyzable residue’ that remains after rigorous application of the Comparative Method will point to areas of lexicon and structure in the various languages where interference is especially likely to have skewed the historical developments. At that stage, it should be possible to satisfy the requisites for establishing a past history of intensive language contact in many parts of the region.

References

Bradley, David

- 1980 Phonological convergence between languages in contact: Mon-Khmer structural borrowing in Burmese. *Berkeley Linguistics Society* 6: 259–267.

Campbell, Lyle

- 1997 *American Indian Languages: The Historical Linguistics of Native America*. Oxford: Oxford University Press.

Dwyer, Arienne

- 1995 Dominant-language influence on serial verb constructions in Salar. Paper presented at the 1995 Linguistic Society of America Annual Meeting, New Orleans.

Evans, Jonathan

- 2008 Is there a Himalayan tone typology? (Preliminary findings). Paper presented at the symposium “Linguistic Substrata in the Tibeto-Burman Area”, Osaka.

Godenzzi, Juan Carlos

- 1988 Lengua y variacion sociolectal: el castellano en Puno. In Luis Enrique Lopez (ed.), *Pesquisas en linguistica andina*, Lima: Consejo Nacional de Ciencia y Tecnologia-Universidad Nacional del Antiplano-GTZ.

Hayashi, Norihiko

- 2008 Historical development of Youle Jino and the linguistic substratum of Tibeto-Burman. Paper presented at the symposium “Linguistic Substrata in the Tibeto-Burman Area”, Osaka.

LaPolla, Randy J.

- 2008 Causes and effects of substratum and superstratum influence, with reference to Tibeto-Burman languages. Paper presented at the symposium “Linguistic Substrata in the Tibeto-Burman Area”, Osaka.

Matisoff, James A.

- 2001 The interest of Zhangzhung for comparative Tibeto-Burman. In Yasuhiko Nagano and Randy J. LaPolla (eds.), *New Research on Zhangzhung and Related Himalayan Languages*, Senri Ethnological Reports 19, 155–180. Osaka: National Museum of Ethnology.

Shirai, Satoko

- 2008 Directional prefixes in nDrapa and neighboring languages: An areal feature of Western Sichuan. Paper presented at the symposium “Linguistic Substrata in the Tibeto-Burman Area”, Osaka.

Takahashi, Yoshiharu

- 2008 On the verbal affixes in West Himalayan. Paper presented at the symposium “Linguistic Substrata in the Tibeto-Burman Area”, Osaka.

Thomason, Sarah G.

2001 *Language Contact: An Introduction*. Edinburgh and Washington, DC: Edinburgh University Press and Georgetown University Press.

Thomason, Sarah G. and Terrence Kaufman

1988 *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.