Language mixing and genetic similarity
The case of Tojol-ab'al

Danny Law
University of Texas at Austin

Definitions of 'mixed' or 'intertwined' languages derive almost entirely from studies of languages that combine elements from genetically unrelated sources. The Mayan language Tojol-ab'al displays a mixture of linguistic features from two related Mayan languages, Chuj and Tseltal. The systematic similarities found in related languages not only make it methodologically difficult to identify the source of specific linguistic features but also mean that inherited similarity can alter the processes and outcomes of language mixing in ways that parallel observed patterns of code-switching between related languages. Tojol-ab'al, therefore, arguably represents a distinct type of mixed language, one that may only result from mixture involving related languages.

Keywords: Mayan languages; Tojol-ab'al; mixed language; code switching; language contact; similarity

There seem to be no data which suggest a period of common development for Chuj and Toholabal. John Robertson (1977: 120)

Tojolabal is almost identical to Chuj. Nicholas Hopkins (2006: 408)

1. Introduction

The discussion to date surrounding the linguistic affiliation of Tojol-ab'al, a Mayan language spoken today by an estimated 55,000 people in Chiapas, Mexico (INEGI 2015), can be summarized as a debate between two conflicting proposals: (1) that Tojol-ab'al is a member of the Q'anjob'alan subgroup of the Mayan language family, closely related to Chuj, but has massive contact-induced similarities with the Tseltalan languages Tseltal and Tsotsil; or (2) that Tojol-ab'al is a member of the Tseltalan subgroup of the Mayan language family, but has massive contact-induced

Diachronica 34:1 (2017), **40–78**. DOI 10.1075/dia.34.1.02law
ISSN 0176-4225 / E-ISSN 1569-9714

similarities with Q'anjob'alan languages, particularly Chuj (See Figure 1 below). I will argue here that neither proposal is correct; nor, as will be shown, is either entirely wrong. An evaluation of Tojol-ab'al's linguistic past rests on the consideration of two key questions: (i) to what genetic source can each element in the language be linked? – since, as will be seen, the language is remarkably mixed with respect to the historical origin of different linguistic features – and (ii) by what processes did this unusual mixture come about?

The identification of 'mixed' or 'intertwined' languages (Michif, Mednyj Aleut, Media Lengua and the like; see Thomason & Kaufman 1988; Bakker & Mous 1994; Matras & Bakker 2003) often starts, as I will do here, with the rejection of a binary question of the type framed above. Mixed languages are not clear linear descendants of just one language. However, mixed languages are more than just cases of languages with an unclear history or genetic affiliation. Whether the phenomenon is described categorically (Bakker & Mous 1994) or as a more abstract type that can be approximated in individual instances to varying degrees (Matras 2003), consensus is that mixed languages should conform to a variety of criteria and that, crucially, conforming to these criteria should be taken as evidence of a common and distinctive process of contact-induced change. It is this shared creative process, rather than the formal structural similarities between mixed languages, that justifies their status as a distinct type of language.

I argue that Tojol-ab'al is mixed in the sense, which originally seems to have motivated the category, of a language with such a thorough mixture of features from donor languages that the question of genetic affiliation is no longer relevant. However, the processes by which that mixture came about appear to have been different from those proposed for other mixed languages. Studies of the emergence of Australian mixed languages (McConvell & Meakins 2005; McConvell 2008; O'Shannessy 2005, 2012, 2015) give some empirical support for the proposal (Myers-Scotton 2000) that canonically compartmentalized mixed languages emerged in the context of bilingual code-switching. If so, we might expect that different attested patterns of code-switching would have corollaries in different types of mixed languages. While the available socio-historical information on the emergence of Tojol-ab'al, as with most other proposed mixed languages, is less direct than what is available for the Australian cases mentioned above, I argue that, at very least, the mixture of features evident in the language is consistent with the conventionalization of a particular pattern of code-switching between related languages, what Muysken (2000) refers to as 'congruent lexicalization'. Crucially, congruent lexicalization would not be expected to give rise to the etymological compartmentalization that has been emphasized as distinctive of mixed languages, a compartmentalization that Tojol-ab'al clearly lacks.

2. Mixed languages

Before detailing the case for Tojol-ab'al as a mixed language, it is worthwhile to examine the concept of a 'mixed language'. In its most trivial sense, the term 'mixed language' could well apply to any human language, since all have at least some mixing of forms from different sources. However, there is consensus among language contact scholars that a particular kind of contact language exists that is categorically different from other outcomes of language contact and that this type of language merits the label of 'mixed language'. Consensus, however, rapidly breaks down in the nitty-gritty of defining exactly what a mixed language is.

Since work by Thomason & Kaufman (1988), which described several mixed languages, including Michif, Mednyj (Copper Island) Aleut and Ma'a, mixed languages have enjoyed increased attention in the literature on language contact. Some authors take the existence of mixed languages as a counterpoint to arguments about universal constraints on linguistic borrowing (Campbell 1993; cf. Moravcsik 1978). Others argue that they are the 'exception that proves the rule', claiming that both the linguistic and the social contexts that produce mixed languages, and the structural details of the languages themselves, suggest that mixed languages are categorically different from other contact phenomena. Only in mixed languages have the normally operating constraints on borrowing been lifted, allowing the marked and remarkable etymological mixture that characterizes them.

Scholars also differ with respect to what languages can rightly be labeled 'mixed', as opposed to those that are merely the result of 'heavy borrowing' (Thomason & Kaufman 1988: 50). Thomason (2001, 2003), for example, emphasizes rupture in historical transmission, labeling as a mixed language any language "whose grammatical and lexical subsystems cannot all be traced back primarily to a single source language" (Thomason 2003: 21). This definition logically extends the umbrella to cover not only prototypical 'intertwined' languages (Bakker & Mous 1994: 4), like Media Lengua, Michif, and Mednyj (Copper Island) Aleut, but also the larger group of pidgins and creoles. The defining criterion, in this view, is that none of these languages can be properly ascribed to a particular position in a particular language family, or, to say it differently, they do not trace their inheritance from a single parent language, a requisite of the Stammbaum model of historical linguistic relationships.

Another common definition of mixed languages specifically excludes pidgins and creoles, claiming that in these languages "it is easy to identify the source of the lexical component, but it is generally impossible to identify a source for the grammatical component (Bakker 2003: 108). This is because pidgins and creoles are argued to involve simplification and radical restructuring of the grammar

(though see, for example, DeGraff 2003, 2004 for critiques of this view). Under this definition, another practical requirement for identifying mixed languages would be the ability to identify whole lexical and morphological components taken from each of the source languages for the bulk of the language.

Mixed languages have also been defined in terms of the relationship of the speakers with the source language. Mixed languages, minimally, would be language "varieties that emerged in situations of community bilingualism …" (Matras & Bakker 2003: 1). However, the degree of bilingualism in a community at the moment in which a mixed language emerges is often inferred from the linguistic evidence of mixture, rather than attested directly. In addition, this criterion would exclude languages like Ma'a, several Para-Romani varieties and numerous 'secret' languages from the category of mixed languages since it is argued that these cases have involved "rather irregular contact with the donor language" (Matras & Bakker 2003: 10) and do not necessarily stem from anything approaching full bilingual ability in the donor language(s).

In fact, the relationship of the speakers of a mixed variety to the source languages has also led some to restrict membership in the category of mixed languages from the opposite end, suggesting that many cases labeled as 'mixed languages' are better understood as pervasive code-switching by fully bilingual communities, though once again the actual patterns of code-switching that led to a mixed language are not often attested directly. Only those cases in which the language form has become codified and obligatory, rather than fluid and optional, can rightly be labeled 'mixed languages' (Auer 1999; Myers-Scotton 2000). In practice, this can complicate the analysis quite a bit, since most proposed mixed languages are spoken entirely by individuals who are also fluent in one or both of the donor languages, and therefore speakers of the mixed language have constant access to the source language(s) to creatively expand and change the mixed language. Michif (Bakker 1997) seems to be the only clear case in the literature in which most speakers are no longer bilingual in either of the donor languages (Cree and French).[1] Tojol-ab'al, if it is a mixed language, would qualify as well, since there

1. It is, however, often the case that speakers of a mixed language are only familiar with one of the donor languages, not both. For example, the Russian/Aleut mixture Mednyj (Copper Island) Aleut appears to have been spoken throughout its existence in a multilingual social context in which Russian served as the wider language of communication (Golovko 1994: 114), but do not necessarily speak Bering or other Aleut languages. Ma'a speakers (aka 'inner Mbugu') also speak a non-mixed (Bantu) language, called 'normal Mbugu' by Mous (1994:176) along with Ma'a, but no southern Cushitic languages appear to be spoken in the community. The Quechua/Spanish mixture 'Media Lengua' is spoken largely by individuals with some degree of fluency in Spanish or Quechua, or both (Muysken 1994: 210).

are many individuals in Tojol-ab'al communities who do not speak either of the apparent source languages.

Finally, scholars often assert that a hallmark of true mixed languages is a distinctive compartmentalization of the subsystems of the language according to the historical source language. The most common division is for one donor language to provide the vast majority of the lexicon (i.e., > 90%). The source of the lexicon is often called the 'lexifier' language (Muysken 1981) while the language that provides the majority of the inflectional morphology, syntactic patterns, etc. (see Matras 2003) is called the matrix or INFL language (Matras 1998).

As will be shown below, Tojol-ab'al does not fulfill most of the above criteria. In addition, the Tojol-ab'al case introduces a further level of complexity into the discussion, since the two primary contributing languages, Tseltal and Chuj (or a Chujean language), are genetically related. To date, almost all of the languages that have been entered into the discussion on mixed languages have involved donor languages that are clearly unrelated (cf. van Bree 1994; Dreyfuss & Oka 1979; Schadeberg 1994). This seems to be, in part, the result of theorists' desire to avoid undue analytical complexity (Bakker 2003: 108). The genetic relatedness of the donor languages not only complicates attempts to accurately quantify the level of contact, since one must sort out similarities due to language contact from those that might be attributable to common inheritance, but it also provides a possible motivation for the unusual mixture of linguistic features evident in Tojol-ab'al, and it may shed additional light on the role that linguistic similarity can play in determining the outcome of linguistic contact.

3. Tojol-ab'al and language contact

The most influential classification of Tojol-ab'al, proposed by Kaufman (1969, 1976), places it as a close sister of Chuj, in the Q'anjob'alan branch of Mayan. Robertson (1977) proposes that Tojol-ab'al is better understood as a sister of the closely related languages Tseltal and Tsotsil, in the Ch'olan-Tseltalan branch of Mayan. Figure 1 summarizes these conflicting configurations.

There are relatively plentiful published statements concerning the genetic affiliation of Tojol-ab'al, but Robertson (1977, 1992), Schumann (1981, 1983) and Dakin (1988) are the only published works with linguistic data and historical arguments regarding the issue. Dakin includes some reference to unpublished work by Terrence Kaufman. Of these sources, only Dakin (1988) directly engages with the data and arguments presented by others, so the published literature does not constitute a debate. As the only published source that has engaged with arguments on both sides, it is telling that Dakin's conclusion concerning the affiliation

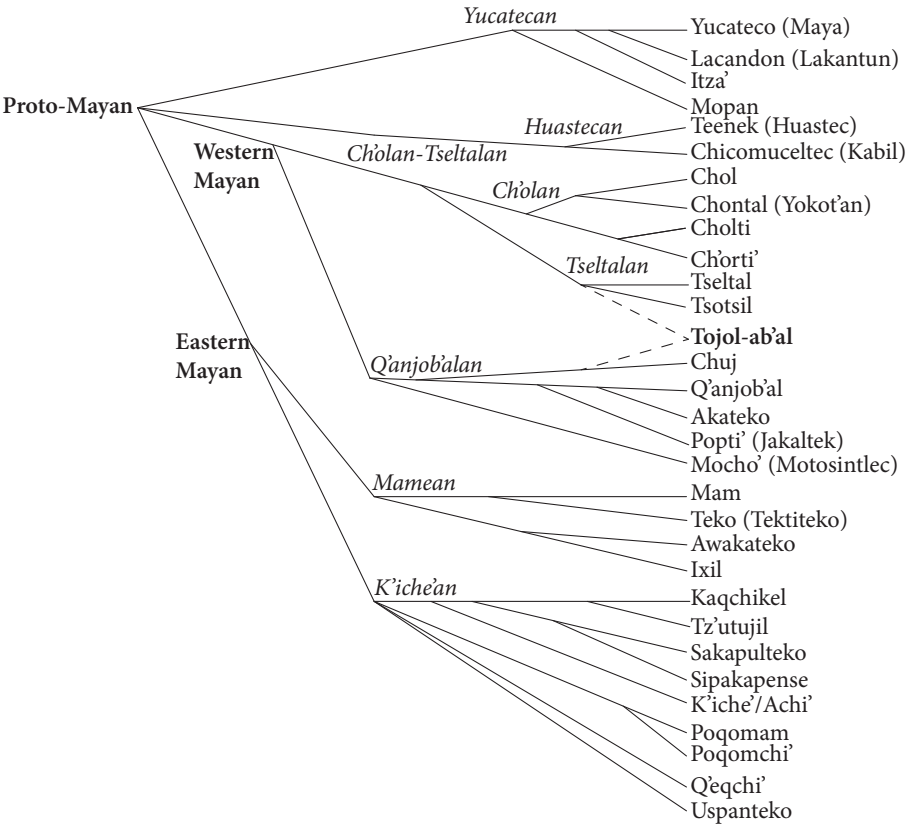

Figure 1. Conflicting placements of Tojol-ab'al in the Mayan family tree

of Tojol-ab'al is simply that it "may yet be premature to decide definitively in favor of one classification over another"[2] (1988: 124). New data and models of language change through contact since Dakin's work open up new possibilities for understanding the linguistic history of Tojol-ab'al.

In §4, I review the linguistic similarities between Tojol-ab'al and Tseltal/Tsotsil on the one hand and Tojol-ab'al and Chuj on the other. The empirical discussion will include data presented in the above-cited sources, as well as data that I have compiled from published descriptions of the relevant languages and my own fieldwork. Based on these data, I argue that the amount and types of materials that Tojol-ab'al shares with each language make it essentially senseless to talk about a single ancestral language.

---

2. My translation of "Todavía puede ser prematuro decidir de manera definitiva por una clasificación sobre otra".

4. Similarities with Chuj and Tseltalan

The evidence of Tojol-ab'al's linguistic history lies primarily in the similarities it shares with Chuj and Tseltal. Interpreting that similarity is complicated by the fact of genetic relatedness between the languages involved. In the following sections, I will detail those similarities in the context of comparative data from other Mayan languages, with an emphasis on the most directly relevant subgroups: Q'anjob'alan and Ch'olan-Tseltalan. Where useful and possible, I will highlight features that can reasonably be traced back to a shared common ancestor of these two subgroups (retentions), in contrast to shared features that appear to be innovative.

4.1 Shared phonological innovations

A logical starting point for investigating the linguistic affiliation of Tojol-ab'al is to look at the shared phonological innovations that might group Tojol-ab'al with one language or another. Since at least the latter half of the 19th century, linguistic innovations, and particularly sound changes, have been central to defining subgroups in language families. At least 12 sound changes from the common language are evident in one or more of the most directly relevant Mayan languages: Tojol-ab'al, Tsotsil, Tseltal, Chuj and Q'anjob'al. Unfortunately every one of these features is demonstrably areally shared. Three of these (*/tʲ/ > /t/, */r/ > /j/ and */k/ > /ʧ/ /_[V_{front}]) applied equally to all of the Ch'olan-Tseltalan and Q'anjob'alan languages. Two more changes, the loss of contrastive vowel length (pM *VV > V) and the merger of */h/ and */x/, are attested in some but not all languages and dialects of languages in both Q'anjob'alan and Tseltalan.

The innovation of a new phoneme /p'/ is found in Tseltalan languages, among others, and not in Tojol-ab'al nor Q'anjob'alan. However, this is clearly an areal innovation, and the fact that Tojol-ab'al and Q'anjob'alan retain a conservative feature is not indicative of a closer genetic relationship in any event. Similarly, the innovation of retroflex fricatives and retroflex affricates cannot be used to group Chuj and Tojol-ab'al with Tseltalan languages because retroflex consonants are an areally shared innovation likely originating from neighboring Mamean languages (Barrett 2002; Robertson 1977).

The fronting of the Proto-Mayan velar nasal */ŋ/ to merge with */n/ and the fronting of Proto-Mayan back velar */q/ and */q'/ to merge with */k/ and */k'/ are similarly unhelpful. The shift of */ŋ/ to /n/ distinguishes Tojol-ab'al from Chuj, while the fronting of the back velars is shared by Chuj and Tojol-ab'al. But, once again, there is distributional evidence that these innovations were areally shared:

only two other Mayan languages besides Chuj (Mocho' and Popti') retain the pM */ŋ/, and the shift of */q/ to /k/ is ongoing in several Q'anjob'alan languages.

Table 1. Relevant phonological innovations from Proto-Mayan

Sound change	Languages involved	Source
pM */r/ > /j/	All relevant languages	Justeson et al. 1985; Law 2014: 33
pM */tʲ/ > /t/	All relevant languages	Law 2013
new /p'/	Tseltalan (not Tojol-ab'al or Q'anjob'alan)	Wichmann 2006; Law 2014: 44
pM *VV > V	All but Mocho	Law 2014: 39
pM *ŋ > n	Tojol-ab'al, Tseltalan, Q'anjob'al, Popti' (not Chuj)	Robertson 1977; Law 2014: 34
pM *q(') > k(')	Tojol-ab'al, Tseltalan, Chuj (not Q'anjob'al)	Law 2014: 43
pM *k > ʧ/_[V_{front}]	Tojol-ab'al, Tseltalan, Q'anjob'alan	Robertson 1977; Law et al. 2014
pM *k > ʧ elsewhere	Tojol-ab'al, Tseltalan, Chuj (not Q'anjob'al)	Robertson 1977; Law et al. 2014
pM *k' > ʧ'	Tseltalan (not Tojol-ab'al or Chuj)	Law et al. 2014
Merged pM */h/,*/x/	Not Chuj, Popti' or some Tseltalan dialects	Law 2014: 41
pM *ʃ > /ʂ	Popti', Q'anjob'al, Mamean	Robertson 1977; Barrett 2002: 312
pM* ʧ(') > / tʂ(')/	Popti', Q'anjob'al, Mamean languages	Robertson 1977; Barrett 2002: 312

Another sound change, the merging of *k(')/ to /ʧ(')/, provides perhaps the best evidence for grouping Tojol-ab'al and Chuj together. The palatalization of */k/ in contexts other than before front vowels (a shared innovation) suggests that both Chuj and Tojol-ab'al should be grouped with Tseltalan. With respect to the lack of palatalization of */k'/, on the other hand (a shared retention), Tojol-ab'al and Chuj pattern with Q'anjob'al. In both cases Chuj and Tojol-ab'al pattern together.

While ambiguous with regard to which branch of Mayan Tojol-ab'al and Chuj should be placed in, this sound change does seem to suggest a particular shared history between Tojol-ab'al and Chuj. However, the evidence is not necessarily as clear as it appears. Shared innovations are more compelling evidence of a linguistic subgroup than shared retentions. In terms of the palatalization of *k('), this would mean that Chuj and Tojol-ab'al would be more appropriately placed with the languages with which they share an innovation (i.e., the Ch'olan-Tseltalan languages), rather than with the languages with which they share a retention (i.e., the Q'anjob'alan languages).

If we were to place Chuj and Tojol-ab'al with Tseltalan on the basis of the shared innovation of */k/ > /ʧ/ before non-front vowels, however, this would mean that the innovation shared among all other Ch'olan and Tseltalan languages of palatalizing */k'/ was areally spread, since it would need to exclude two members of that subgroup (Chuj and Tojol-ab'al). Purely in terms of numbers of shared phonological innovations, this account, which requires assuming the areal diffusion of the sound change */k'/ to /ʧ'/ in Ch'olan and Tseltalan languages, is no simpler than assuming that the shared phonological innovation */k/ to /ʧ/ that linked Chuj and Tojol-ab'al to each other and to Ch'olan-Tseltalan was areally diffused. Indeed, Law et al. (2014) propose, based on comparative linguistic and hieroglyphic evidence, that the palatalization of */k/ and */k'/ was diffused areally in the Maya Lowlands. Thus, no configuration of genetic relationships avoids hypothesizing the areal diffusion of a substantial number of phonological innovations.

The phonological innovations pertinent to these languages, then, only highlight the extensive effects of language contact not only on Tojol-ab'al but on all of the languages in the study. No clear evidence has been garnered to place Tojol-ab'al in one branch or another, but the phonemic inventory of Tojol-ab'al, the result of the layering of numerous areally spread sound changes, has come to be very similar to both Chuj and Tseltalan as a result of contact-induced change.

4.2 Shared morphosyntactic innovations

From a comparative standpoint, one of the most remarkable and perplexing facts about Tojol-ab'al is the strong similarities that it has with both Chuj and the other Q'anjob'alan languages on the one hand and Tseltal/Tsotsil on the other, in terms of grammatical morphology. In this section, we will examine evidence concerning the origins of the morphology and morphosyntax of Tojol ab'al. Obviously, since all of the languages involved belong to the Mayan language family, part of this similarity is due to common inheritance from Proto-Mayan. However, a careful consideration of these similarities in the context of the rest of the family allows us, in many cases, to identify which subgroup or language was the likeliest originator of innovations from Proto-Mayan in Tojol ab'al and which Tojol-ab'al retentions of Proto-Mayan features are conserved in either of the relevant subgroups. Tables 2 and 3 below summarize some of the salient grammatical similarities between Tojol-ab'al and Tseltal (Table 2) and Chuj (Table 3). In these tables, 'innovation' means that the shared form is not reconstructible to Western Mayan, the common ancestor of the Q'anjob'alan and Tseltalan subgroups, nor to Proto-Mayan; 'retention' means that the shared form can be reconstructed to either Western Mayan or Proto-Mayan. Published mentions of these similarities are noted as well.

Table 2. Some Tojol-ab'al grammatical features shared with Tseltal and not Chuj

Feature	PM?	Source
Consistently postverbal set B	INNOVATION	Law 2014; Robertson 1977
Irrealis *-uk*	INNOVATION	Law 2011
Positional mediopassive *-an*	RETENTION	Dakin 1988
Antipassive *-wan*	INNOVATION	Dakin 1988; Law 2011
Lacks *-chaj* passive	RETENTION	Dakin 1988
Causative *-es/-tes*	RETENTION	Kaufman (Dakin 1988)
Unmarked completive intransitive	COINCIDENCE	Dakin 1988; Law 2011
Potential imperative *-an/ -anik*	INNOVATION	Robertson 1977
No agent focus suffix *-on*	INNOVATION	Robertson 1977
No *-ni* in dependent clauses	INNOVATION	Robertson 1977
Agentive RN *-u'un*	INNOVATION	Law 2011
Comitative preposition *sok*	INNOVATION	Law 2011
Negative existential *mayuk/me'yuk*	INNOVATION	Law 2011
Phrase-final topic clitic *=i/=e*	RETENTION?	Law 2011; Dakin 1988
Phrase-final distal clitic *=a*	RETENTION?	Law 2011
1PL inclusive suffix *-tik*	INNOVATION	Robertson 1977; Law 2011
1PL exclusive suffix *-tik+1.ABS*	INNOVATION	Law 2011
1PL > unmarked for number	INNOVATION	Robertson 1977; Kaufman (Dakin 1988)
Incompletive intransitive *x-*	INNOVATION	Law 2011
Completive transitive *x-*	INNOVATION	Law 2011
Plural *-ik* (pos. and adj.)	INNOVATION	Law 2011
Plural for nouns *-tik*	INNOVATION	Law 2011
Plural imperative *-ik* (from **-eq*)	INNOVATION	Law 2011

Table 3. Some Tojol-ab'al grammatical features shared with Chuj or Q'anjob'alan and not Tseltal

Feature	HISTORY	Source
Positional stative *-an*	INNOVATION?	Dakin 1988
Potential *oh=*	INNOVATION	Schumann 1981; Kaufman (Dakin 1988)
Participle *-an*	RETENTION	Schumann 1981;
Interrogative/dubitative *ama*	INNOVATION	Schumann 1981; Kaufman (Dakin 1988)
Progressive *wan*	INNOVATION	Law 2011; Schumann 1981
Transitive perfect *-unej/-nak*	RETENTION	Kaufman (Dakin 1988); Law 2011
Negative existential *me'ey, ma'ay*	RETENTION	Law 2011
Intransitive status suffix *-i*	RETENTION	Law 2011
Transitive status suffix *-a*	RETENTION	Law 2011
Passive *-j*	RETENTION	Law 2011
Mediopassive *-x*	RETENTION	Law 2011
Noun classifiers	INNOVATION	Law 2011
Nominal plural for humans *-e'*	INNOVATION	Law 2011
3PL absolutive *-e'*	INNOVATION	Law 2011
General preposition *b'a(y)*	INNOVATION	Law 2011

4.2.1 *Person marking*

Space does not allow a thorough discussion of all of the features in Tables 2 and 3, but several examples are explored in detail below. A good starting point for comparing the grammatical morphology of Tojol-ab'al with that found in other Mayan languages is the system of person marking. There are marked similarities in the paradigms of person marking for Tojol-ab'al and Tsotsil/Tseltal. Tojol-ab'al, like most Mayan languages, has two distinct sets of person markers that, in most contexts, follow an ergative-absolutive patterning. The so-called 'set A' pronouns prefix to transitive verbs and agree in person and number with the grammatical agent. They also inflect nouns to reference the possessor and are used to indicate several other grammatical relations by inflecting 'relational nouns', a common category in Mesoamerican languages (Campbell et al. 1986). Set A markers in Tojol-ab'al and most Mayan languages have two allomorphs, one that prefixes to vowel-initial stems, and one to consonant-initial stems.

The second set of person markers, 'set B', are obligatory in all predicates. In stative predicates and intransitive verbs, they reference the grammatical subject. In transitive verbs, they reference the grammatical object. Both paradigms of set A pronouns and the set B paradigm are given below (Tables 4–6) for each of the languages, along with the reconstructed paradigms for Proto-Mayan proposed by Kaufman & Norman (1984; 'K&N' below) and Robertson (1992; 'JSR' below):

Table 4. Ergative (set A) pronouns: Preconsonantal set

Set A (/__C)	**1**	**2**	**3**	**1PL (INC)**	**1PL (EXCL)**	**2PL**	**3PL**
CM (JSR)	*nu-	*a-	*ru-	*qa-	–	*e-	*ki-
pM (K&N)	*nu-	*aa-	*u-	*qa-	–	*ee-	*ki-
Q'anjob'al	*hin-*	*ha-*	*s-*	*ko-…(heq)*	*ko-…hon(on)*	*he-*	*s-*
Chuj	*hin-*	*ha-*	*s-*	*ko-…(hek)*	*ko-…-hoŋ*	*he-*	*s-*
Tojol-ab'al	***h-***	***ha-***	***s-***	***h-…-tik***	***h-…-tikon***	***ha-…-ex***	***s-…(-e')***
Tsotsil	*j-*	*a-*	*s-*	*j-…-otik*	*j-…-otikotik*	*a-…-ik*	*s-…(-ik)*
Tseltal	*j-*	*a-*	*s-*	*j-…-tik*	*j-…-yotik*	*a-…-ik*	*s-…(-ik)*

Table 5. Ergative (set A) pronouns: Prevocalic set

Set A (/__V)	**1**	**2**	**3**	**1PL (INC)**	**1PL (EXCL)**	**2PL**	**3PL**
CM (JSR)	*w-	*aw-	*r-	*q-	–	*er-	*k-
pM (K&N)	*w-	*aaw-	*r-	*q-	–	*eer-	*k-
Q'anjob'al	*w-*	*h-*	*y-*	*j-…(heq)*	*j-…hon(on)*	*hey-*	*y-*
Chuj	*w-*	*h-*	*y-*	*k-…(hek)*	*k-…(hoŋ)*	*hey-*	*y-*
Tojol-ab'al	***k-***	***haw-***	***y-***	***k-…-tik***	***k-…-tikon***	***haw-…-ex***	***y-…(-e')***
Tsotsil	*k-*	*av-*	*y-*	*k-…-otik*	*k-…-otikotik*	*av-…-ik*	*y-…(-ik)*
Tseltal	*k-*	*aw-*	*y-*	*k-…-tik*	*k-…-yotik*	*aw-…-ik*	*y-…(-ik)*

Table 6. Absolutive (set B) pronouns

Set B	1	2	3	1PL (INC)	1PL (EXCL)	2PL	3PL
CM (JSR)	*-in	*-at	*-Ø	*-o'ŋ	–	*-ex	*-eb
pM (K&N)	*-iin	*-at	*-Ø	*-o'ŋ	–	*-ix/*-ex	*-eb
Q'anjob'al	*(h)in*	*(h)ach*	*-Ø*	*(h)on… (heq)*	*(h)on…(h)on(on)*	*(h)ex*	*-Ø (heb')*
Chuj	*(h)in*	*(h)ach*	*-Ø*	*(h)oŋ… (hek)*	*(h)oŋ…-(h)oŋ*	*(h)ex*	*-Ø (heb')*
Tojol-ab'al	*-on*	*-a*	*-Ø*	*-otik*	*-otikon*	*-ex*	*-Ø (-e')*
Tsotsil	*-on / -i-*	*-ot / -a-*	*-Ø*	*-otik / -i-…-otik*	*-otik / -i-…-otikotik*	*-oxuk / -a-…-ik*	*-Ø (-ik)*
Tseltal	*-on*	*-at*	*-Ø*	*-otik*	*-yotik*	*-ex*	*-Ø (-ik)*

These data illustrate clearly that the system of person marking in Tojol-ab'al is much more similar to Tseltal/Tsotsil than it is to Chuj. Most notably, Tojol-ab'al, Tsotsil and Tseltal all share the innovative spread of the first-person plural to become the generic first-person marker and postclitics to indicate first-person plural exclusive and inclusive for non-third-person referents based on the morpheme *-tik*.

4.2.2 *Position of the absolutive*

Another similarity between Tseltal and Tojol-ab'al is the position of the set B (absolutive) markers relative to the verb. There is evidence that set B markers in Proto-Mayan came after the stem in participles, stative predicates, unmarked (completive) aspect and imperatives and came before the stem for verbs that were overtly marked for aspect (Robertson 1992: 53). This is essentially the pattern found in both Q'anjob'al and Chuj, as well as in some dialects of Tsotsil, though in Q'anjob'al the third-person plural *heb'* follows the verb in active contexts. Tojol-ab'al and Tseltal, on the other hand, both suffix set B person markers universally, as shown in Table 7 below.

Table 7. Position of set B marker relative to predicate

	Q'anjob'al	Chuj	Tojol-ab'al	Tseltal	Tsotsil
NVP	After	After	After	After	After
Perfect	After	After	After	After	After
Active	Before/after	Before	After	After	Before

Law (2009, 2011) shows that pronoun borrowing among languages connected to the Lowland Mayan sphere of influence was fairly common. The shared pattern for the placement of set B markers described is found in all of the Yukatekan and Ch'olan languages (see Law 2011, ch. 4; Bricker 1977). The degree of similarity in person marking between Tojol-b'al and Tseltalan, however, is beyond that seen in any other Mayan language.

4.2.3 *Plural*

Tojol-ab'al and Tseltalan also share a great deal in terms of the way in which plurality is managed. In most Mayan languages plurality is optionally expressed, with a preference for marking plurality on nouns higher in the animacy hierarchy. Number is often marked on both nouns and verbs, as well as adjectives in some languages. Number agreement on verbs is generally expressed along with person in the two sets of person markers common to Mayan languages.

Tojol-ab'al and Tseltal are like other Mayan languages in this respect, though, along with several other Mayan languages, they have developed a series of plural enclitics so that person and number are not expressed by the same morpheme. In fact, number agreement may have been lost in Lowland languages so that all plurality is nominal or pronominal (see Mateo-Toledo 2008: 49 for this analysis in Q'anjob'al). Plural markers for Tojol-ab'al, Tseltal, Tsotsil, Chuj and Q'anjob'al are summarized in Table 8.

Table 8. Summary of plurals

	Tsotsil	Tseltal	Tojol-ab'al	Chuj	Q'anjob'al
Imperative	*-ik*	*-ik*	*-ik*	*-ek*	*-eq*
Adjective	*-ik*	*-ik*	*-ik*	–	*-eq*
Positional		*-ik, -ajtik*	*-ik*	–	
3rd-person	*-ik*	*-ik*	*-e'*	*-e'*	*heb'*
Nouns	*-etik*	*-etik*	*-tik, jumasa'*	*heb'*	*laq*
2nd-person	*-ik*	*-ik*	*-ex*	–	*-ex*
Nouns (human)	*-ab'*	*-ab'*	*-e'*	*-(h)eb'*	*heb'*
General numeral suffix	*-eb', -ib'*	*-eb'*	*-e'*	*-eb'*	*-eb'*

Tseltal and Tsotsil share with Tojol-ab'al the plural suffix *-tik* and the expanded functions of the plural suffix *-ik*, which is elsewhere restricted to the imperative mood. While Tseltal and Tsotsil have expanded the function of *-ik* even beyond what we see in Tojol-ab'al, the commonality between the two is clear. The suffix *-tik* is likewise transparently similar, as was discussed earlier.

The similarity of plural markers in Tojol-ab'al and the Tseltalan languages becomes even more pronounced upon consideration of Chuj, which is markedly different and more conservative. In Chuj, there is a cognate of Tojol-ab'al's *-ik* (also found in Tseltal and Tsotsil, as well as numerous other Mayan languages) in the form of a plural imperative suffix *-ek*, almost certainly a Proto-Mayan retention. However, this is very restricted in usage: it does not mark plurality on adjectives and positionals, as in Tojol-ab'al and the Tseltalan languages, though Q'anjob'al does have a suffix *-eq* as a plural of adjectives. Finally, its form does not reflect the idiosyncratic vowel change shared by Tojol-ab'al and the Tseltalan languages. Chuj also lacks any semblant of *-tik* for marking plurality.

The primary plural marker in Chuj, however, is *(h)eb'*, which is clearly cognate with Tojol-ab'al *-e'* and which is probably the number suffix *-eb'* in Tseltal and Tsotsil as well. Chuj and Tojol-ab'al, like other Q'anjob'alan languages, have seen a merging of the general number classifier and a historically separate third-person plural agreement marker. In Tojol-ab'al, this suffix is restricted to human referents. In Chuj, it can optionally mark plural on inanimates. The Tojol-ab'al and Chuj forms are cognate with the Tseltalan form *-ab'* and a retention of the Proto-Mayan plural form used for plural absolutive agreement. In Chuj, *heb'* is preposed to noun classifiers to indicate plurality of nouns, and it follows verbs to indicate plurality in third-person verbal agreement.

Thus, Tojol-ab'al shares two of its plural markers with Tseltal and Tsotsil, both of which are clearly innovative in their distribution and their form. In keeping with its 'mixed' appearance, however, Tojol-ab'al also shares a plural marker with Chuj (and the other Q'anjob'alan languages). As mentioned above, this suffix is a retention of a Proto-Mayan plural form, a reflex of which is preserved in Tseltal and Tsotsil, but both the exact form, and the third-person verbal agreement function, are shared by Chuj and Tojol-ab'al.

4.2.4 *Aspectual system*

Tojol-ab'al also shares a remarkable number of features with Tseltal and Tsotsil in the system of aspect, as well as having some shared features with Chuj. These are summarized in Table 9 below.

Table 9. Tense/aspect/mood

	Q'anjob'al	Chuj	Tojol-ab'al	Tseltal	Tsotsil
COMPLETIVE					
TRANS	*(ma)x-*	*x-*	*Ø*	*la*	*l-* (~ *i-* 3rd person)
INTR	*(ma)x-*	*x-*	*Ø*	*Ø*	*l-* (~ *i-* 3rd person)
PROXIMATE	*Ø*	*Ø*	–	–	–
INCOMPLETIVE					
TRANS	*ch(i)-*	*tz-*	*(wa) x-*	*ya*	*t(a) x-*
INTR (3rd)	*ch(i)-*	*tz-*	*(wa) x-*	*(ya) x-*	*t(a) x-*
INTR (non-3rd)			*(wa) la-*	*(ya) x-*	*t(a) x-*
PROGRESSIVE	*lanan*	*wal, wan*	*wan*	*yak*	*yak*
POTENTIAL	*(h)oq-*	*oj-*	*oj-*	–	–

4.2.4.1 *Completive*

The first similarity to note is in the completive aspect. In Tojol-ab'al and, with intransitive verbs, in Tseltal, the completive aspect is unmarked. However, it is not clear what to make of this similarity. The Proto-Mayan completive was most likely unmarked (Robertson 1992: 66), so that the similarity between Tseltal and

Tojol-ab'al would seem to be a retention. However, Colonial Tseltal data show that the Colonial Tseltal completive involved a preverbal clitic *u-* for both transitives and intransitives (Robertson 1992: 184). The fact that an attested prior stage of Tseltal had a clear overt marker for the completive, which was subsequently lost, seems to suggest that, if the unmarked completive is a similarity due to contact, it would have to be the result of recent contact. However, we might also speculate that a conservative, unmarked completive did exist, perhaps in a more restricted function, in Colonial Tseltal but that it has been missed or is simply unattested in the available Colonial materials (see, for example, Mateo Toledo's (2008) identification of an unmarked completive in Q'anjob'al, where no such form had been noted previously). It is also possible that Tojol-ab'al maintained the conservative completive form, while Tseltalan underwent several different changes, ultimately ending back where it began, by sheer coincidence.

While the similarity between Tseltal and Tojol-ab'al in the completive may be inconclusive, the difference between *Chuj* and Tojol-ab'al with respect to the completive marking does seem significant. Chuj, as well as Q'anjob'al, Popti', Akateko and, curiously, most of the K'iche'an languages, indicate the completive aspect with a preverbal marker *x-* (not related to the Tojol-ab'al incompletive, to be discussed below). Q'anjob'al and Chuj also have an unmarked completive/past form, the precise semantics of which are unclear. Robertson (1992: 66) reconstructs a proximate particle *'ix, common as a second-position clitic with the meaning 'soon', 'recently' or 'already'. He proposes that the completive marker in Q'anjob'alan and K'iche'an is derived from this particle. The fact that the meaning reconstructed by Robertson for *'ix now appears to be expressed in these languages with zero-marking, while the more generic completive is marked with *x-*, suggests that these forms switched values (i.e., markedness reversal). If Tojol-ab'al indeed belongs to the Q'anjob'alan subgroup, it either expanded the unmarked form to be the general completive (in effect, returning to a previous state of affairs) or the form of the completive in the other Q'anjob'alan languages (*x-*) developed some time after the separation of Q'anjob'al from other members of that subgroup, including Chuj, which would suggest that the *x-* completive was areally diffused among Q'anjob'alan languages.

4.2.4.2 *Incompletive*

In the incompletive aspect, things are somewhat clearer. All of the languages shown in Table 9 have reflexes of the same Proto-Mayan incompletive aspect marker *k(i)- (Robertson 1992). However, Tojol-ab'al shares the same idiosyncratic mutation of this marker found in Tseltal and Tsotsil: *x-*, while the Q'anjob'alan languages have simple affricativization to *ch-*. It is interesting too that first- and second-person intransitives in Tojol-ab'al have a different form, *la-*, in the incompletive, much as we find for the completive aspect in Tsotsil. These forms have different etymological

sources. The *l-* in the Tsotsil completive is a reduction of the auxiliary *laj-* 'to complete', the source of *la* in Tseltal as well. The *la-* in Tojol-ab'al has been linked with the *la reconstructed for the 'future' form in Proto-Mayan (Robertson 1992).

4.2.4.3 *Progressive and potential*

In other respects, the system of aspectual distinctions in Tojol-ab'al is very similar to Chuj, particularly with respect to both the progressive auxiliary, and the potential or 'future'. The progessive is *wan* in Tojol-ab'al, and both *wan* and *wal* can be found in Chuj. This form is similar to the Lowland progressive auxiliary found in Chol from Tumbalá (*woli*), Cholti (<*yual*>), Ch'orti' (*war*) and possibly even Yucatecan (*walak* in Mopan and Colonial Yucatec).

The potential or future prefix is clearly a Q'anjob'alan innovation. It derives from the historical optative suffix *-oq, which was prefixed and adopted the related function of 'potential', while the old suffix remained in its prior location with its historical function. In Tojol-ab'al, the form that was prefixed and the current form of the optative suffix in the language are not the same (*oj-* vs. *-uk*). This will be discussed in more detail below.

4.2.4.4 *Perfect*

The perfect is another aspectual form that can be used to understand the relationship between Chuj and Tojol-ab'al (Table 10 below). The perfect is different from other features of these languages in that it is indicated with a suffix on the verb. In this sense, it is more like a marker of mood or what Kaufman (1990: 72) calls the 'status' of a verb. It does not co-occur with other markers of aspect in most Mayan languages (one apparent exception being Poqomchi'). In some languages, we can formally distinguish between perfect participles, which function like non-verbal predicates, and verbal perfects, which maintain all of the associated grammatical roles of a verb. In many Mayan languages, however, there is no distinction between these two constructions in terms of the perfect suffix used, and it is often very difficult, in practice, to determine if a given instance of a perfect suffix is verbal or non-verbal, particularly with intransitives and in secondary predications.[3]

For the purposes of determining the genetic affiliation of Tojol-ab'al, this distinction is not crucial. Even with a very imprecise understanding of the perfect in these languages, we can still observe that Tojol-ab'al displays certain similarities with Q'anjob'alan languages that it does not share with Tseltal. Tojol-ab'al's

3. This lack of clarity means that the analysis of the perfect varies not only from language to language but also from linguist to linguist when describing the same language. A careful comparative semantic and structural analysis of the perfect in Mayan languages is beyond the scope of this paper, but might help clarify the matter.

transitive perfect *-unej* is a retention of a Proto-Mayan intransitive perfect form *-Vnaq. A reflex of this morpheme is also found in the Q'anjob'al intransitive perfect *-naq*. Chuj has extended the use of the perfect *-nak* to transitives as well as intransitives. This extension to transitives is an innovation that Tojol-ab'al seems to have shared before later restricting *-unej* to transitives with the innovation of a new intransitive perfect *-el*. Note that the change of morpheme final /k/ to /j/ in Tojol-ab'al is irregular but fairly common, particularly in grammatical morphemes. Tseltal and Tsotsil use a reflex of the Proto-Mayan transitive perfect suffix, *-em,* in an innovative function marking the intransitive perfect, and they use an innovative *-oj* for transitive perfects.

4.2.4.5 *Irrealis and imperative*

While the perfect in Tojol-ab'al is shared with Chuj, the modal suffixes for marking the irrealis as well as the imperative, to a lesser degree, both show innovative characteristics of use and form that are shared with Tseltal and Tsotsil. The Proto-Mayan intransitive imperative *-aŋ (SG) and *-aŋ-ek (PL) is, for the most part, preserved in all of the languages here, with the one caveat that Tojol ab'al follows Tseltal and Tsotsil in having a high vowel *-ik* for the plural rather than *-ek* (see the section on plurals above for more discussion).

The optative of Tojol-ab'al, however – in its form, in its distribution of use and in an irregular fact of its paradigm related to the imperative – shows strong similarities with Tseltalan. The first thing to note is that Tojol-ab'al manifests the same distinctive mutation of the Proto-Mayan optative/irrealis suffix *-oq in the form *-uk*. This is in spite of the fact that the preposed derivative of this suffix, used to mark the future, displays the conservative low vowel *oj-*, a feature that Tojol-ab'al shares with Chuj and the other Q'anjob'alan languages.

Even more interesting for our purposes is the fact that the *-uk* optative suffix is not perfectly regular across the paradigm. As Robertson (1977: 111) notes, in Tseltal, Tsotsil and Tojol-ab'al, the second-person optative form is derived from the imperative. Rather than the combination *-uk-a (toj), *-uk-at (tse) or *-uk-ot (tso), we get *-an*, and in place of their respective plural forms (*-uk-ex [toj and tse], *-uk-oxuk [tso]) we find *-an-ik*. This irregularity in the paradigm is likely to be rather unstable, and Polian (2004: 113) reports *-uk-at* for the Tseltal variety spoken in the town of Oxchuc. The fact that such a marked innovative irregularity is shared by these three languages is significant. These forms, as well as those for the perfect and imperative, are summarized in Table 10 below.

Another significant aspect of the *-uk* suffix in Tojol-ab'al has to do with its syntactic and semantic functions. All of the languages under investigation here seem to use *-uk/-ok* both for irrealis/optative and to mark dependent clauses; however, the marking of dependent clauses is central in Q'anjob'alan, while the irrealis/

optative function is more salient and productive in Tseltal, Tsotsil and Tojol-ab'al (Polian 2007). In addition, while *-uk/-ok* is used in all of these languages to mark negation in non-verbal predicates, in Chuj *-ok* marks negation in verbal predicates as well.

4.2.5 *Category suffixes*

Another noteworthy similarity between Tojol-ab'al and Chuj can be seen in the root category suffixes. There are three suffixes in particular to note here: a positional 'adjective' or stative predicate suffix and two root thematic or category suffixes, one for root transitives and one for root intransitives. As will be seen, Tojol-ab'al looks like Chuj and Q'anjob'alan in all three cases.

Table 10. Perfect, irrealis, imperative and interrogative

	Q'anjob'al	Chuj	Tojol-ab'al	Tseltal	Tsotsil
PERFECT					
TRANS	*-b'il*	*-nak*	*-unej*	*-oj*	*-oj*
INTR	*-naq*	*-nak*	*-el*	*-em*	*-em*
IRREALIS/OPTATIVE					
1st & 3rd persons	*-oq*	*-ok*	*-uk*	*-uk*	*-uk*
2nd person	*-oq*	*-ok*	*-an (PL -anik)*	*-an (PL -anik)*[4]	*-an (PL anik)*
IMPERATIVE					
TRANS	*a' ~ -V', -j*	*-a' ~ -ah/ -V' ~ -Vh*[5]	*-an (PL -anik)*	*-a (all persons; PL -a(w)ik*	*-o (3rd ABS) -Ø (1st & 2nd)*
INTR	*-an (PL -an-eq)*	*-aŋ (PL -aŋ-ek)*	*-an (PL -anik)*	*-an (PL -anik)*	*-an (PL anik)*

The root category of positionals common to all Mayan languages generally cannot be expressed without some sort of suffix; the precise suffix depends on the semantic and morphosyntactic context of the form. Positionals can be derived with a suffix into what have traditionally been referred to as 'positional adjectives' but are now understood, in most Mayan languages, to be stative predicates, which predicate upon one of the arguments of a main predicate (see Aissen & Zavala Maldonado 2010). The stative predicate suffix for positionals in the Q'anjob'alan languages

4. Kaufman (1971: 104) gives this as the form for Tseltal. Polian (2004: 113) has *uk-at* as well. It may be that the more transparent form *-uk-at* is a recent innovation in the Tseltal of Oxchuc, described by Polian, the result of regularization of an irregular paradigm.

5. The vowel used in the transitive category suffix depends on the vowel of the root: if unrounded, then /a/; if rounded, then it matches the vowel quality of the root (Maxwell 1982: 132).

and Tojol-ab'al is *-an*. In Tseltal and Tsotsil, as well as the Ch'olan and Yukatekan languages, it is a vowel-harmonic *-V_1l* suffix. Given this distribution, it seems reasonable to suppose that the state of affairs in Proto-Mayan was similar to what we find in K'iche' (López Ixcoy 1997: 201), where positional roots take *-V_1l*, unless the root has /l/ or /r/, in which case the root takes *-an*. If this is the case, then both Tseltalan (and broader Lowland) and Q'anjob'alan are innovative. Tojol-ab'al clearly patterns with Chuj and Q'anjob'alan in this respect.

The other two category type suffixes that Tojol-ab'al shares with Q'anjob'alan are more clearly retentions of Proto-Mayan forms. These are, first, the intransitive category suffix, which affixes to intransitive roots, and in some languages all intransitive verbs, in all contexts that don't call for some other suffix (i.e., perfect aspect, optative or imperative mood, non-third person, in languages that suffix the set B person markers) and, second, the transitive category suffix, which is parallel in distribution to the intransitive category suffix but which occurs on transitive verbs (one form for root transitive verbs and another for derived transitives). In Tojol-ab'al, Chuj and the other Q'anjob'alan languages (and many others) the intransitive category suffix is *-i* (from Proto-Mayan *-ik), while the transitive category suffix in these languages is *-a'* (toj), *-a'* (chu, q'an; vowel harmonic *-V_1'*, if the root vowel is rounded) for root transitives, and *-Vj* or *-j* for derived ones. The Proto-Mayan form of the root transitive category suffix was most likely a root-vowel harmonic vowel, so the use of /a/ across the board in Tojol-ab'al, and with unrounded vowels in Chuj and Q'anjob'al, is innovative.

The contrast with Tseltal and Tsotsil with respect to these markers is obvious, since in both of these languages the transitive and intransitive status markers were entirely lost, an innovation not shared with Tojol-ab'al. See Table 11 for a summary of these status markers in the various languages.

Table 11. Category and root class suffixes

	Q'anjob'al	Chuj	Tojol-ab'al	Tseltal	Tsotsil
Positional pred	*-an*	*-an*	*-an*	*-V_1l*	*V_1l*
Intransitive category suffix	*-i*	*-i*	*-i*	*none*	*none*
Transitive category suffix	*a' ~ -V', -j*	*-a' ~ -ah/ -V' ~ -Vh*[6]	*-a(w), -aj*	*none*	*none*
Category suffixes phrase final?	*yes*	*yes*	*yes and no*[7]	–	–

6. The vowel used in the transitive category suffix depends on the vowel of the root: if unrounded, then */a/*; if rounded, then it matches the vowel quality of the root (Maxwell 1982: 132).

7. Yes in third person and with CVC roots; no in first & second and with CVCC roots.

4.2.6 *Voice*

The system of voice, in general, seems to show a greater connection between Tojol-ab'al and Chuj than what we've seen in other aspects of the inflectional morphology of these languages (Table 12).

Table 12. Voice

	Q'anjob'al	Chuj	Tojol-ab'al	Tseltal	Tsotsil
Passive					
Root	*-lay*	*-aj*	*-j*	*-ot*	*-e*
Derived	*-lay*	*-ax*	*-j*	*-ot*	*-at*
Mediopassive	–	–	*-x*	*-j-*	*Ø*
Other passive	*-chaj*	*-chaj*	–	*ich' (aux)*	*chi' (aux)*
Antipassive	*-waj*	*-waj*	*-wan*	*-wan*	*-van*

The passive in both of these languages is indicated with an *-(a)j* suffix. The main difference with respect to passive voice is that a morpheme *-(a)x* is used on derived transitives in Chuj whereas that same form has a mediopassive function in Tojol-ab'al; the *-j* passive in Tojol-ab'al has been extended to inflect both root and derived transitives.

But even in the voice system, Tseltalan influence is not entirely absent. While all of the antipassive forms listed here are cognate to a degree, Tojol-ab'al, Tseltal and Tsotsil all display an innovative form *-wan,* historically created as a combination of the two Proto-Mayan antipassive suffixes *-(V)w and *-Vn. The antipassive form *-waj* in both Chuj and Q'anjob'al is likely a reflex of the *-(V)w antipassive.

Note in Table 12 that a *-Vn antipassive form is not the productive antipassive in any of the languages here, at least in the grammatical function that Smith-Stark (1978) called the 'absolutive antipassive', in which the verb is made intransitive and the remaining overt argument is the semantic agent of the verb while the semantic patient is not specified. This does not mean that the *-Vn suffix cannot be found in any of these languages. In fact, as Robertson (1977: 112) notes, Chuj and other Q'anjob'alan languages make particularly productive use of this suffix in a large range of contexts, including cases in which the agent of a transitive verb has been moved to focus position (as also happens with content questions and often in negation), when a transitive verb phrase is relativized and the agent of this relative clause is the same as the noun phrase that the relative clause modifies, or in transitive complement clauses that are not marked for aspect.

Tseltalan and Tojol-ab'al, like several Ch'olan and Yucatecan languages, however, do not have any sort of distinctive marking for focused agents or relative clauses, and they certainly show no signs of the innovative use of this suffix in

complement clauses generally. Once again, it seems that this similarity between Tojol ab'al and the Tseltalan languages is a consequence of larger patterns of regional linguistic exchange.

Table 13. Comparison of contexts for -Vn antipassive suffix

(transitive agent)	Q'anjob'al	Chuj	Tojol-ab'al	Tseltal	Tsotsil
Agent extraction	*-on-i*	*-an-i*	*unmarked*	*unmarked*	*unmarked*[8]
Relative clause	*-on-i*	*-an-i*	*unmarked*	*unmarked*	*unmarked*[9]
Complement clause	*-on-i*	*-an-i*	*unmarked*	*unmarked*	*unmarked*

4.2.7 *Prepositions*

Mayan languages generally have a small set of prepositions (in many cases only one). Most locative and relational meanings are expressed with relational nouns, which are functionally equivalent to prepositions in English but morphologically like possessed noun phrases. Tojol-ab'al has two prepositions, the general preposition *b'a* and the comitative preposition *sok*. The general preposition is not cognate with prepositions in Chuj, but it is related to the benefactive/locative preposition in Q'anjob'al, *b'ay*, which is a Q'anjob'alan innovation.

(1) Tojol-ab'al (Curiel 2007: 57, my translation)
kan-ø-ta *lap-an* ***b'a*** *bigro* ***b'a*** *y-olom* ja= ts'i' =i
remain-B3-already put.on-POS PREP glass PREP A3-head DET= dog =TOP
"The dog remained stuck in the jar up to his head."

(2) Q'anjob'al (Mateo Toledo 2008: 60)
Max-ø *w-aq'* *te* *on* ***b'ay*** *cham* *winaq*
COMPL-B3SG A1SG-give CLF avocado PREP CLF man
"I gave the avocado to the old man."

The other preposition in Tojol-ab'al, however, is shared uniquely with Tseltal. This is the comitative preposition *sok*, which is almost certainly a frozen form of what was formerly a relational noun with a third-person possessive prefix (*s-* in both

8. The suffix *-on* is optionally used in Tsotsil in agent extraction structures. Aissen (1999) argues that this has functionally become an inverse marker in Tsotsil. It is unclear if other related languages with this optional structure have also developed an inverse function.

9. Robertson (1977: 112) says that Tsotsil uses an *-on* suffix in agent relative clauses in the dialect of Zinacantan, but not in Chamula. Haviland (1988: 342–360), which is based on the dialect of Zinacantan, does not show any examples of the use of the *-on* suffix in this context. This may indicate that it can be optionally used in Zinacantan.

Tojol-ab'al and Tseltal). In fact, the contexts in which it occurs can all be analyzed as third person. Polian (2004: 54) proposes a Tseltalan etymology for this form based on the Tseltalan root *joy "to accompany" and the optative suffix *-uk*: *s-joy-uk > *sok* (cf Kaufman 1971: 118). It is also possible that it derives from the root *-mok* (*s-mok > sok*), since in Tojol-ab'al the form *s-mok* varies with *sok*, and in non-third persons *-mok*, with a possessive prefix, is the form of the comitative for first and second person.

(3) Tojol-ab'al (Curiel 2007: 83, 54; my translation)
yaj =ni nupan-y-on ***sok*** *jun sapatista ja= k-e'n =i*
since =EMPH marry-INTR-B1 with one zapatista DET= A3-PRON =TOP
lek lek wan-ø 'ek'-el k-uj
well well ICP-B3 experience-INF A1-RN:AGEN
"Since I got married with a Zapatista, things are going pretty well."

(4) *te'y-a j-**mok**-tikon =a*
DEM+EXT-B2 A1-with-1PL.EXCL =DIST
"Here you are, with us."

(5) Tseltal (Polian 2004: 53, 2013: 86; my translation)
Jich k'ax-Ø j-wokol-tik ***sok*** *te j-me' j-tat-tik namey*
thus pass-B3 A1-suffering-PL with DET A1-mother A1-father-PL long.ago
"Thus we suffered with our parents in the olden days."

(6) *At'ej-**o**n ta j-lum**a**l sok ja'-**a**t.*
work-B1[COMPL] PREP A1-town with FOC-B2
"I worked in my town with you."'

4.2.8 *Negation*

The system of negation is formally somewhat different in Tojol-ab'al than in other Mayan languages. It appears to have been innovated by recruiting the interrogative/hypothetical mode particle *mi* to be the generic marker of negation. Both Tseltal and Chuj have reflexes of the Proto-Mayan negation *ma' for most forms of negation. In addition, Chuj has a derived negative form for nonverbal predicates, *maŋ*. All three languages also have a negative existential: Chuj has *ma'ay*; Tseltal has *mayuk*. Interestingly, Tojol-ab'al has two variants of this morpheme, one parallel to Chuj, *me'ey* (from *mi-'ay), and one parallel to Tseltal, *meyuk* (from *mi-'ay-uk). In this case, the form *ma'ay/me'ey* is the more conservative form, and *me'yuk/mayuk* is an innovative form derived from *ma'ay/me'ey with the addition of the irrealis suffix *-uk*.

Table 14. Negation in Tseltal (Polian 2004:105, 210, 212), Tsotsil (Aissen 1987: 6, 12), Tojol-ab'al (Curiel 2007: 78) and Chuj (Maxwell 1982: 179)

	Tseltal	**Tsotsil**	**Tojol-ab'al**	**Chuj**
Verbal	*ma*	*mu*	*mi*	*ma … -ok*[10] *(-laj)*
Nonverbal	*ma … -uk*	*mu … -uk*	*mi*	*maŋ … -ok (-laj)*
Existential	*mayuk*	*mu'yuk, ch'abal*	*me'yuk, me'ey*	*ma'ay*
Imperative	*ma*	*mu*	*mok*	*?*

Negation in Tojol-ab'al appears to be consistent with both Tseltal and Chuj negation. The negation particle *mi* and the loss of the *-uk* or *-ok* irrealis marker in nonverbal negation are innovations that set Tojol-ab'al apart from both Chuj and Tseltal/Tsotsil. In the case of existential negation, forms from both languages appear to have been retained.

4.3 Lexicon

Another area that merits examination is the lexicon. Several documented mixed languages, such as Media Lengua, have a lexicon that is primarily derived from one source, while grammatical morphology is predominately from another. In the following sections, we will examine the 'core' or 'basic' vocabulary, based on Swadesh's famous lists, as well as a lexicostatistical comparison and network analysis of a more extended lexicon of approximately 1200 lexical items, as reported in Adell & Law (2015).

4.3.1 *Basic vocabulary*

Clearly, from the preceding review, we would be hard pressed to identify a single source for the grammatical morphology of Tojol-ab'al. The vocabulary is similarly problematic. In his investigation of the classification of Tojol-ab'al, Schumann (1981) compared the Swadesh 100 list of basic vocabulary for Tojol ab'al, Chuj, Tseltal and Tsotsil. The result was the following statistics:

Tojol-ab'al Basic Vocabulary Schumann (1981: 160–161)

Shared with Tseltal:	65%
Shared with Tsotsil:	60%
Shared with Chuj:	69%

Based on these statistics, Schumann concluded (1981:164):

10. The suffix *-ok* does not appear when the verb is transitive (Maxwell 1982: 178).

> It is clear that lexical similarities arise with much greater frequency between Tojol-ab'al and Chuj, though without becoming as apparent as those of Tseltal in relation to Tsotsil; in any case the lexical similarities of Tojo-ab'al point towards Chuj and not towards any other language of the Mayan family.[11]

While it is true that Chuj has a higher percentage of shared lexicon with Tojol-ab'al than Tsotsil and Tseltal do, it does not seem accurate to say that these shared lexical items "arise with much greater frequency between Tojol-ab'al and Chuj". What these statistics seem to show, to the contrary, is that the lexical similarities between Tojol-ab'al and Chuj, Tseltal and Tsotsil, are remarkably parallel, with Chuj only sharing four more terms from the basic word list than Tseltal does. The results are remarkably similar to those obtained from looking at similarities in the grammar.

My own comparison of the Swadesh 100 word list for Chuj, Tojol-ab'al and Tseltal agrees with Schumann's statistics, and allows us to bring out additional detail. These results are summarized in Figure 2 below.

One thing that the lexical data show clearly is the complexity introduced by the fact that these languages are genetically related. Half of the items in the word list are shared by all three languages. Chuj and Tseltal are also fairly similar with respect to the number of additional similarities each contributes. Chuj has more similarities with Tojol-ab'al in the basic vocabulary than Tseltal does, but the difference (68% vs. 65%) is hard to interpret as anything like the prototypical massive relexification that is often associated with language mixing. Neither language is obviously the 'lexifier' language.

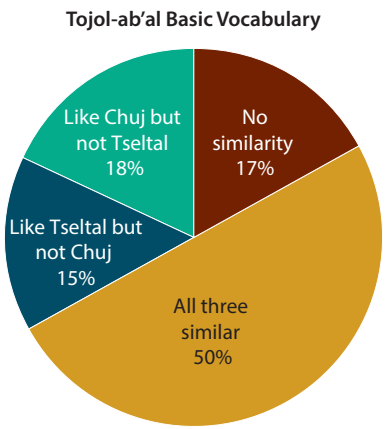

Figure 2. Tojol-ab'al basic vocabulary in comparison with Tseltal and Chuj

11. My translation of "Resulta evidente que las aproximaciones léxicas se dan con mucha mayor frecuencia entre el tojolabal y el chuj, aunque sin llegar a ser tan manifiestas como lo son las del tzeltal en relación a las del tsotsil; de todas formas, las aproximaciones léxicas del tojolabal apuntan hacia el chuj y no hacia otra lengua de la familia maya".

4.3.2 *Extended vocabulary*

If the basic vocabulary of Tojol-ab'al is slightly more similar to Chuj, it appears that the extended vocabulary is somewhat more similar to Tseltal. Adell & Law (2015) present a detailed analysis of a corpus of approximately 1200 lexical items for all of the languages of the Q'anjob'alan and Ch'olan-Tseltalan subgroups. The lexicostatistical comparison of each language, in Table 15 below, highlights Tojol-ab'al's uniquely ambiguous lexical profile.

The chart shows that, overall, Tojol-ab'al is more lexically similar to Tseltal (41%) than any other Mayan language. Adell & Law (2015) found this to be true for the dataset as a whole, as well as for a variety of different subsets of the data, including different parts of speech and different semantic fields. Chuj, on the other hand, at 28% similarity with Tojol-ab'al, is fifth in terms of overall similarity with Tojol-ab'al, after Tseltal, Tsotsil, Chol and Cholti, but it fits comfortably in degree of similarity with the other Q'anjob'alan languages. Tsotsil, a close relative of Tseltal, and Chol, which has had extensive contact with Tseltal, are likely to be high on the list because of their relationship with Tseltal rather than with Tojol-ab'al. Cholti, on the other hand, was a highly influential language in the Lowland Mayan language contact area. The lexical similarity between Cholti and Tojol-ab'al may well be the result of direct contact and lexical borrowing into Tojol-ab'al from Cholti (Adell & Law 2015).

Table 15. Percentage of items in each language shared with other languages[12]

	CHT	CHR	CHN	CHL	TSE	TSO	TOJ	CHJ	QAN	AKA	POP	MCH
CHT	–	54%	47%	49%	36%	34%	37%	26%	22%	21%	20%	20%
CHR	54%	–	34%	34%	25%	23%	23%	20%	17%	16%	16%	15%
CHN	47%	34%	–	43%	25%	24%	23%	18%	15%	15%	15%	16%
CHL	49%	34%	43%	–	39%	35%	33%	22%	20%	19%	18%	17%
TSE	36%	25%	25%	39%	–	67%	41%	24%	22%	21%	20%	19%
TSO	34%	23%	24%	35%	67%	–	37%	24%	22%	20%	20%	20%
TOJ	37%	23%	23%	33%	41%	37%	–	28%	23%	22%	23%	21%
CHJ	26%	20%	18%	22%	24%	24%	28%	–	46%	43%	36%	24%
QAN	22%	17%	15%	20%	22%	22%	23%	46%	–	63%	51%	22%
AKA	21%	16%	15%	19%	21%	20%	22%	43%	63%	–	50%	22%
POP	20%	16%	15%	18%	20%	20%	23%	36%	51%	50%	–	24%
MCH	20%	15%	16%	17%	19%	20%	21%	24%	22%	22%	24%	–

12. Abbreviations for languages, which are used here and throughout the paper, are as follows: CHT=Cholti, CHR=Ch'orti', CHN=Chontal, CHL=Chol, TSE=Tseltal, TSO=Tsotsil, TOJ=Tojol-ab'al, CHJ=Chuj, QAN=Q'anjob'al, AKA=Akateko, POP=Popti', MCH= Mocho'.

4.3.3 *Discussion of lexical similarity*

It is important to note that, while Tojol-ab'al is clearly more lexically similar to Tseltal than it is to Chuj, it stands out among the languages in Table 15 because it does not clearly fit, in terms of percent similarity, into one of the three subgroups (Ch'olan, Tseltalan and Q'anjob'alan). Other than the highly divergent Mocho', it is the only language in the sample that does not have greater than 45% similarity with any other language in the extended lexicon. The contrast between the greater similarity with Chuj in the basic vocabulary and the greater similarity with Tseltal in the extended vocabulary is also noteworthy. However, the amount of difference, in each case, is not enough to support anything like wholesale relexification.

The fairly ambiguous and obviously mixed results of the analysis of the Tojol-ab'al lexicon lead us to open once again the more theoretical question of how to determine if a language is 'mixed' and whether a mixed language is categorically different from cases of heavy, or even extremely heavy, borrowing. The general sense is that there is something essentially different about mixed languages. Some even argue that this difference is easily quantifiable. For example, in the introduction to the first edited volume dedicated entirely to mixed languages, Bakker & Mous (1994: 5) hypothesized that "extreme borrowing never exceeds roughly 45% of the lexicon, whereas in some mixed languages the proportion of 'foreign' lexical elements is close to or over 90%". They go on to note that "there do not seem to be languages with a proportion of borrowed items between 45% and 90%, so that there is no continuum between languages with heavy borrowing and mixed languages" (Bakker & Mous 1994: 5).

The Tojol-ab'al data definitely establish that this kind of wholesale relexification (>90% similarity) did not take place. If we include all similarity and do not attempt to distinguish inheritance from possible borrowing, the basic vocabulary figures (59.7% and 65.9%) fall firmly in this 'no man's land' between 45% and 90%. The extended vocabulary percentages, even including inherited similarities, at highest (41%), are still within the range Bakker and Mous give for heavy borrowing, even if all of that similarity were due to borrowing, which it is not. If we exclude the similarities shared by Tojol-ab'al, Chuj and Tseltal (presumably all inherited similarities), the numbers in basic vocabulary (11.3% and 17.5%) barely qualify as heavy borrowing, and if we exclude from these numbers basic vocabulary items that are likely retentions from Proto-Mayan, the numbers drop even further. Only three of the eleven similarities in basic vocabulary shared by Tsotsil and Tojol-ab'al, but not Chuj, are clearly innovative forms, and only seven of the seventeen forms that Chuj and Tojol-ab'al share are innovative.

The claim that the difference between language mixing and heavy borrowing is one of kind and not a matter of a continuum has been questioned elsewhere (Stolz 2003). The Tojol-ab'al data do not provide a direct refutation of this claim,

but they do problematize the validity of such lexical statistics as a definitional criterion for mixed languages. Because the criterion cannot readily apply to related languages, as the Tojol-ab'al data have shown, we are forced either to accept genetic unrelatedness as another defining requirement of mixed languages or to seek other means for defining mixed languages as a distinct category.

5. Historical context of Tojol-ab'al

In the case of Tojol-ab'al, we can imagine a variety of hypothetical historical scenarios that could lead to such a mixture. For example, a community of Tseltal speakers may have become bilingual in Chuj, or a community of Chuj speakers may have immigrated into the Tseltal area and become bilingual in Tseltal, or speakers of both languages may have relocated to a single, mixed-origin community. Any of these possibilities could logically have led to the language mixing in Tojol-ab'al. In order to help evaluate and correctly interpret the conflicting and confusing historical linguistic data of Tojol-ab'al, it would be valuable to have clear non-linguistic evidence regarding the history of the Tojol ab'al people. Unfortunately, the available historical and ethnographic record tapers off quickly into the murkiness of conjecture.

The lack of a consistent name for the language, or the speakers of that language, exacerbates the problem. Campbell (1988, ch. 3) summarizes the history of references to the language. He notes that the earliest secure linguistic material written in the Tojol-ab'al language, dating to 1775, refers to the language as Chanabal. This source is a liturgical and confessional text written by the Dominican friar Domingo Paz titled *Confesionario y Doctrina Christiana en lengua Chanabal de Comitán y Tachinulla en las Chiapas* (Published in Ruz 1989: 33–73, along with a 'Chanabal' confessionary from 1813). Here we have a good example of the confusion surrounding language names in Chiapas, since within this manuscript *en lengua Chanabal* we have texts in Tojol-ab'al (from Comitán), Mocho' (from Tachinulla) and (the now extinct) Cotoque, or Chicomuseltek. Clearly the author of these texts was aware that they were different languages, but precision in use of language names does not seem to have been particularly important.

Tojol-ab'al is also referred to in the later Colonial literature as Chaneabal, Chañabal, Trokek, Casdal, Jojlabal, Jocolabal (Campbell 1988: 153) or Comiteco, after the municipal head where it was reported to be spoken, Comitán. Nearly a century prior to the writing of the Chanabal manuscript, in the year 1686, Fray Matías Martínez was granted license to preach in the Chañabal language, which presupposes its existence in the region (Contreras Garcia 2001: 83). Assuming this is the same Chanabal of Fray Paz' confessional, we can trace a Tojol-ab'al presence in

Chiapas to at least the middle of the 17th century. However, assuming that attempts to link the Tojol-ab'al with other groups that are mentioned in ethnohistorical records (the Coxoh, for example; see G. Lenkersdorf 1986: 46–47 and Campbell & Gardner 1988 on this question) are unfounded, it is far from clear where the Tojol-ab'al were during the first century and a half of the Spanish conquest.

One proposal (Ruz 1981: 43; Contreras Garcia 2001: 83) is that, upon the arrival of the Spaniards, the Tojol-ab'al, like other indigenous groups in the region, including for example the Manché and Lacandon Chol, fled into the jungles and mountains and stayed there, either below the radar of the Spanish, so to speak, or else simply folded into the larger group of 'wild Indians', of various ethnicities, that were all simply called 'lacandones'. Sometime in the mid 17th century (prior to the 1686 mention of Chanabal), they would have come out of the mountains and yielded to Spanish rule. If this is the case, there is no historical record of it, in spite of the fact that such an event would have been considered quite a coup by the Spanish missionaries, who were constantly trying to get the rebel Indians out of the jungles and mountains and into easily controlled towns.

Another possibility is that the Tojol-ab'al (Chanabal) were originally from near the Chuj region in Guatemala and that they did not migrate to Chiapas until the mid 17th century. Such a migration into the heart of a region already under the control of the Spanish also seems unlikely to go without mention by contemporary missionaries and historians. The historical details, it seems, are beyond our reach.

This line of inquiry, however, brings us to another important question about the history of the Tojol-ab'al. Whether they happened before or after the Spanish conquest, the striking similarities between Tojol-ab'al and Chuj make it clear that these languages have extensive shared history. Whether this was in the form of Tseltalan language speakers interacting with Chuj or in the form of speakers of a Chujean language engaging in intensive interaction with speakers of Tseltal, a historical connection with Chuj speakers is unavoidably necessary to explain the form of the language. Campbell & Gardner (1988) and, in the same volume on the linguistics of Southeastern Chiapas, Adams (1988) describe ethnographic evidence of the same historical connection implied by the linguistic structure of Tojol-ab'al. Traditionally (though it is unclear if this tradition continues today), the Tojol-ab'al participated in a yearly pilgrimage to certain salt water springs near San Mateo Ixtatán, in the department of Huehuetenango, Guatemala, a Chuj-speaking region four days' journey from Trinitaria, the starting point of the *romería*. According to some Chuj traditions, the Tojol-ab'ales used to live near Ixtatán and attempted to take over the salt mines in that area. The Chuj are said to have used the help of cougars to scare the Tojol-ab'ales away (Campbell 1988: 160–161).

However, a historical connection with the Tseltal is also apparent in this pilgrimage, since the necessary prayers for the journey are all in Tseltal and a Tseltal

speaker is generally hired to pray during the pilgrimage (Adams 1988: 184). To further confuse the matter, the same group of Tojol-ab'al, again led by Tseltal *rezadores*, engages in another yearly pilgrimage to Oxchuc, in the heart of Tseltal country, a three-day journey from Trinitaria. Clearly, the ethnographic description of these religious practices, like the language itself, is a testament to a history of strong connections with both the Chuj and the Tseltal.

The above-mentioned explanation of the Chuj/Tojol-ab'al relationship provides anecdotal evidence that the Tojol-ab'al migrated at some point, given the degree of linguistic difference between Tojol-ab'al and other languages, most likely some time prior to the Conquest, from the Chuj region. This would be consistent with a scenario in which a group of Chujean speakers left their homeland and moved among the Tseltal. However, it is important to emphasize that, even if we accept the scenario in which a group of Chuj or Chujean speaking people migrated into what had been Tseltal speaking region and established strong ties with them, the problem of the genetic affiliation of Tojol-ab'al does not go away. The metaphor of genetic descent in language cannot erase the fact that language is not intrinsically tied to biology. We could just as easily speak of a group of Chuj speakers shifting to Tseltal, but a Tseltal full of Chujean features, as we could a group of Chuj speakers maintaining their language but replacing large portions of the inflectional morphology, syntax and lexicon with features from Tseltal.

6. Tojol-ab'al in the context of Lowland language contact phenomena

Campbell (1988: 154), in a discussion of the genetic classification of Tojol-ab'al, speculates that Tojol-ab'al may be more appropriately placed with Chuj than with Tseltal-Tsotsil because "the similarities shared with Tzeltal-Tzotzil seem more easily explicable as borrowings, due to diffusion, while the similarities shared with Chuj and other Kanjobalan languages seem much more arbitrary, interrelated, and less likely to be due to contact". He goes on to argue that, "Tojolabal (as well as Tzeltal-Tzotzil) received much influence from Cholan, the principal language of the Classic Maya civilization". This mutual influence from a single prestige language may, he argues, account for many of the similarities between the languages.

The additional evidence now available to us, as discussed above, establishes that Campbell was right to highlight larger regional patterns of contact in the shaping of Tojol-ab'al. This is particularly evident in the phonology of Tojol-ab'al. At least five areally diffused phonological changes have made Tojol ab'al phonology more similar to Tseltal and Tsotsil, as well as other Lowland and non-Lowland languages, one of which (the merging of */ŋ/* and */n/*) made Tojol-ab'al different from Chuj. At least one Lowland sound change, the introduction of the new contrastive

sound *p'*, made Tojol-ab'al less like Tseltal and Tsotsil, since these languages participated in that areally spread innovation while Tojol-ab'al did not. Evidence from the extended lexicon also brings to the fore the issue of contact. Tojol-ab'al shares high levels of similarity with the Ch'olan language Cholti, a similarity that Adell & Law (2015) interpret as the result of extensive contact with that Lowland language. Tojol-ab'al is also lexically similar to Chol, though this is arguably similarity mediated by Tseltal.

Regional contact effects are also plausible for the grammatical morphology discussed, but, particularly in this case, regional contact does not explain the extraordinarily high level of features from both Tseltal and Chuj. Of the 22 grammatical features surveyed above that are shared by Tojol-ab'al and Tseltal and not Chuj, 8 are features that are shared with other (Lowland) languages as well, demonstrably because of contact after the breakup of their common language. All of these features are also found in Tseltal, and so they could have been acquired in Tojol-ab'al through Tseltal. Two of the innovations that Tojol-ab'al shares with Chuj (the progressive auxiliary *wan* and the use of the nominal plural *-e'* for humans) are also likely areally diffused innovations that have impacted much more than just Tojol-ab'al and Chuj. Once again, however, as these are features also found in Chuj, the possibility remains that they are present in Tojol-ab'al because they are present in its source language, Chuj. Thus, while some areal features in Tojol-ab'al may be the result of areal influence on the language, the uniquely high degree of linguistic hybridity in Tojol-ab'al makes areal diffusion explanatorily insufficient.

7. Code-switching, compartmentalization and processes

The question of whether or not Tojol-ab'al can be labeled a mixed language should not distract attention from the underlying question that gives importance to the label: Are the historical processes that led to the current state of the Tojol-ab'al language clearly different from those that led to the formation of other 'mixed languages' or, conversely, to the processes of language change in 'normal' transmission? The processes through which mixed languages are formed have received some attention in the literature, though often based on only slightly more ethnographic information than we have for Tojol-ab'al. For example, scholars have debated the role different social factors – ethnic and community identity, or intermarriage between speakers of each language – may play in the creation of mixed languages (Golovko 2003; Bakker 1997). The role of code-switching in the creation of mixed languages has also been a matter of debate (Myers-Scotton 2002, 2003; Auer 1999). And recent work on emerging mixed languages in Australia gives something

approaching direct evidence of mixed languages emerging in the context of pervasive code-switching (McConvell & Meakins 2005; O'Shannessy 2005, 2012, 2015).

As outlined in §5, direct evidence for the process through which Tojol-ab'al arose is extremely limited. Historical records do not provide a clear picture of when Tojol-ab'al began to be spoken, and the modern day linguistic context of Tojol-ab'al is dramatically different because of the pervasiveness of Spanish. Perhaps the best we can say is that what little we know about the social context of Tojol-ab'al over the last 500 years or so does not refute the scenario presented here. The notes in §5 about ritual activities that suggest a dual orientation towards both Chuj and Tseltal ritual space is consistent with the hybrid group identity that has been argued to be key in the emergence of mixed languages.

The fact that both of Tojol-ab'al's donor languages came from the same linguistic ancestor is not merely a definitional desideratum for mixed languages. I argue here that it is a crucial fact in order to understand why Tojol-ab'al is so etymologically mixed. In Tojol-ab'al, there does not appear to be a handy division in terms of one grammatical subsystem or another. Rather, Tojol-ab'al displays a general mixture of features at every level that I have investigated here. It seems problematic to exclude a language from the category of 'mixed languages' because it is too mixed, but at the same time the difference in mixture clearly bespeaks very different circumstances for the emergence of that mixture. I argue that the difference can be understood by looking at patterns of code-switching involving related languages.

In Muysken's (2000) typology of code-switching patterns, he notes that the two more widely acknowledged code-switching types – insertion and alternation – do not fully capture the empirical facts. A third type of code-switching, he argues, is necessary to explain patterns found in code-switching between closely related or highly similar languages. He calls this third type 'congruent lexicalization'. According to Muysken, congruent lexicalization "is akin to language variation and style shifting: switching is grammatically unconstrained and characterizable in terms of alternative lexical insertions …. Some cases of word-internal mixing can be viewed as congruent lexicalization." (Muysken 2000: 221). "In congruent lexicalization", he says, "both languages contribute to the grammatical structure of the sentence, which in many cases, is fully shared by the languages involved. The vocabulary comes from two or more different languages, but may also be shared" (Muysken 2000: 127). This implies, as seems to be the case for Tojol-ab'al, that there is not a single INFL or matrix language in congruent lexicalization, but that each language is an equal contributor. Because shared grammatical structure is important for congruent lexicalization to happen, there is a relationship between congruent lexicalization and convergence: "Linguistic convergence feeds

into congruent lexicalization, and the two processes may reinforce each other" (Muysken 2000: 221).

The characteristics that Muysken highlights for identifying congruent lexicalization code-switching include a general linear and structural equivalence between the varieties being mixed, the use of multi-constituent code-switching and non-constituent mixing (also known as 'ragged mixing'; see Poplack 1980) and switches involving all lexical categories, including function words, mixed collocations and relatively numerous homophonous diamorphs (interlingually identified morphemes; see Haugen 1956: 46). For a more comprehensive list and comparison to insertion and alternation code-switching, see Muysken (2000: 229–231). In summary, congruent lexicalization is a pattern in which code-switches are common at all points within and across clauses and even within words. To repeat Muysken's somewhat tongue-in-cheek assessment, "Basically, anything goes in congruent lexicalization" (Muysken 2000: 128). This is a fair assessment of the pattern of mixing described here for Tojol-ab'al.

The fact that an observed pattern of code-switching that is typically restricted to situations involving related languages would produce utterances mixed at the level that we find for Tojol-ab'al is consistent with the proposal that mixed languages arise through the crystallization of community-wide code-switching patterns. In fact, the mounting evidence that this hypothesis about the emergence of mixed languages is on the right track (McConvell & Meakins 2005; McConvell 2008; O'Shannessy 2005, 2012, 2015) encourages scholars to search mixed languages for the same kinds of patterns documented in code-switching. Tojol-ab'al provides just such a case: it not only displays the pattern of mixture typical of congruent lexicalization but also meets the conditions that have been argued to license this type of code-switching, i.e., code-switching between closely related languages.

8. Conclusion

Is Tojol-ab'al a mixed language? The grab-bag mixture of elements apparently from different branches of the Mayan language family seems to argue in favor of such a label. The description of the empirical and comparative facts about Tojol-ab'al presented here is, in and of itself, of interest for exploring the ways in which languages change in contact. Under several prominent definitional criteria in the literature, Tojol-ab'al does not fit the 'structural prototype' (Matras 2003) of a mixed language: it is not neatly compartmentalized according to the etymological source of linguistic material in different linguistic subsystems, and, at least arguably, we can trace all of its linguistic subsystems back to a common source (Proto-Mayan).

However, I argue that languages like Tojol-ab'al invite us to reconsider current definitions of mixed languages. The typologies proposed for mixed languages to date have largely ignored mixing between related languages, as well as, for the most part, languages in known *Sprachbünde*. These more complex situations, like that of Tojol-ab'al, are precisely the contexts in which the typologies proposed to date would be most likely to break down. Regional and local attitudes and social norms regarding language mixing and language and identity can have a profound role in constraining and motivating different linguistic creations. The fact that Tojol-ab'al emerged in a regional context in which areal diffusion, including grammatical borrowing, was widespread, is surely relevant to understand its development. Also relevant is the fact that the languages that contributed to the current shape of Tojol-ab'al, Chuj and Tseltal are genetically related. Inherited similarity has been shown to be relevant both to processes of language contact (Epps et al. 2013; Law 2014) and patterns of code-switching (Muysken 2000). If Tojol-ab'al, like other mixed languages, arose out of patterns of community code-switching, the thorough mixture of features from contributing languages is reasonable given the relatedness of those contributing languages. Indeed, the case of Tojol-ab'al suggests that, if patterns of code-switching are indeed relevant in the emergence of mixed languages, we may expect to see more mixtures of this type.

Acknowledgements

The research presented here would have been much poorer were it not for generous support from a University of Texas summer research assignment and a Javits Fellowship from the U.S. Department of Education. I also gratefully acknowledge the valuable feedback I received, at various stages, from Roberto Zavala Maldonado, José Gomez Cruz, Nora England, John Robertson, Lyle Campbell, Peter Bakker, Pattie Epps, and the editors of *Diachronica*. All errors and misrepresentations that remain are, regretfully, my own.

Abbreviations

1	first person	FOC	focus
2	second person	ICP	incompletive
3	third person	INC	inclusive
A	'set A' (ergative) person marker	INF	infinitive
ABS	absolutive	INTR	intransitive
AGEN	agent	NVP	non-verbal predicate
B	'set B' (absolutive) person marker	PL	plural
CLF	classifier	POS	positional

COMPL	completive
DEM	demonstrative
DET	determiner
DIST	distal
EMPH	emphatic
EXCL	exclusive
EXT	existential
PREP	preposition
PRON	independent pronoun
RN	relational noun
S	singular
TOP	topic
TRANS	transitive

References

Adams, Walter Randolph. 1988. Religious practices of southeastern Chiapas and Tzeltal-Tojolabal interactions. In Lyle Campbell, *The linguistics of Southeast Chiapas, Mexico* (Papers of the New World Archaeological Foundation 50), 183–198. Provo, UT: New World Archaeological Foundation, Brigham Young University.

Adell, Eric & Danny Law. 2015. Lexical similarity and the history of Tojol-ab'al. Paper presented at CILLA 7, Austin, Texas, October 29–31, 2015.

Aissen, Judith. 1987. *Tzotzil clause structure*. Boston: D. Reidel Publishing Company. doi:10.1007/978-94-009-3741-3

Aissen, Judith. 1999. Agent focus and inverse in Tzotzil. *Language* 75(3). 451–485. doi:10.2307/417057

Aissen, Judith & Roberto Zavala Maldonado (eds.). 2010. *La predicación secundaria en lenguas de Mesoamérica*. México: CIESAS.

Auer, Peter. 1999. From code-switching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism* 3(4). 309–332. doi:10.1177/13670069990030040101

Bakker, Peter & Maarten Mous (eds.). 1994. *Mixed languages: 15 case studies in language intertwining*. Amsterdam: Instituut voor Fundamenteel Onderzoek naar Taal en Taalgebruik.

Bakker, Peter. 1997. *A language of our own: The genesis of Michif, the mixed Cree-French language of the Canadian Métis*. Oxford: Oxford University Press.

Bakker, Peter. 2003. Mixed languages as autonomous systems. In Yaron Matras & Peter Bakker, (eds.), *The mixed language debate: Theoretical and empirical advances*, 107–150. Berlin: Mouton de Gruyter.

Barrett, Rusty. 2002. The Huehuetenango sprachbund and Mayan language standardization in Guatemala. In Mary Andronis, Erin Debenport, Anne Pycha & Keiko Yoshimura (eds.), *Proceedings of the 38th Chicago Linguistics Society: The panels*, 309–318. Chicago: Chicago Linguistics Society.

Bricker, Victoria. 1977. *Pronominal inflection in the Mayan languages* (Middle American Research Institute Occasional Papers 1). New Orleans: MARI, Tulane University.

Campbell, Lyle & Brant Gardner. 1988. Coxoh. In Lyle Campbell (ed.), *The linguistics of southeast Chiapas, Mexico* (Papers of the New World Archaeological Foundation 50), 315–338. Provo, UT: New World Archaeological Foundation, Brigham Young University.

Campbell, Lyle. 1988. *The linguistics of southeast Chiapas, Mexico* (Papers of the New World Archaeological Foundation 50). Provo, UT: New World Archaeological Foundation, Brigham Young University.

Campbell, Lyle. 1993. On proposed universals of grammatical borrowing. In Henk Aertsen & Robert Jeffers (eds.), *Historical linguistics 1989*, 91–110. Amsterdam: John Benjamins. doi:10.1075/cilt.106.08cam

Campbell, Lyle, Terrence Kaufman & Thomas C. Smith-Stark. 1986. Mesoamerica as a linguistic area. *Language* 62. 530–570. doi:10.1353/lan.1986.0105

Contreras García, Irma. 2001. *Las etnias del estado de Chiapas*. México: UNAM-IIB.

Curiel Ramírez del Prado, Alejandro. 2007. *Estructura de la información, enclíticos y configuración sintáctica en Tojol'ab'al*. MA thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social, México.

Dakin, Karen. 1988. Las lenguas kanjobalanas, polémicas clasificatorias. In Aurora M. Ocampo (ed.), *Studia Humanitas: Homenaje a Rubén Bonifaz Nuño*, 111–133. México: Universidad Nacional Autónoma de México.

DeGraff, Michel. 2003. Against creole exceptionalism. *Language* 79(2). 391–410. doi:10.1353/lan.2003.0114

DeGraff, Michel. 2004. Against creole exceptionalism (redux). *Language* 80(4). 834–839. doi:10.1353/lan.2004.0178

Dreyfuss, G. R. & Djoehana Oka. 1979. A new kind of language hybrid? *Pacific Linguistics* (Series A) 57. 247–274.

Epps, Patience, John Huehnergard & Na'ama Pat-El. 2013. Introduction: Contact among genetically related languages. *Journal of Language Contact* 6. 209–219. doi:10.1163/19552629-00602001

Golovko, Evgeniy V. 1994. Mednyj Aleut or Copper Island Aleut: an Aleut-Russian mixed language. In Peter Bakker & Maarten Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, 113–121. Amsterdam: Instituut voor Fundamenteel Onderzoek naar Taal en Taalgebruik.

Golovko, Evgeniy V. 2003. Language contact and group identity: The role of 'folk' linguistic engineering. In Yaron Matras & Peter Bakker (eds.), *The mixed language debate: Theoretical and empirical advances*. 177–207. Berlin: Mouton de Gruyter.

Haugen, Einar. 1956. *Bilingualism in the Americas*. American Dialect Society. University, Ala.: University of Alabama Press.

Haviland, John B. 1988. It's my own invention: A comparative grammatical sketch of Colonial Tsotsil. In Robert Laughlin, *The great Tzotzil dictionary of Santo Domingo Zinacantan* (Smithsonian Contributions to Anthropology Number 31), 79–121. Washington: Smithsonian Institution Press.

Hopkins, Nicholas. 2006. Review of *The linguistics of Maya writing*, Søren Wichmann (ed.). *Anthropological Linguistics* 48(4). 404–412.

Instituto Nacional de Estadística y Geografía (INEGI). 2015. Encuesta Intercensal. http://cuentame.inegi.org.mx/poblacion/lindigena.aspx (accessed Dec. 26, 2016).

Justeson, John S., William M. Norman, Lyle Campbell & Terrence Kaufman. 1985. *The foreign impact on Lowland Mayan language and script*. New Orleans: Middle American Research Institute, Tulane University.

Kaufman, Terrence. 1969. Some recent hypotheses on Mayan diversification (Working Paper 26). *Language Behavior Research Laboratory*, University of California, Berkeley.

Kaufman, Terrence. 1971. *Tzeltal phonology and morphology*. Berkeley: University of California Press.

Kaufman, Terrence. 1976. Archaeological and linguistic correlations in Mayaland and associated areas of Meso-America. *World Archaeology* 8(1). 101–118. doi:10.1080/00438243.1976.9979655

Kaufman, Terrence. 1990. Algunos rasgos estructurales de los idiomas Mayances con referencia especial al K'iche'. In Nora C. England & Stephen R. Elliott (eds.), *Lecturas sobre la lingüística Maya*, 59–114. La Antigua Guatemala: CIRMA.

Kaufman, Terrence & William M. Norman. 1984. An outline of Proto-Cholan phonology, morphology and vocabulary. In John Justeson & Lyle Campbell (eds.), *Phoneticism in Mayan hieroglyphic writing* (Institute for Mesoamerican Studies 9), 77–166. Albany: State University of New York, Albany.

Law, Danny. 2009. Pronominal borrowing among the Maya. *Diachronica* 26(2). 214–252. doi:10.1075/dia.26.2.03law

Law, Danny. 2011. *Linguistic inheritance, social difference, and the last two thousand years of contact among Lowland Mayan languages*. Ph.D. diss., University of Texas at Austin.

Law, Danny. 2013. Mayan historical linguistics in a new age. *Language and Linguistics Compass* 7(3). 141–156. doi:10.1111/lnc3.12012

Law, Danny. 2014. *Language contact, inherited similarity and social difference: The story of linguistic interaction in the Maya Lowlands*. Amsterdam: John Benjamins. doi:10.1075/cilt.328

Law, Danny, John Robertson, Stephen Houston, Marc Zender & David Stuart. 2014. Areal shifts in Classic Mayan phonology. *Ancient Mesoamerica* 25(2). 357–366. doi:10.1017/S0956536114000273

Lenkersdorf, Gudrun. 1986. Contribuciones a la historia colonial de los Tojolabales. In Mario Humberto Ruz (ed.), *Los legítimos hombres: Aproximación antropológica al grupo tojolabal*, vol. 4, 13–102. Mexico: UNAM.

López Ixcoy, Candelaria Dominga (Saqijix). 1997. *Ri Ukemiik ri K'ichee' Chii': Gramática K'ichee'*. Guatemala: Cholsamaj.

Mateo Toledo, Eladio. 2008. *The family of complex predicates in Q'anjob'al (Maya): Their syntax and meaning*. Ph.D. diss., University of Texas at Austin.

Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36. 281–331. doi:10.1515/ling.1998.36.2.281

Matras, Yaron. 2003. Mixed languages: Re-examining the structural prototype. In Yaron Matras & Peter Bakker (eds.), *The mixed language debate: Theoretical and empirical advances*, 151–175. Berlin: Mouton de Gruyter. doi:10.1515/9783110197242.151

Matras, Yaron & Peter Bakker (eds.) 2003. *The mixed language debate: Theoretical and empirical advances*. Berlin: Mouton de Gruyter. doi:10.1515/9783110197242

Maxwell, Judith. 1982. *How to talk to people who talk funny: The Chuj (Maya) solution*. Ph.D. diss., University of Chicago.

McConvell, Patrick. 2008. Mixed languages as outcomes of code-switching: Recent examples from Australia and their implications. *Journal of Language Contact* THEMA 2. 187–212. doi:10.1163/000000008792525327

McConvell, Patrick & Felicity Meakins. 2005. Gurindji Kriol: A mixed language emerges from code-switching. *Australian Journal of Linguistics* 25(1). 9–30. doi:10.1080/07268600500110456

Moravcsik, Edith. 1978. Universals of language contact. In Joseph Greenberg (ed.), *Universals of human language*, 94–122. Stanford: Stanford University Press.

Mous, Maarten. 1994. Ma'a or Mbugu. In Peter Bakker & Maarten Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, 175–200. Amsterdam: Instituut voor Fundamenteel Onderzoek naar Taal en Taalgebruik.

Muysken, Pieter. 1981. Halfway between Quechua and Spanish: The case for relexification. In Arnold Highfield & Albert Valdman (eds.), *Historicity and variation in creole studies*, 52–78. Ann Arbor: Karoma.

Muysken, Pieter. 1994. Media Lengua. In Peter Bakker & Maarten Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, 207–211. Amsterdam: Instituut voor Fundamenteel Onderzoek naar Taal en Taalgebruik.

Muysken, Pieter. 2000. *Bilingual Speech: A typology of code-mixing*. Cambridge: Cambridge University Press.

Myers-Scotton, Carol. 2000. What matters: The out of sight in mixed languages. *Bilingualism, Language and Cognition* 3. 119–21. doi:10.1017/S1366728900280218

Myers-Scotton, Carol. 2002. *Contact linguistics, bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198299530.001.0001

Myers-Scotton, Carol. 2003. What lies beneath: Split (mixed) languages as contact phenomena. In Yaron Matras and Peter Bakker (eds.), *The mixed language debate: Theoretical and empirical advances*, 73–106. Berlin: Mouton de Gruyter.

O'Shannessy, Carmel. 2005. Light Warlpiri: A new language. *Australian Journal of Linguistics* 25(1). 31–57. doi:10.1080/07268600500110472

O'Shannessy, Carmel. 2012. The role of code-switched input to children in the origin of a new mixed language. *Linguistics* 50(2). 305–340.

O'Shannessy, Carmel. 2015. Typological and social factors influencing a new mixed language, Light Warlpiri. In Gerald Stell & Kofi Yakpo (eds.), *Code-switching between structural and sociolinguistic perspectives*, 289–304. Berlin: De Gruyter.

Polian, Gilles. 2004. *Elements de grammaire du Tseltal*. Ph.D. diss., Université Paris III: Sorbonne Nouvelle.

Polian, Gilles. 2007. El sufijo *uk* y la diversificación de la modalidad *irrealis* en tseltal. *Conference on Indigenous Languages of Latin America 3*. 1–26. Austin, Texas: CILLA.

Poplack, Shana. 1980. Sometimes I'll start a sentence in Spanish y termino en español. *Linguistics* 18. 581–618. doi:10.1515/ling.1980.18.7-8.581

Robertson, John S. 1977. A proposed revision in Mayan subgrouping. *International Journal of American Linguistics* 43(2). 105–120. doi:10.1086/465466

Robertson, John S. 1992. *The history of tense/aspect/mood/voice in the Mayan verbal complex*. Austin: University of Texas Press.

Ruz, Mario Humberto. 1981. *Los legítimos hombres: Aproximación antropológica al grupo tojolabal*, vol. 1. Mexico: Centro de Estudios Mayas, UNAM.

Ruz, Mario Humberto (ed.). 1989. *Las lenguas del Chiapas colonial: manuscritos en la Biblioteca Nacional de Paris*. Mexico: Instituto de Investigaciones Filologicas, Seminario de Lenguas Indigenas, UNAM.

Schadeberg, Thilo C. 1994. KiMwani at the southern fringe of KiSwahili. In Peter Bakker & Maarten Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, 239–244. Amsterdam: Instituut voor Fundamenteel Onderzoek naar Taal en Taalgebruik.

Schumann, Otto G. 1981. La relación lingüística Chuj-Tojolabal. In Ruz, Mario Humberto (ed.), *Los legítimos hombres: Aproximación antropológica al grupo Tojolabal*, 129–169. Mexico: Centro de Estudios Mayas, UNAM.

Schumann, Otto G. 1983. Algunos aspectos de la relación Chuj-Tojolabal. In Lorenzo Ochoa & Thomas A Lee. Jr. (eds.), *Antropología e historia de los Mixe-Zoques y Mayas: Homenaje a Frans Blom*, 355–363. México: Universidad Autónoma de México; Provo, UT: Brigham Young University.

Smith-Stark, Thom. 1978. The Mayan antipassive: Some facts and fictions. In Nora C. England (ed.), *Papers in Mayan linguistics*, 169–187. Columbia, MO: University of Missouri.

Stolz, Thomas. 2003. Not quite the right mixture: Chamorro and Malti as candidates for the status of mixed language. In Yaron Matras & Peter Bakker (eds.), *The mixed language debate: Theoretical and empirical advances*, 271–314. Berlin: Mouton de Gruyter.

Thomason, Sarah. 2001. *Language contact: An introduction*. Edinburgh: University of Edinburgh Press. doi:10.1016/Bo-08-043076-7/03032-1

Thomason, Sarah. 2003. Social factors and linguistic processes in the emergence of stable mixed languages. In Yaron Matras & Peter Bakker (eds.), *The mixed language debate: Theoretical and empirical advances*, 21–40. Berlin: Mouton de Gruyter.

Thomason, Sarah & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Van Bree, Cor. 1994. The development of so-called Town Frisian. In Peter Bakker & Maarten Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, 69–82. Amsterdam: Instituut voor Fundamenteel Onderzoek naar Taal en Taalgebruik.

Wichmann, Søren. 2006. Mayan historical linguistics and epigraphy: A new synthesis. *Annual Review of Anthropology* 35. 279–294. doi:10.1146/annurev.anthro.35.081705.123257

Résumé

Les définitions des langues « entremêlées » ou « mixtes » proviennent presque toutes d'études de langues qui combinent des éléments provenant de sources génétiquement non apparentées. La langue maya tojol-ab'al présente un mélange de caractéristiques linguistiques de deux langues mayas apparentées, le chuj et le tseltal. Les similitudes systématiques qui se retrouvent dans des langues apparentées rendent non seulement difficile, méthodologiquement, l'identification de la source des caractéristiques langagières spécifiques, mais encore ces similitudes héritées peuvent-elles également modifier les processus et les résultats du mélange des langues, aboutissant à des situations proches des schémas observés d'alternance codique entre langues apparentées. Par conséquent, le tojol-ab'al représente un type distinct de langue mixte, qui peut résulter du croisement de langues apparentées.

Zusammenfassung

Definitionen von „gemischten" oder „verflochten" Sprachen stammen fast ausschließlich aus Studien zu Sprachen, die Elemente aus genetisch nicht verwandten Quellen kombinieren. Die Maya-Sprache Tojol-ab'al weist eine Mischung aus sprachlichen Elementen von zwei verwandten Maya-Sprachen auf, Chuj und Tzeltal. Die systematischen Ähnlichkeiten dieser verwandten Sprachen machen es nicht nur methodisch schwierig, die Quelle der spezifischen sprachlichen Merkmale zu identifizieren. Die genealogische Ähnlichkeit kann auch die Prozesse und Ergebnisse von Sprachmischung in einer Weise verändern, so dass sie bestimmten Mustern von Code-Switching zwischen verwandten Sprachen gleicht. Tojol-ab'al stellt daher einen bestimmten Typ Mischsprache dar, der wohl nur aus einer Mischung verwandter Sprachen hervorgehen kann.

Author's address

Danny Law
University of Texas at Austin
305 E. 23rd Street, B5100
AUSTIN, TX 78712
USA
dannylaw@austin.utexas.edu