Language Contact and Language Documentation: Whence and Whither?

Lyle Campbell
University of Hawai'i at Mānoa

1. Introduction

Most of my work in recent times has been with documentation of endangered languages. Language documentation typically involves languages in contact. Nevertheless, we mostly talk about language documentation in terms of a single language, typically an endangered language, although in almost all cases an "endangered language" involves languages in contact. Given that, it is strange that language documentation has mostly been directed to single languages, aimed at structural analysis of individual languages, at preparing and archiving corpora, or at revitalization efforts dedicated to a single language, with little direct attention paid to the language contact involving the particular language being documented. The goal of this paper is to address language contact in the context of language documentation, and language documentation in the context of language contact. The ultimate goal is to examine some general claims about language contact, and to examine how language documentation research can contribute to understanding of aspects of language contact. Several specific claims about language contact are discussed and assessed here in the context of language documentation.[1]

1 Views about how to define language documentation vary. Himmelmann's (1998, 2006) early definition contrasted language description and language documentation; for him language documentation "aims at the record of the linguistic practices and traditions of a speech community" (Himmelmann 1998:9-10), and "language documentation may be characterized as radically expanded text collection: (Himmelmann 1998: 2). Himmelmann's (2006:1) definition was: "a language documentation is a lasting, multipurpose record of a language"; it "is primarily concerned with the compilation and preservation of linguistic primary data and interfaces between primary data and various types of analyses based on these data." Several others favored similar views; for example, Woodbury's (2011:159) is often cited. Given this definition, it is hardly surprising that many, according to Himmelmann (2012:187) himself, thought that this approach means that:

> "Documentary linguistics is all about technology and (digital) archiving.
> Documentary linguistics is just concerned with (mindlessly) collecting heaps of data without any concern for analysis and structure.
> Documentary linguistics is actually opposed to analysis."

The paper has two threads.
[1] I look at the impact from language documentation on several claims about language contact (with examples mostly from my own work with indigenous languages of the Americas, with also examples from others).

[2] I make recommendations for how the language documentation research can contribute to understanding of language contact, and for how taking a broader perspective on language contact can improve language documentation projects.

The invitation to participate in this conference led me to some stocktaking, to re-thinking several hypotheses or claims about language contact and language change in the context of language documentation, including several claims that I have made in the past. After much additional work in language contact, I ask myself, what do I think about these things today?

2. Language documentation and the documentation of language contact

In this section, I consider the implications for claims about language contact involving several cases encountered in language documentation work.

2.2. Implications of extensive language contact in Misión La Paz

Several aspects of what I report here are based on findings from language documentation research in Misión La Paz, Salta Province, Argentina (henceforth MLP).[2] Three indigenous languages are spoken in MLP: Chorote, Nivaclé (a.k.a Chulupí, Ashluslay), and Wichí (formerly called Mataco). All three are members of the Matacoan language family, diversified on the order of Germanic languages. At the time the fieldwork research began, there were about 650 inhabitants, but it grew to over 850 while

---

In contrast, many linguists follow the Boasian view that language documentation includes language description and analysis, with a grammar and a dictionary, as well as corpora (texts). As Rehg (2007:15) puts it, language documentation 'involves the development of high-quality grammatical materials and an extensive lexicon based on a full range of textual genres and registers, as well as audio and video recordings, all of which are fully annotated, of archival quality, and publicly accessible'. For discussion of what adequate language documentation is, see Rhodes et al. (2007:3).

Opinions differ; however as Himmelmann (2012) explains in his revised view, there is also agreement, but with differences of emphasis. Some scholars give greater prominence to a large number of recordings representing many genres and on the technology for recording and archiving, while others give more attention to description, to analysis, which includes a grammar and dictionary. We can summarize that adequate language documentation aims at a transparent record of a language where that record includes language analysis and the production of a grammar and a dictionary, along with the rich corpus of recordings, and there is no sharp dichotomy between language documentation and language description.

2 Language documentation research in Misión La Paz was supported by the grant, "Description of Chorote, Nivaclé and Kadiwéu: Three of the Least Known and Most Endangered Languages of the Chaco," from the Endangered Languages Documentation Programme (Rausing Charitable Fund), School of Oriental and African Studies, London University (co-principal investigators Lyle Campbell, Verónica Grondona, and Filomena Sandalo).

this study was being conducted (and to c. 1,000 today). The indigenous people there live in very poor socioeconomic conditions, and they maintain much of the traditional culture.

MLP is located on the Pilcomayo River, across from Paraguay and about 20 km downriver from Bolivia. Here, I focus on changes involving language contact and claims about convergence.

Some background on the multilingualism in MLP is relevant for understanding the language changes described below. Speakers and hearers in conversations in MLP are typically not speaking the same language to one another. Instead, people communicate regularly with speakers of different languages, but very often not replying in the same language as the one addressed to them. Each participant in a conversation typically speaks his or her own language, while the other participants in the conversation reply in their own language, in dual-lingualism. Linguistic exogamy is also practiced in MLP – one marries someone who speaks a different language. Here, each spouse speaks his or her own language and is addressed in and understands the other spouse's language in return – a spouse does not accommodate by speaking the other spouse's language; each maintains and uses his or her own language. In general, people identify with a single language and speak it with all others. They claim to understand but not speak one or both of the other two indigenous languages in MLP. Nevertheless, these languages are spoken around them constantly and they usually have perfect comprehension of the languages that they claim not to speak. In most families, multilingual, dual-linguistic conversations are going on all day long every day. This stable combination of dual-lingualism with linguistic exogamy appears to be unique in the world. (See Campbell & Grondona 2010 for details.)

The common view about languages in intensive contact is that they should undergo structural convergence, becoming more similar to one another, and should not undergo changes that make them less similar. For example, Bloomfield' (1933: 476) declared: "When two speech communities are in continuous communication, linguistic convergence is expected, and any degree of divergence requires an explanation." This quote is often repeated, and many similar citations could be added. For example, in famous cases from India, different languages in intensive contact have changed to become more structurally similar to one another, so that rather exact one-to-one structural matching in morpheme-by-morpheme translations is possible (see Gumperz & Wilson 1971, Nadkarni 1975).

The question is, do the languages of MLP, in intensive contact, tend to converge, as expected? The answer appears to be "no."

Against expectations, the three indigenous languages in MLP show no obvious evidence of changes that make them structurally more similar to one another; rather, they have undergone changes that make them structurally more different. Changes in the languages of MLP bear significantly on claims about convergence in contact situations. This is illustrated by the following examples.

All three languages have or had /ɬ/, phonemic voiceless 'l', as in Nivaclé *ɬuʔp*, Wichí *ɬup*, Chorote *lop*/*xlop* 'nest', to cite one set of cognate forms. However, Chorote in MLP has changed. Speakers no longer have /ɬ/; rather they have changed it to a consonant cluster of /x/ + voiced /l/, which alternates with just plain /l/ (with no /x/) in some contexts, especially word-initially and word-finally, as in: *xlop* / *lop* 'nest', *xlaʔa* / *laʔa* 'fruit', *xlam* / *lam* 'he', *xloma* / *loma* 'day', *axlu* / *alu* 'iguana', *samexl* / *samel* 'we', etc.

This change has taken place in spite of the fact that these Chorote speakers are in constant intensive contact with the speakers of the other two languages that preserve their original voiceless "l", /ɬ/. To change this /ɬ/ when the other languages maintain it goes against expectations. It is expected that if the other languages in the intensive contact situation have voiceless "l" there would be pressure on Chorote not to lose or change its voiceless "l" but instead to remain structurally similar with regard to this trait to the other two languages with which Chorote is in intensive contact.

In another example, both Nivaclé and Wichí have contrastive first-person plural inclusive vs. exclusive pronominal forms, as seen in the contrasts in Nivaclé between the (a) and (b) pairs in (1), (2), and (3).

(1) a. kas-waʔtša
1PL.INCLUSIVE.POSSESSIVE-PRONOMINAL.ROOT
'we' (all of us)

b. yi-waʔtša-ʔeł
1.POSSESSIVE-PRONOMINAL.ROOT-PL.EXCLUSIVE
'we' (but not you)

(2) a. katsi-tata
1PL. INCLUSIVE. POSSESSIVE-father
'our father' (of all of us)

b. yi-tata-ʔeł
1. POSSESSIVE-father-PL. EXCLUSIVE
'our father' (but not yours)

(3) a. šta-sekkis
1PL. INCLUSIVE.ACTIVE-scrape
'we scrape it' (all of us)

b. xa-sekkis-eł
1ACTIVE-scrape-PL. EXCLUSIVE
'we scrape it' (but not you)

The inclusive-exclusive contrast in Wichí is seen in the difference between the (a) and (b) forms in examples (4) through (6).

(4) a. n-ʔameł
1PL.INCLUSIVE.POSSESSIVE-PRONOMINAL.ROOT
'we' (all of us)

b. no-łamel, o-łamel
1PL.EXCLUSIVE.POSSESSIVE-PRONOMINAL.ROOT
'we' (but not you)

(5) a. ła-čoti
1PL.INCLUSIVE.POSSESSIVE-grandmother
'our grandmother' (of us all)

b. n-čoti
1PL.EXCLUSIVE.POSSESSIVE-grandmother
'our grandmother' (but not yours)

(6) a. yaʔ-lɑn
ACTIVE.1PL.INCLUSIVE-kill
'we kill it' (all of us)

b. na-lɑn
ACTIVE.1PL.EXCLUSIVE-kill
'we kill it' (but not you) (Wichí examples from Terraza 2008).

However, Chorote in MLP has lost this inclusive-exclusive contrast in first person plural pronouns which the language once had and now has only a non-contrastive first-person plural. This is seen in the comparison of the single Chorote form in (7a) and (8a) with the Nivaclé contrasting forms in (7b)-(7c) and (8b)-(8c).

(7) a. Chorote: si-ʔleh
1PL.POSSESSIVE-language
'our language'

b. Nivaclé: kas-kliʔš
1PL.INCLUSIVE.POSSESSIVE-language]
'our language' (INCLUSIVE)

c. Nivaclé: xa-kliʔš-eł
1SG.POSSESSIVE-language-PL.EXCLUSIVE
'our language' (EXCLUSIVE)

(8) a. Chorote: a-lan-a
we-kill-SUFFIX
'we killed it'

b. Nivaclé: šta-klɑn
1PL.ACTIVE.INCLUSIVE-kill
'we kill it' (INCLUSIVE)

c. Nivaclé: xa-klɑn-eł
1SG.ACTIVE-kill-PL.EXCLUSIVE
'we kill it' (EXCLUSIVE)

Again, Chorote would not be expected to lose such a morphological contrast that is so salient in the other two languages which speakers of Chorote hear and understand constantly in MLP.

The third example involves a change in Nivaclé. In both Chorote and downriver dialects of Nivaclé, when active verbs appear with the 'prospective' (future intent, PROS) morpheme, they are required to take the pronominal prefixes for stative verbs – even when an inherently active verb is involved – as seen in the contrasts in Chorote between (9a) and (9b) and between (10a) and (10b).

(9) a. **a**-ʔwešiy
1ACTIVIE-hunt.and.gather
'I hunt and gather' (ACTIVE)

b. **si**-ʔwešiy=**ayi**
1STATIVE-hunt=**PROS**
'I'm going to hunt and gather' (STATIVE)

(10) a. **hi**-kapehnan
2ACTIVE-cook
'you cook' (ACTIVE)

b. **in**-kapehnan=**ayi**
2STAT-cook=**PROS**
'you are going to cook' (STATIVE)

This construction, in which the 'prospective' requires stative pronominal agreement affixes on the verb (regardless of whether the verb is inherently active or stative), is the original state of affairs for these languages. However, MLP Nivaclé has changed. The corresponding construction in Nivaclé does not takes (nor does it allow) stative pronominal agreement markers with the prospective, but can only bear the active pronominal agreement markers with inherently active verbs, as in (11a). The same utterance but with a stative subject prefix (as required in the other languages) is ungrammatical here, as seen in (11b).

(11) a. **xa**-woʔ=**xayu**
1ACTIVE-hunt.for=**PROS**
'I'm going to hunt for it'

b. ***tsi**-woʔ=**xayu**
1STATIVE-hunt.for=**PROS**
(same intended meaning as in (11a))

It is expected that, because of intensive language contact, the restriction on pronominal agreement markers with prospective that holds in Chorote and other Nivaclé dialects (several speakers of which also live in MLP) would be maintained in MLP (Upriver) Nivaclé. However, this is not what happened, and in spite of assumed influence from these other languages and dialects with which Nivaclé is in intensive language contact to maintain this restriction, MPL Nivaclé changed.[3]

3 Nivaclé, it should be noted, has a 'desiderative' clitic morpheme that is similar in form to the 'prospective', and this desiderative does require stative pronominal agreement markers even when

(See Campbell & Grondona 2010 for discussion of other examples.)

Cases such as these in MLP challenge us to investigate more thoroughly what happens in situations of intensive, intimate language contact. None of these changes just reported would be known without the recent language documentation undertaken in MLP, and these findings contribute not just to understanding the structure and history of these particular languages, but also have several implications for broader claims about language change.

It is mistaken to think that languages in intensive contact must necessarily change only in the direction of more convergence and never in the direction of more divergence. While the expectation has always been that heavy language contact would result in more similarity among the languages in contact and not in greater difference, recently other cases also have been recognized of some aspects of language contact resulting in diversification rather than convergence (see Ellison & Micelli in press, Evans In press, and Kühl & Braunmüller 2014; cf. aslo Thomason 2007).[4] It is important to document unusual and unique facts about language contact, socio-cultural facts such as language choice and language use, as we document endangered languages. It is important to document cases of language contact (multilingual language usage and choice), as in the MLP case, not just individual languages within multilingual settings. It is not just straight linguistic information that we stand to lose with loss of endangered languages, but facts with consequences for claims about language contact. With the increasing threat from Spanish and immigration into MLP, MLP's pattern of multilingualism is decaying and will probably be lost in the near future. If that had happened without it being documented, knowledge of the existence of this arguably unique pattern of multilingualism and language choice (dual-lingualism + linguistic exogamy) would have been lost, a very unfortunate loss, because the MPL case reveals much about the range of possibilities for language choice in multilingual communities and can tell us much about various claims in the literature concerning language contact and multilingualism.

The examples from MLP just presented make clear that the assumed pressure towards convergence in intensive language contact did not prevent these three languages from becoming more distinct from one another. Rather, they have undergone changes that result in greater difference among the three, while no changes towards convergence are evident in the languages of MLP. In short, it is mistaken to claim that languages in intensive contact must necessarily change only in the direction of more structural similarity and never in the direction of more divergence.

3. Borrowed sounds, without distributional restrictions of original language

In "Language contact and sound change" (Campbell 1976), I presented examples of borrowed sounds and I hypothesized that:

---

the verb root is inherently active: *=jayu*, *=jatsu* 'desiderative'. Difference between the 'desideraative' forms and the 'prospective' construction can be seen in the comparison of (1) with (2):

(1) tsi-mô=**xayu** [1STATIVE-sleep=DESIDEERATIVE] 'I want to sleep' ('I'm sleepy')

(2) xa-mô7 **xayu** [1ACTIVE-sleep PROSPECTIVE] 'I am going to sleep').

The 'desiderative', however, is very rare, and is not the same as the 'prospective', which is frequent and matches the prospective of the other related languages.

4 I thank Patience Epps drawing these publications to my attention.

> Since the borrowed segments seem generally to lack in the borrowing language distributional restrictions holding in the donor languages, this suggests the hypothesis for further testing that borrowed segments generally lack such distributional restrictions of the donors. (Campbell 1976:83.)

I discussed several cases of such borrowed sounds. For example, retroflex consonants were borrowed from Dravidian to Indo-Aryan. In Proto-Dravidian, retroflex segments could not occur word-initially, though they can occur word-initially in Indo-Aryan (cf. Burrow 1971[:]36, Zvelebil 1970[:]77). The distribution of clicks in Zulu (and some other Southern Bantu languages) are less restricted than in the so-called "Khoisan" languages from which they borrowed the clicks. (Campbell 1976:83.)

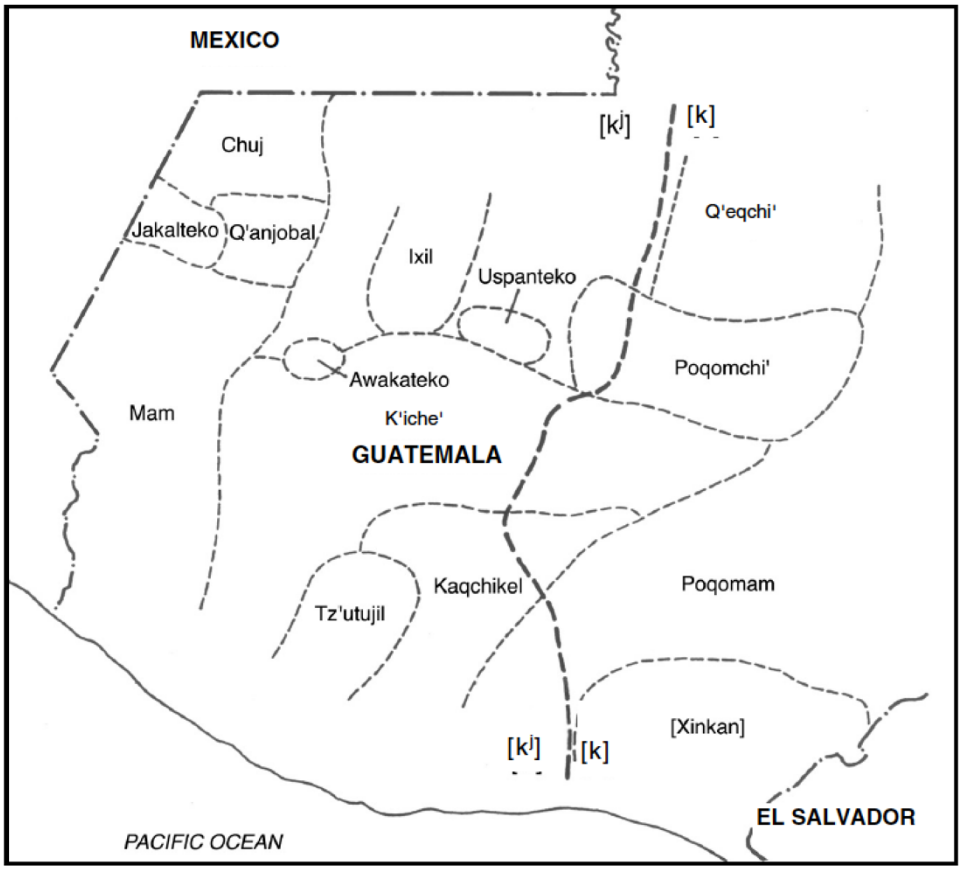
In retrospect, I now believe more firmly that this hypothesis is probably correct; it should be investigated further.

I would now add a further research question: Is it the case with borrowed morphemes or grammatical constructions that certain co-occurrence restrictions and distributional restrictions that hold in the donor language may not necessarily be maintained in a borrowing language along with the borrowed grammatical morphemes or constructions?

4. Borrowed phonological rules (constraints)

I also presented evidence the phonological rules could be borrowed (diffused) (Campbell 1976, 2013a:69-71). Some examples are the borrowing of vowel harmony from Turkish into Asia Minor dialects of Greek, borrowing from Breton of final-consonant devoicing in French in Quimper, first-syllable stress diffused among many of the languages in the Baltic area, and others. A particular interesting case is the diffusion of the rule that palatalizes velar stops (*k, k'*) when a uvular (*q, q', X*) is the next consonant in the root in several Mayan languages. This began in Mamean languages and spread to Mocho' and to K'ichean languages, as for example in K'iche': //k'aq // → k^{y}'aq 'flea', //ke:χ// → k^{y}e:χ 'deer'. A few other forms that show the result of this rule are: *kyaq* 'red', *k'yaq* 'flea', *ikyaq'* 'guava', *iškyaq'* 'fingernail', *ikyaχ* 'axe', and *k'yaχ* 'flour'. Interestingly the rule, diffusing from west to east, reached and affected only western dialects of several K'ichean languages but not eastern dialects of these same languages, as seen in Map 1.

Map 1: Diffusion of the velar palatalization rule in K'ichean languages (redrawn after Campbell 1977 map 1)



It is important to investigate further the diffusion of phonological rules (or constraints) in language contact.

5. Retention due to language contact

I have argued that not only could sounds, sound changes, and phonological rules be borrowed in language contact, but that also language contact can result in resistance to changes that otherwise would take place. Several examples from the Spanish of the Americas illustrate this. Most varieties of New World Spanish underwent what is called *yeísmo*, the change of palatal "l" (spelled <ll>) to the glide *y* ($l^y > y$), as in *calló* 'he/she/it shut up' vs. *cayó* 'he/she/it fell', both now pronounced as /kayó/. An exception to this change, however, is the Spanish of the Andes and adjacent regions, where indigenous languages with large numbers of speakers have phonemic /l^y/. The hypothesis is that Andean Spanish resisted the change of $l^y > y$ because of contact with Quechua and Aymara, preserving distinctive /l^y/ in Spanish because of the salience of this sound in those indigenous languages.

Another example is the assibilated "r" of many Spanish speakers in Guatemala, found

also in Mayan languages of Guatemala and in certain Peninsular Spanish dialects. It is argued that contact with K'ichean languages which have the assibilated "r" resulted in preservation of a pronunciation not found in the great majority of other Spanish dialects, but known in some Peninsular varieties.

A final example involves the preservation in Guatemalan, Chiapan, and Yucatecan Spanish of the pleonastic possessive due to contact with Mayan languages. The pleonastic possessive is seen in cases such as *tengo un mi caballo* 'I have a horse' (literally 'I have one my horse') which is equivalent to, for example, K'iche' *k'o xun nu-kye:x* 'I have a horse' (literally 'exists one my-horse'). The retention of this pleonastic possessive is found only in Spanish dialects that are in contact with Mayan language. This construction, once was more widely used in Peninsular Spanish, now is no longer known there (Martin 1978).

Again, retention (or resistance to change) due to language contact deserves much more investigation.

6. Claims about possible changes in endangered languages and their implications

Documentation of endangered languages can provide insights into language change, in particular into how they change when they are severely threatened. Investigation of endangered languages has raised questions about the **nature of language change** itself. The examples discussed here involve the linguistic behavior of semispeakers in severely endangered languages.

6.1. Does sound change in endangered languages have to be regular?

Though in non-endangered languages sound change is generally considered regular (Campbell 1996, 2013a, Labov 1994), the answer to this question appears to be "no," sound change in endangered languages does not necessarily have to be regular. Changes in endangered languages sometimes do not change all instances of a particular sound in defined contexts in the same way, sometimes changing the sound in some words and sometimes not changing the sound in other words. For example, in Tlahuica (a.k.a. Ocuilteco, an Otomanguean language of Mexico), fully competent native speakers voice stops after nasals, but semispeakers of the language sometimes voice (e.g. *nd*) and sometimes do not (e.g. *nt*), irregularly. Semispeakers of Cuisnahuat Nawat (a.k.a. Pipil, a Uto-Aztecan language of El Salvador) irregularly sometimes devoiced and sometimes did not devoice final /l/, though fully competent speakers always have the voiceless allophone word-finally (see Campbell & Muntzel 1989 for these and other examples). These are not regular changes.

Examples such as these go against the Neogrammarian regularity hypothesis that sound laws suffer no exceptions. However, given that the regularity of sound change holds well in non-endangered languages, we would not give up this valuable principle just because sometimes the speech of semispeakers may fail to conform, just as we would not abandon an otherwise well supported linguistic principle if we found violations of it only in the speech of adult second-language learners or of persons with speech pathologies.

6.2. Can endangered languages change in ways not possible in fully viable languages?

Another question is, can endangered languages change in ways not normally possible in non-endangered languages? The answer appears to be "yes." For example, some semispeakers of Jumaytepeque (a Xinkan language of Guatemala) arbitrarily glottalized essentially every possible consonant (C > C') (Campbell & Muntzel 1989:189) – this is not a normal sound change, and it would not be expected in fully viable languages for a change to result in only glottalized consonants with no plain counterparts. It appears in cases such as this that semispeakers become aware of sounds not found in their dominant language but that they have not learned exactly where these sounds occur, and consequently they hyper-correct, employing these "exotic" sounds (exotic from the point of view of the dominant language) with great frequency and in ways inconsistent with their distribution in the fully viable form of the endangered language.

For example in the Jumaytepeque case, it is impossible that a viable language would change all its plain voiceless stops into glottalized stops, leaving no plain voiceless stop counterparts in the language. This would violate the universal that the presence of glottalized consonants implies the presence also of plain non-glottalized counterparts in languages (C' ⊃ C). Changes in fully viable languages do not violate linguistic universals (cf. Labov 1994).

In another example, semispeakers of Teotepeque Nawat overgeneralized voiceless "l", losing plain voiced "l" entirely. In viable Nawat (Pipil), voiceless "l" is an allophone of /l/ word-finally. Teotepeque semispeakers failed to learn the context that conditioned the pronunciation of voiceless "l" and thus changed l > ɬ everywhere, as for example in *peeɬu* 'dog', *čakaɬin* 'shrimp', *čiɬtik* 'red' – none of which had voiceless "l" in the speech of fully competent speakers (see Campbell 1985). Such a change is not an expected and not normal; it is typologically strained.

In another example from Teotepeque Nawat, semispeakers changed *š̻* > *r* (*š̻* is a voiceless retroflex non-apical laminal fricative; it changed to a trilled "r"). Native Nawat has no *r* sounds; the change *š̻* > *r* is unnatural and unexpected. This change appears to involve speakers imposing the prejudices of the dominant language onto Teotepeque Nawat. Local Spanish, the dominant language, has a stigmatized variant *š̻* of its trilled "r". The stigma of the assibilated *š̻* pronunciation of *r* in local Spanish caused these Teotepeque Nawat speakers to shift the native Nawat sound to match Spanish prestige, replacing their *š̻* with trilled "r", for example *ručit* < *š̻učit* 'flower. This change is not natural and would not be expected in non-endangered languages. (For these and other examples, see Campbell & Muntzel 1989.)

These two claims – of the existence of irregular sound changes and of unnatural sound changes in endangered languages – have important implications for historical linguistics and deserve more extensive investigation.

7. Implications linguistic areas

Defining linguistic areas has almost always proven contentious at best (Campbell 2006, in press, Campbell et al. 1986). Language documentation work in the Chaco region (as in section 2.2. above) has revealed that there is considerable sharing of structural traits involving Chaco languages, traits such as grammatical gender, SVO basic word order, rich systems of demonstratives and verbal directional, lack of overtly marked verbal tense, and active-stative verb alignment (Campbell 2013b). However, these shared traits do not come

together to reveal a cohesive geographical area. This case has implications for defining linguistic areas generally.

Most of the shared traits in the Chaco are found extending also into languages well beyond the Chaco, and others are characteristic of only a few of the languages within the region. Only a few of the traits seem true of a majority of Chaco languages, but none is unique to the area and some are quite commonplace in the world, for example SVO word order. This raises the question, is the Chaco a legitimate linguistic area, or, if it is, how should a Chaco Linguistic Area be defined?

Tupí-Guaranían illustrates some of the problems for defining a Chaco linguistic area and indeed for linguistic areas in general. Opinion has diverged about whether Tupí-Guaranían languages should be considered members of a Chaco linguistic area, though no grounds for excluding Chiriguano and Tapiete (Tupí-Guaranían languages in the region) have been found (see Comrie 2010:86 for discussion). Tupí-Guaranían shares most of the Chaco traits just listed. Since Tupí-Guaranían extends far beyond the Chaco region, with a large presence in Amazonia but with few representatives also in the Chaco, its inclusion in a Chaco linguistics area would extend the putative Chaco linguistic area far beyond the geographically defined Chaco region – i.e. hardly a "Chaco" area if defined in this way. If Tupí-Guaranían is included, since these languages also share many traits with languages of the Amazonian area (see Campbell 2006, Campbell & Grondona 2012), how could we establish what belongs to the Chaco linguistic area and what to the Amazonian linguistic area, and how are the two to be distinguished? If Tupí-Guaranían is not included, the areal definition of the Chaco as defined on the basis of shared traits is compromised, since many of the shared traits seemingly reflective of a Chaco area are also found Tupí-Guaranían languages and in other neighboring languages well beyond the Chaco region. This sprawling, overlapping, or twining of shared traits among Chaco languages and languages of Amazonia, of the Andes, and elsewhere complicates any attempt to define a Chaco linguistic area with recognizable boundaries.

To deal with problems of the sort represented in the distribution of the traits encountered among languages of the Chaco, I have proposed distinguishing kinds of linguistic areas (Campbell 2013b, in press). One is the Linguistic Area *Sensu Stricto* (LASS), a geographical region defined by shared diffused traits mostly contained within and shared across the languages of a clearly delimited geographical area. The other is the Trait-Sprawl Area (TSA), an entity where the individual traits can pattern in disordered ways, some crossing some languages while others cross other languages, with some extending in one direction, others in another, with some overlapping others in part of their distribution but also not coinciding in other parts of their geographical distribution, some extending also into other linguistic areas. The focus of the TSA is the actual diffused traits themselves rather than their geography.[5]

5 Note that Comrie et al. (2010:89) declare their faith that the Chaco is a linguistic area by acknowledging but dismissing the complications just discussed. Comrie et al. (2010:125) conclude by calling the Chaco a core-periphery (núcleo-periferia) type linguistics area. Their approach appears to be a declaration of interest in shared traits found in and around Chaco languages, no matter what the ultimate geographical distribution of the various traits might be. I support such a view if it is taken to mean that what is important are the diffused changes themselves, not attempts to set up linguistic areas defined by the overall reach of these traits (Campbell 2006, in press).

There does appear to be considerable diffusion of structural traits involving Chaco languages, but these do not come together in such a way as to suggest a cohesive geographical area. Rather, they show varying linkages with languages and regions outside the Chaco on all sides, while at the same time often not linking all Chaco languages. This evidence is too weak to declare a Chaco linguistic area *sensu stricto*, a LASS. However, the sprawling of trait among Chaco languages and beyond is a good example of a *trait-sprawling area* (TSA), the concept advocated here to avoid definitional problems with the concept of linguistic area, focussing on actual diffused traits and on answering the question, "what happened?", rather than on fixing a geographical area to be defined by its shared linguistic traits.

8. Conclusions

In this paper I have advocated for recognizing the close connection between documentation of endangered languages and language contact, and for documenting language contact along with research to document endangered languages, and I have considered several cases of the kinds of language-contact phenomena that can be found in documenting languages. I looked back at several claims about language contact (the "whence" of the title), and I made recommendations for future investigation ("whither).

Claims that should be investigated more thoroughly include:

[1] Change towards greater similarity is not as necessary in intensive language contact as usually claimed; change towards greater divergence is also possible, as illustrated by the MLP case. What can happen in intensive contact needs more attention.

[2a] Borrowed sounds typically enter the borrowing language without some of the distributional restrictions that hold on the sound in the donor language.

[2a] The hypothesis should be tested that borrowed morphemes and grammatical constructions may also enter the borrowing language without some of the distributional restrictions that hold on them in the donor language.

[3] Phonological rules can be borrowed; how they spread needs to be investigated further (as in the K'ichean palatalization example).

[4] Some changes can be resisted due to influence from other languages in language contact situations (as in the several Spanish examples cited here).

[5] In highly endangered languages, sound change need not be regular or natural (as they typically are in non-endangered languages).

[6] Discoveries in language documentation projects have implications for how we define linguistic areas: as documentation of the Chaco languages shows, languages can share traits but still not fit traditional views of a linguistic area. I argue for TSA (trait-sprawling area), contrasted with LASS (linguistic area *sensu stricto*), to allow ways out of the morass of constantly spinning our wheels trying to define and defend linguistic areas.

References

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart & Winston.
Burrow, Thoms. 1971. Spontaneous cerebrals in Sanskrit. *Bulletin of the School of Oriental and African Studies*, *34*, 538-559.
Campbell, Lyle. 1976 Language contact and sound change. *Current Progress on Historical Linguistics:* In William M. Christie, Jr., (Ed.), *Proceedings of the Second International Conference on Historical Linguistics* (pp. 181-94). Amsterdam: North Holland Publishing.
Campbell, Lyle. 1977. *Quichean linguistic prehistory*. (University California Publications in Linguistics, 81.) Berkeley: University of California Press.
Campbell, Lyle. 1985. *The Pipil language of El Salvador*. Berlin: Mouton de Gruyter.
Campbell, Lyle. 2006. Areal linguistics: a closer scrutiny. In Yaron Matras, April McMahon & Nigel Vincent (Eds.), *Linguistic areas: convergence in historical and typological perspective* (pp. 1-31). Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.
Campbell, Lyle. 2013a. *Historical linguistics: an introduction*. (3rd edition). Edinburgh: Edinburgh University Press, and Cambridge, MA: MIT Press.
Campbell, Lyle. 2013b. Language Contact and Linguistic Change in the Chaco. In Ana Suelly Arruda Câmara Cabral & Jorge Domingues Lopes (Eds.), *Encontro internacional arqueologia e linguística histórica das línguas indígenas sulamericanas*, Special issue of *Revista Brasileira de Linguística Antropológica* 5, *259-291*.
Campbell, Lyle. In press. Why is it so hard to define a linguistic area? In Raymond Hickey (Ed.) *Handbook of areal linguistics*. Cambridge: Cambridge University Press.
Campbell, Lyle & Verónica Grondona. 2012. Languages of the Chaco and Southern Cone. In Lyle Campbell & Verónica Grondona (Eds.), *The indigenous languages of South America: a comprehensive guide* (pp. 625-668). Berlin: Mouton de Gruyter.
Campbell, Lyle & Martha Muntzel. 1989. The structural consequences of language death. In Nancy Dorian (Ed.), *Investigating obsolescence: studies in language death* (pp. 181-196). Cambridge: Cambridge University Press.
Campbell, Lyle, Terrence Kaufman & Thomas Smith-Stark. 1986. Mesoamerica as a linguistic area. *Language, 62*, 530–70.
Comrie, Bernard, Lucía A. Golluscio, Hebe González & Alejandra Vidal. 2010. El Chaco como área lingüística. In ed. by Zarina Estrada Fernández & Ramón Arzápalo Marín, (Eds.), *Estudios de lenguas amerindias 2: contribuciones al estudio de las lenguas originarias de América* (pp. 85-132). Hermosillo, Sonora: Editorial Universidad de Sonora.
Ellison, Mark & Luisa Miceli. In press. Contact-induced lexical differentiation: Bilingual monitoring leads to a loss of shared lexical forms. *Language*.
Evans, Nicholas. In press. Linguistic divergence under contact. In Michela Cennamo & Claudia Fabrizio (Eds), *Selected Papers from the 22nd International Conference on Historical Linguistics*. Amsterdam/Philadelphia: John Benjamins.
Gumperz, John J. & Robert Wilson. 1971. Convergence and creolization: A case from the Indo-Aryan/Dravidian border in India. In Dell Hymes (Ed.), *Pidginization and creolization of languages* (pp. 151–67). Cambridge: Cambridge University Press.
Himmelmann, Nikolaus. 1998. Documentary and descriptive linguistics. *Linguistics, 36*, 161–95.
Himmelmann, Nikolaus. 2006. Language documentation: What is it and what is it good for? In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (Eds.), *Essentials of language documentation* (pp. 1-30). Berlin: Mouton de Gruyter.
Himmelmann, Nikolaus. 2012. Linguistic data types and the interface between language

documentation and description. *Language Documentation and Conservation, 6*, 187-207. 〈http:// n rc.hawaii.edu/ldc/〉
Kühl, Karolin & Kurt Braunmüller. 2014. Linguistic stability and divergence: an extended perspective on language contact. In K. Braunmüller, S. Hoder & K. Kühl (eds.), *Stability and divergence in language contact* (pp. 4-38). Amsterdam: John Benjamins.
Labov, William. 1994. *Principles of linguistic change: internal factors*. Oxford: Blackwell.
Martin, Laura. 1978. Mayan influence in Guatemalan Spanish: a research outline and test case. In Nora C. En- gland, *Papers in Mayan* linguistics (pp. 106-26). Columbia: Museum of Anthropology, University of Missouri.
Nadkarni, Mangesh V. 1975. Bilingualism and syntactic change in Konkani. *Language, 51*, 672-683.
Rehg, Kenneth L. 2007. The language documentation and conservation initiative at the University of Hawai'i at Mānoa. Documenting and revitalizing Austronesian languages. In D. Victoria Rau & Margaret Florey (Eds), *Language Documentation and Conservation, special publication no. 1, documenting and revitalizing austronesian languages*, 13–24. (https://scholarspace.manoa.hawaii.edu/handle/10125/1350)
Rhodes, Richard, Lenore A. Grenoble, Anna Berge & Paula Radetzky. 2007. Adequacy of documentation. (A preliminary report to the CELP.) Linguistic Society of America Committee on Endangered Languages and their Preservation.
Thomason, Sarah G. 2007. Language contact and deliberate change. J*ournal of Language Contact, 1*, 41-62.
Woodbury, Anthony. 2010. Language documentation. In, Peter Austin & Julia Sallabank (Eds.), *The Cambridge handbook of endangered languages* (pp. 159–186). Cambridge: Cambridge University Press.
Zvelebil, Kamil V. 1970. *Comparative Dravidian phonology*. The Hague: Mouton.

Author's address:
Lyle Campbell
Department of Linguistics, University of Hawai'i at Mānoa
1890 East-West Rd, Honolulu, HI 96822 USA
Campbell@hawaii.edu