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.[*]

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]

---

[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**	*semantic shift*	semantic extension	
semantic confusion	**3**	loan shift	*semantic extension*	
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	**5**			
derivative	**5**			
compound	**5**			
creations induced creations (non-borrowings)	**2**		part of loan translations	loan creations
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 [skaɪ] 'sky' and [skreɪpɚ] 'scraper' to signify 'very tall building'. If we were to use (near-)synonyms [hɛvɨnz] 'sky' ("the heavens poured down rain") and [greɪzɚ] '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):

FORM **[plætəpʊs]**

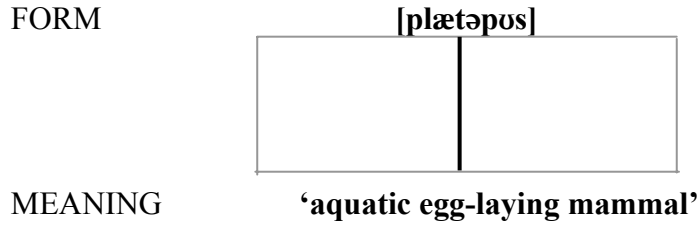MEANING **'aquatic egg-laying mammal'**

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.

F **[hɔg]** **[pɪg]** **[oɪŋkɚ]** **[hɔg]**

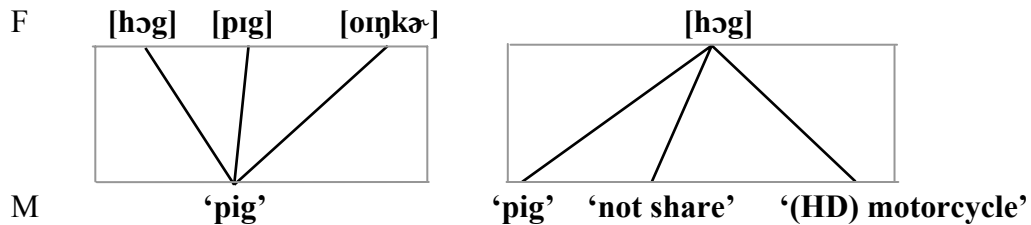M **'pig'** **'pig'** **'not share'** **'(HD) motorcycle'**

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.
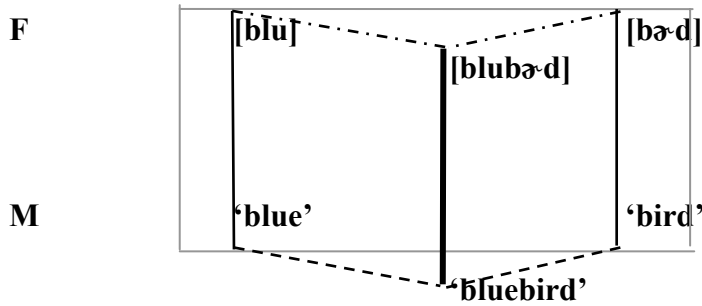

Figure 3. Compositional compound.

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.
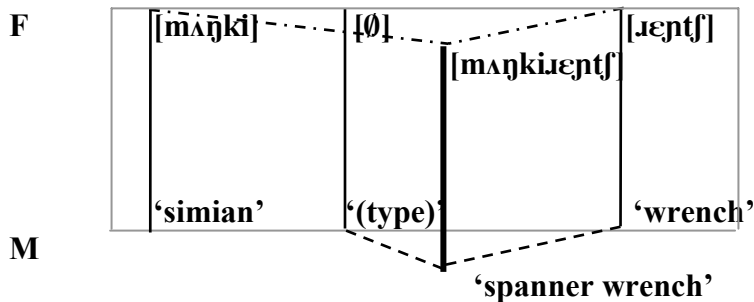

Figure 4. Idiomatic compound.

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 metaphorically[5] 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.
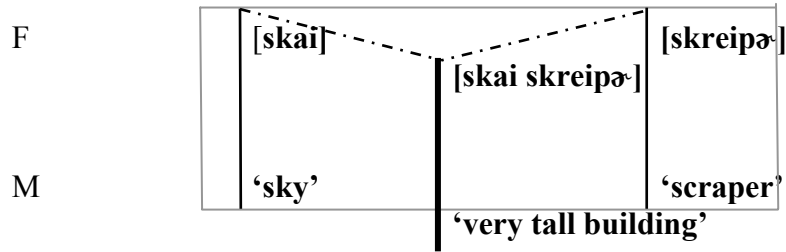

Figure 5. Metaphoric compound.

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_{ROOT} + 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*.
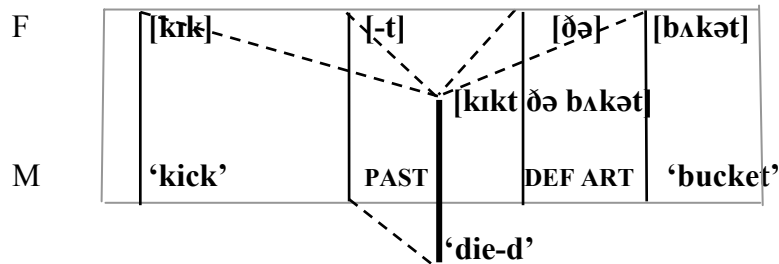
Figure 6. Phrasal compound.

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.
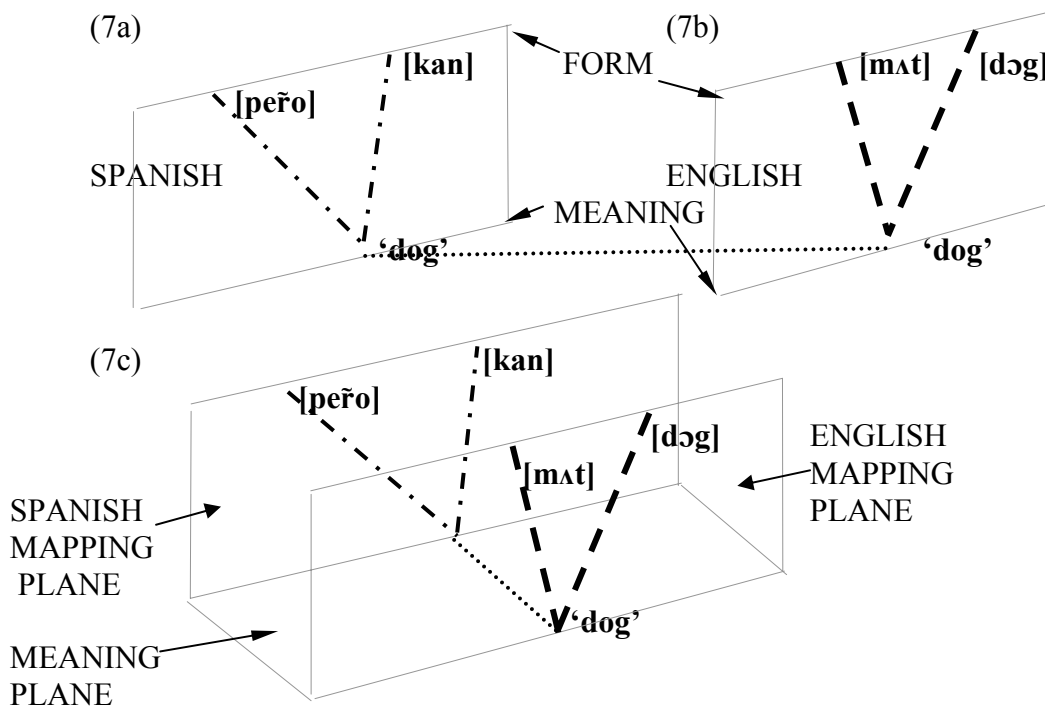
Figure 7. Spanish and English mappings of 'dog'.

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.
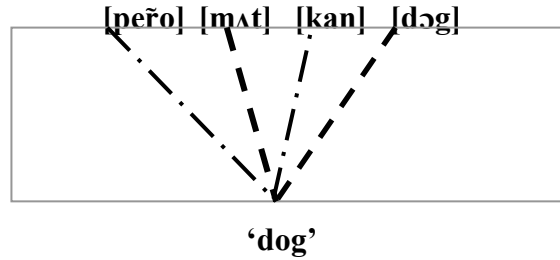
[peřo] [mʌt] [kan] [dɔg]

'dog'

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 [buřito] vs. English [bəɹirou], 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.

[tako]

SPANISH   **[tʰakoʊ]**

ENGLISH

'taco'
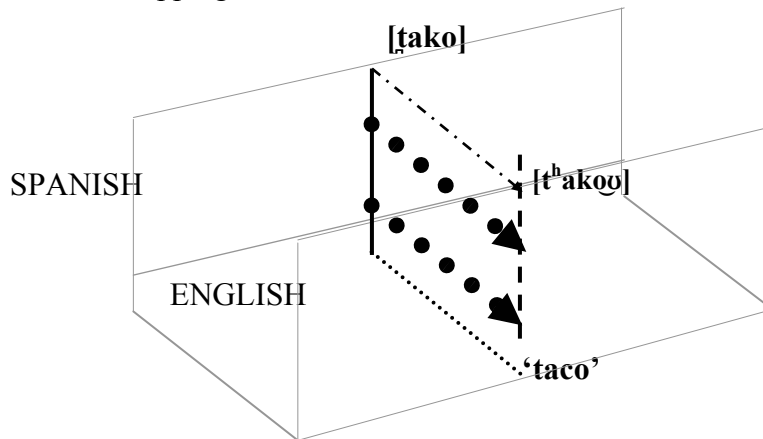
Figure 9. Pure loan.

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* [xusˈɣa(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" [baikiŋgu̥] 'buffet', probably semantically from 'smorgasbord' or from the name of a restaurant chain.[7]
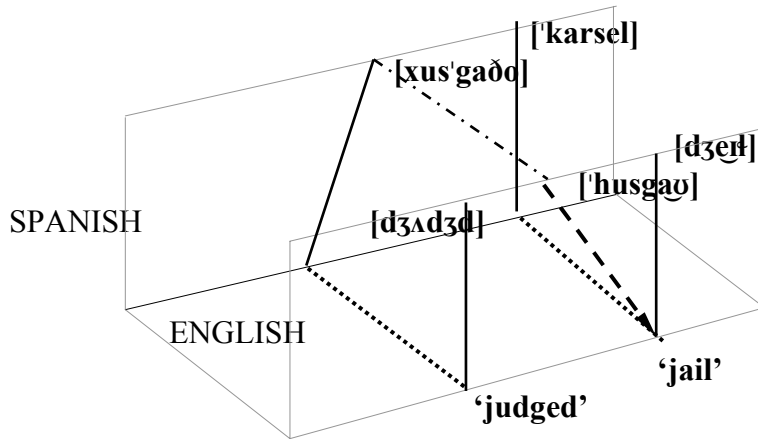
Figure 10. Polyphony through borrowed form.

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.
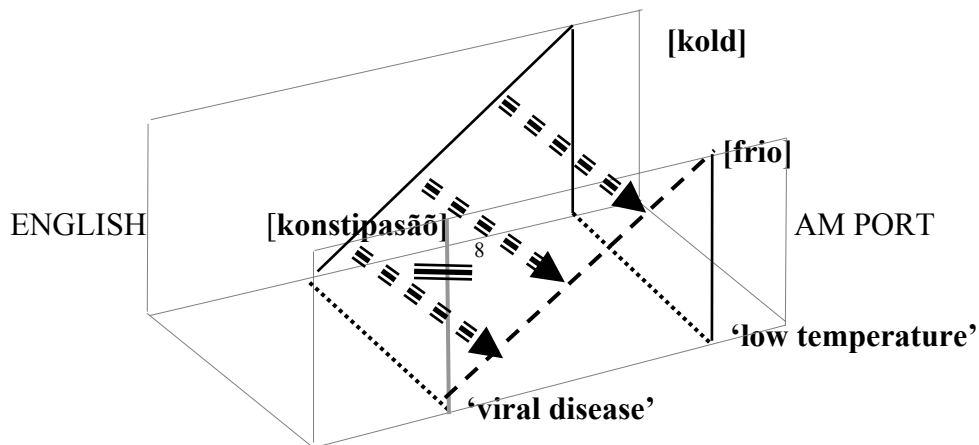
Figure 11. Borrowed mapping.

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

---

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 morphosyntaxis, but the borrowed polyphony of *dereito* 'right' as "correct" occurs only in this phrasal frame.
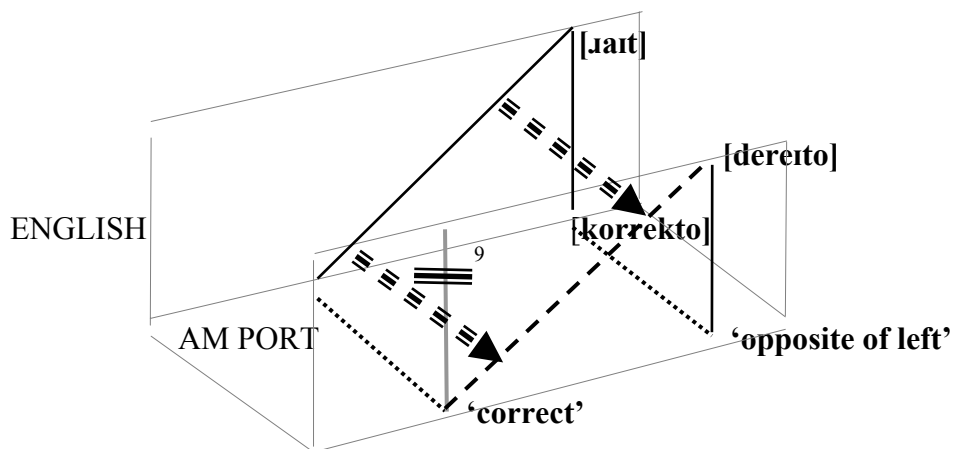
Figure 12a. Borrowed mapping.
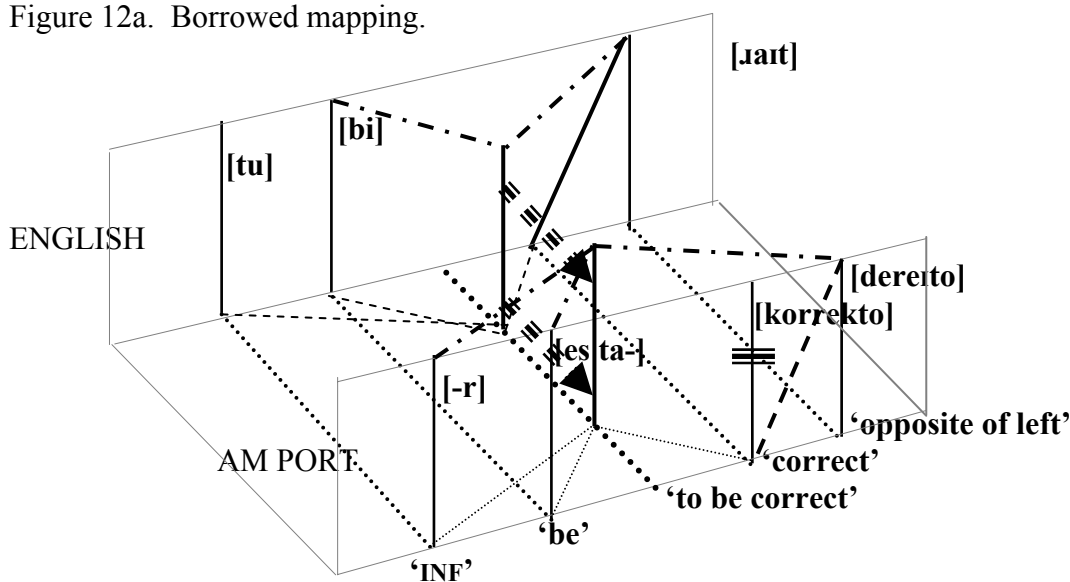
Figure 12b. Multi-faceted borrowing of mapping.

---

[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.
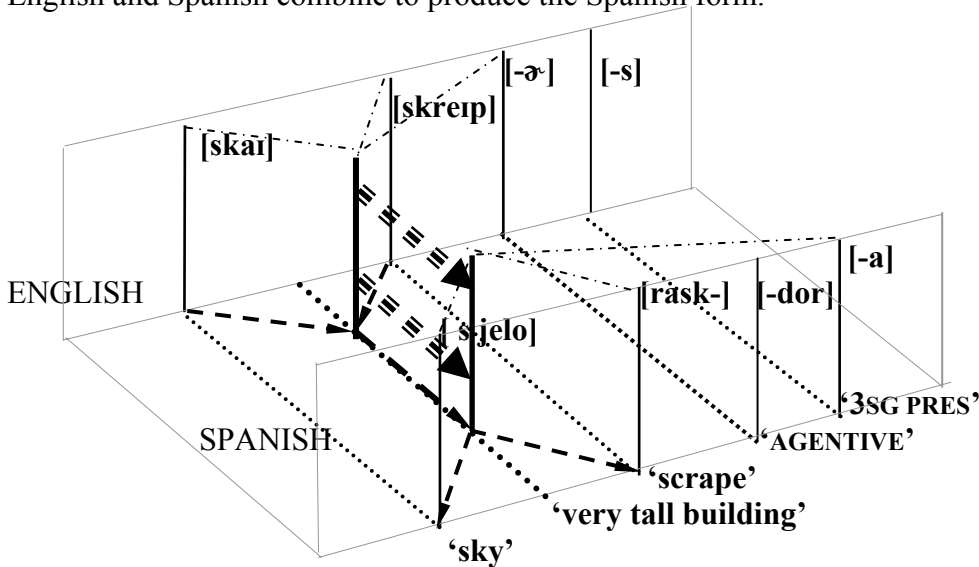
Figure 13. Partial mapping borrowing.

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.
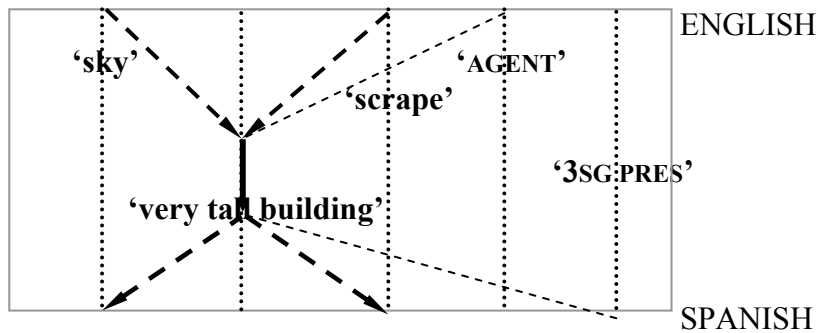
Figure 14. Partial mapping congruence.

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 [bɔs][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.
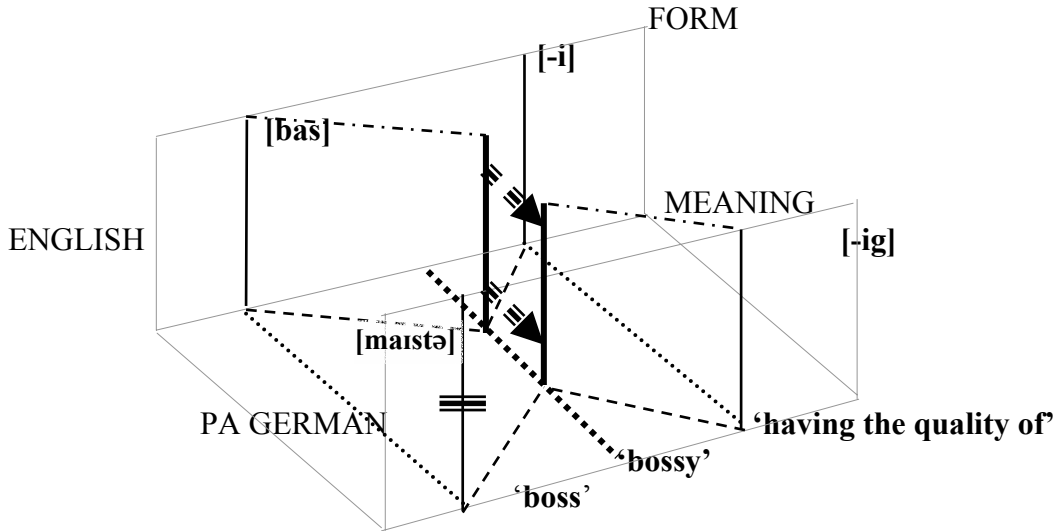
Figure 15. Super blend.

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 mono-semous 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 kíʔnì lǎ* 'without' ← Spanish *sin* 'without' + *kíʔnì 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 [wʊdčʌk] for Algonquian [wučak] 'woodchuck'
e. [LF − BF] Atepec Zapotec [xaɣwar (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* [læ̃ʒəʁi], commonly pronounced in Standard American English (SAE) with a correctly nasalized but incorrectly placed first vowel and a "French" ending [lɑ̃ʒəɹeɪ], 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' [paɹˈteɪ] 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 *taco*
Loanblends	partial morphemic import		
stem	mixed monomorpheme		E→AmN *kårna* 'corner'
derivative	mixed types of morphemes		E→PaG *bassig* 'bossy'
compound	mixed free morphemes		E→PaG *blaumǝpaɪ* 'plum pie'
LOANSHIFTS	borrower word changes meaning		
Extensions	semantic loan (added meaning)		
homologous	stimulus	= form, ≈ meaning	E→QFr *librarie* = 'library'
		= form, ≠ meaning	E→AmP *grosseria* = 'grocery'
homophonous			
synonymous		≠ form, = meaning	E→AmP *frío* = 'cold (disease)'
Creation (calque)	imported arrangement		
literal	identical arrangement		E→AmP *estar dereito* 'to be right'
	approx. arrangement		E→Spanish *rascacielo* 'skyscraper'
approximation			
CREATIONS			
Induced creation	imported meaning only		Pima "downward tassles" 'oats'
Hybrid creation	imported meaning, mixed form		Yaqui *lios nooka* "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 signi-ficant 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 con-gruent. 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 dif-ference 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 exis-tence of a mapping holds true although in some ways the source of the mapping is much more difficult to ascertain for words like PaG *blauməpaɪ* '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 mean- ing, 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 *grosería* '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 exten- sion**, 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 map- ping 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 *¢ih* '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 ¢*ih*.

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 *grosería* '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 *grosería*, 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 distinctions[17] are obliterated through the influence of partial interlingual synonymity", as in AmP *livraría* '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 *blauməpaɪ* '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 *blauməpaɪ* 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] Diccionários 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	stimulus = form	LF=BF, BM+LM, BA[22]
synonymous G	stimulus ≠ 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
Lender — **Lender**	**Lender**	loanword **A**		
	Mixed		loan compound **D** loan derivative **C**	
	Borrower	calque **J**	*loan extension* **H** *loan shift* **I**	
Mixed	**Lender**		loan homologue **G**	
	Mixed	loan homophone **F**	blended stem **B** analogue **E** (loan homonym)	
	Borrower		**loan rendition K**	*loan mapping*
Borrower	**Lender**		semantic confusion	

	Mixed	*created blend* **M**
		  (hybrid creation)
			reverse substitution **N**
			creation compound **M**
			  (blended compound)	

	Borrower	induced creation **L**
		  (*creation*)
			creation extension **H**
			creation shift **I**
				native
				  vocabulary

(MAPPING is the label running vertically on the left of the table.)

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 Otomanguean 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.

AZ form	Spanish form	English gloss
gázhú[24] [ɣážú]	*ajo* [ášo]	garlic
kuléká	*clueca* [kluéka]	brood hen
ánjeli [áŋxeli]	*angel* [áŋxel]	angel
kumárí	*comadre*	godmother
gutzĩlu	*cuchillo* [kučiλo]	candle
kurûtzi	*cruz* [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.

AZ form	Spanish form	English gloss
gwinda	*guinda*	burnt red color (of livestock)
demasiádú	*demasiado*	very much, a whole lot
dilijensia	*diligencias* 'investigation'	investigation
krióyú	*criollo*	native (to the town)

Table 5. "Gap" loans [LF, BM, =LM, LA]**.**

---

[24] In these forms, several processes of adaptation are occurring: 'garlic' [ašo] *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]

AZ form	*Spanish form*	*English gloss*
kampaníyú	*campanillo* 'little bell'	brown-backed solitaire (bird)
kuyóté	*coyote*	coyote (AZ *tzawiyo*)
jagwar	*jaguar*	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 *demasiádú* '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' *yáʔà*) with the Spanish form.

AZ form	*Spanish form*	*English gloss*
berde limón	*verde limón*	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).

AZ form	*Spanish form*	*English gloss*
pitu kanúá	*pico canoa* '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.

AZ form	Spanish form	English gloss
kustíyú kiaʔ	*costilla* 'rib' + 'my'	my wife
layaʔ jwísíú	'tooth' + *juicio* 'wisdom'	wisdom tooth
wêʔ ...kwerda	'to give' + *cuerda* 'cord'	to encourage s.o. to speak
tsiʔnu ora	'twelve' + *hora* '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, *tsiʔnu 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 *lawiʔ tsá*, "middle [of] day", is a clear demonstration that a lexeme existed previous to Spanish contact, but the synonymity of these two forms may differ in punctuality.[28]

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

AZ form	Spanish form	English gloss
ébèkkíá … fotu	'to take out' + *foto* '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' *vγázo* 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 (onomatopoetic) *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.

AZ form	Spanish form	English gloss
gútèʔ...kwenta	'to give' + *cuenta* '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.

AZ form	Spanish form	English gloss
kuttsí ettséʔ	*cochi* '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.

AZ form	Original meaning	Spanish meaning(s) added to AZ form
íchittha	to raise, lift (*levantar*)	to conduct (a census); to give/bear (false testimony); to take minutes/document
íthella	to command (*mandar*)	to send (a letter)
llaʔ	leaf (*hoja*)	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

87

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.

	MEANING			
FORM	**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'.

AZ form	*Spanish form*	*English gloss*
alambre yettséʔ	*alambre* 'wire' + 'spine'[30]	barbed wire
zhúáʔ xtíla	maize + *(Ca)stilla* '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.

AZ form	*Spanish form*	*English gloss*
ìyyà zhubànà karnérú	'flower' + 'tail' + *carnero* 'lamb'	lamb's tail (*Sedum spp.*)

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 wellaʔắréʔè* '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**:

AZ form	Spanish form	English gloss
gútèʔ...mensaje	'to give' + *mensaje* '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**.

AZ form	Spanish form	English gloss
gúni...bwelta	'to make/do' + *vuelta* 'turn; occasion'	to go for a walk
thú...ttu nesesidad	'to have' + 'a' + *necesidad* 'poverty'	to be in poverty
éttíá lista kìʔ	'call out' + *lista* '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**				
	Lender	**BA≠ LA**		**BA=LA**	**Borrower**
		(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

AZ form	*Spanish form*	*English gloss*
lachasuélá	*la hacha* 'the axe' + *azuela* 'adze'	adze
lásu kabréstú	*laso* 'loop' + *cabresto* '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.

AZ	*Spanish form and gloss*	*Current AZ gloss*
duĕndé	*duende* 'goblin'	malignant
kosku, josku	*josco, hosco* 'dark red color of animals'	well, very well; of beautiful color
kwáyú	*caballo* 'horse' [kaβáyo]	colt
lúkkú	*loco* 'crazy'	surly, "snooty"
maskiʔ	*mas que* 'more than'	don't do it!; it doesn't matter
rruínu	*ʔruín* 'ruin'	affected, simpering
(ka) uxtísíá	(PL) + *justicia* '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 *maskiʔ* 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' *laʔuxtísíá* 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 *zhúá xtílá* 'wheat', is that the word for 'strawberry' *dígáʔ 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 laʔ- (*laʔgo* '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 *laʔ* is probably not a confusion.

AZ form	Spanish form	English gloss
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**.

AZ form	Spanish form	English gloss
kuttsí kwíní	*cochi* 'pig' + 'small & fat'	pig (small breed)
kuttsí lúlaʔá	*cochi* '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.

AZ form	Spanish form	English gloss
nékútóʔ[33]	*conejo* 'rabbit' + DIM	'daddy' "little rabbit"
sópaní	*sopas* '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.

AZ form	Spanish form	English gloss
éyakka ... arrepentir	'to compose (oneself)' + *arrepentir* 'to repent'	to regret; to repent
gúdètàʔ...lístá	'to make inclined' + *lista* 'ready'	to make ready
gáppá...interés	'to have' + *interés* 'interest'	to be interested
gúni...prueba	'to make/do' + *prueba* 'proof'	to test, to prove
gúni...remédíú	'to make/do' + *remedio* 'remedy'	to cure

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

---

[33]In *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.

AZ form	Spanish form	English gloss
fwértèní	*fuerte* 'gravity' + 'done'	serious
tántuání	*(en)tanto(que)* 'inasmuch as' + 'done'.	inasmuch as
gú-kodia	POT+CAUS - *joder* '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 *étzeʔétzeʔ* 'to reconcile' ← *étzeʔ* 'to meet' (an ultimate meeting of the minds). Reduplication as an intensive is a common pattern in, e.g., Jamaican Creole[34] *blæk* 'black', *blæk blæk* 'very black', and the most intense would be the limit.

AZ form	Spanish form + AZ morpheme	English gloss
desdebá	*desde* 'out of' + *bá* 'from'	out of, from
astaʔna	*hasta* 'until' + *na* 'until/since'	until, till
áttu buéltá	*áttu* 'other/again' *vuelta* 'time/occasion'	again
para kiʔní	*para que* 'so that' + *kiʔní* 'because'	so that
porkiʔní	*porque* 'because' + *kiʔní* 'because'	because
sin kiʔní lǎ	*sin que* 'without'+ *kiʔní lǎ* '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 ìxxìʔ* 'bobcat' "forest/wild housecat", which coexists with the fully AZ *betziagá*, and the other is *kuttzí ìxxìʔ* '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.

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

AZ form	*Spanish form*	*English gloss*
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.

AZ form	*Morphology*	*Morphology gloss*	*Gloss*
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.

AZ form	*Original meaning*	*Added meaning*
gú, gû	sweet potato	potato (*papa*)
gúxataʔ	to flatten/smash	to iron (*planchar*)
íthi …(žîtsiʔ)	to wring (breast, teat)	to milk (*orderñar*)
ìyyà	metal	syringe (*jeringa*)
kíttsá	to thunder	to shoot (a gun) (*disparar*)
tseni	indolent, apathetic	scarecrow (*espantapájaro*)

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

---

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

[36] Terrence Kaufman (p.c.).

94

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.

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

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.

MAPPING	FORM	MEANING			
		Lender	**Equal**	**(Added)**	**Borrower**
Lender	**Lender**	Loanwords	Gap Loan/ Replacement	Loanword Extension	Form Shift Calque
Lender	**Mixed**		Calque Blend	Loan Blend Extension	Calque Blend Loanblend Shift
Lender	**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
Borrower	**Mixed**		Creation Blend Compound/ Derivative/ Stem; Semantic Reduplications	Loan Homonyms (BF=LF)	Created Blends (−BF) Creation Blend Shift Reverse Substitution
Borrower	**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 *zhúáʔxtíla* '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. PAPEN. 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 Titsaꞌ Quiꞌ Riꞌu*. 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.