Distinguishing Contact-Induced Change from Language Drift in Genetically Related Languages

T. Mark Ellison
Psychology
University of Western Australia
Mark.Ellison@uwa.edu.au

Luisa Miceli
Linguistics
University of Western Australia
lmiceli@cyllene.uwa.edu.au

Abstract

Languages evolve, undergoing repeated small changes, some with permanent effect and some not. Changes affecting a language may be independent or contact-induced. Independent changes arise internally or, if externally, from non-linguistic causes. En masse, such changes cause isolated languages to drift apart in lexical form and grammatical structure. Contact-induced changes can happen when languages share speakers, or when their speakers are in contact.

Frequently, languages in contact are related, having a common ancestor from which they still retain visible structure. This relatedness makes it difficult to distinguish contact-induced change from inherited similarities.

In this paper, we present a simulation of contact-induced change. We show that it is possible to distinguish contact-induced change from independent change given (a) enough data, and (b) that the contact-induced change is strong enough. For a particular model, we determine how much data is enough to distinguish these two cases at $p < 0.05$.

1 Introduction

Evolutionary change happens when structures are copied, the copying is inexact, and the survival of copies is uncertain. Many structures undergo this kind of reproduction, change and death: biological organisms, fashions, languages. Often evolutionary change leaves little or no trace, except for those copies which are present at the moment. In these cases, determining the evolutionary history of a family of structures involves comparing surviving copies and making inferences from where they correspond and where they differ.

Language is, for the most part, one of those cases. Most languages have not had a writing system until recently, and so their history has left no direct trace. Since the 18th century, linguists have been comparing languages to reconstruct both common parents and individual histories for these languages (Jones, 1786; Schleicher, 1861; Brugmann, 1884, for example).

In this paper, we hope to contribute to this effort by presenting a formal model of a particular kind of evolutionary change, namely **contact-induced change**, and placing limits on when its past presence can be inferred from synchronic evidence.

Contact-induced change can happen when speakers of different languages come in contact, or where there is a sizeable group of bi- or multilinguals. We distinguish two different types. One type, **contact-induced assimilation** (CIA) changes languages so that they become more similar to each other. This is the type of contact-induced change that is most obvious and that has been best studied. The consensus is that it can affect all sub-systems of a language depending on the intensity of contact (see eg. Thomason & Kaufman 1988). The other type, less frequently noticed and only recently receiving attention (see eg. François 2011, Arnal 2011), is **contact-induced differentiation** (CID) where the change acts specifically to make the languages less similar. This type of contact-induced change predominantly affects the parts of a language which speakers are most conscious of being distinct: the phonological forms of morphemes and words.

It is hard to isolate contact-induced change in

Proceedings of the EACL 2012 Workshop on Computational Models of Language Acquisition and Loss, pages 1–9,
Avignon, France, April 24 2012.

related languages from the effects of common inheritance or normal independent drift. In languages in contact over a long period of time, it is impossible to tell whether the dropping of any single cognate is the result of chance variation or the action of a differentiation process. Likewise, if languages are compared using a single-valued measure of similarity (such as fraction of cognates in a Swadesh list), the effects of more or less contact-induced changes cannot be distinguished from a greater or lesser time-depth since the common ancestor. This is shown in figure 1.

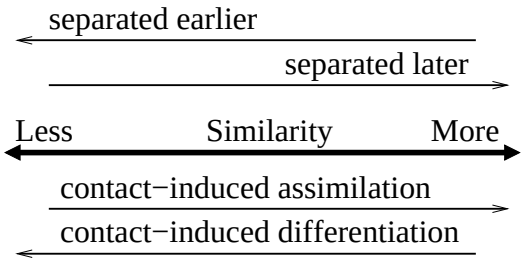

Figure 1: shows the problem of identifying contact-induced change between related languages. Contact-induced assimilation and having a more recent common ancestor can both account for language similarities. Contact-induced differentiation accounts for less similarity, but so does positing a remoter common ancestor that allows time for more independent drift resulting in greater differentiation without contact. A single similarity measure is insufficient to separate time-depth from contact-induced change.

Contact-induced change is, however, different from independent drift. If it is detectable at all, it will be because it creates different counts of synonyms and different proportions of cognates, than drift alone. Thus, with enough data, it should possible to distinguish the effects of time-depth and contact-induced change. This paper presents the results of a simulation to determine just how much data would be enough.

1.1 Overview

Section 2 discusses contact-induced change, and CID in particular. While it is easy to find instances of CIA, eg. borrowing a word from one language to another, it is harder to find unarguable cases of CID. They can be found, however, and some of these are discussed in section 2.2.

Section 3 describes language as a bundle of relations. Language changes can then be modelled as changes in these relations. A formal account of independent and contact-induced changes in relations is given, as the underpinnings for the next section.

This next section (section 4) investigates how much data is needed to develop 95% certainty that contact-induced change has occurred as opposed to independent change alone. As might be expected, the weaker the CIA or CID pressure, the more evidence needed to distinguish the types of change.

The final section considers the implications of the research, and situates it within a larger programme of investigation into contact-induced change.

1.2 Terminology

This paper uses terms from mathematics and linguistics. The term **relation** will only be used in its mathematical sense of a potentially many-to-many association from elements in one set, the **domain**, to elements in another, the **range**. An association between a domain element and a range element will be called a **link**. We introduce the term **doppels** to describe words from different languages which have had a common origin, or are so similar that they might be presumed to have a common origin. These differ from **cognates** in two ways. Although cognates must have had a common origin, doppels need not – they may just look like they do. Also, where there is a common origin, cognates must have evolved with the language as a whole, while doppels may be the result of borrowing. Etymologically, **doppel** is a doppel of the German **Doppel**, *duplicate, copy, double*.

2 Contact-Induced Change in Natural Languages

It is impossible to study language history without being aware of the impact of contact on languages all around the world, not least in the current age of globalisation. However, while the most transparent and best known process of contact-induced assimilation, word borrowing, has been a focus in historical linguistics, some other assimilatory phenomena and almost all differentiating processes are only recently receiving attention.

2.1 Contact-Induced Assimilation

Contact-induced assimilation (CIA) describes any process which causes two languages to become more similar. The increased similarity could be

the result of: more doppels between the languages, due to one language borrowing from another; convergent phonology, as a large community of bilinguals use a single phonemic inventory for both languages; or convergent syntax and morphology. This last may occur as the speech of weak bilinguals, dropping rich morphology and using a lot of word-for-word translations in their non-native tongues, impacts the entire community.

English itself exemplifies the extent to which borrowing can make languages similar. Finkenstaedt and Wolff (1973) found that Latin and French (including Old Norman) have each contributed more words to Modern English than its Germanic parent language has. English speakers consequently often find it easier to learn a Romance language than a Germanic one.

Metatypy (Ross, 2006) is one type of contact-induced change at the grammatical level. Languages engaged in metatypy, such as Kannada and Marathi in the Indian village of Kupwar, can come to have (nearly) identical grammatical and morphological organisation; the languages only differ in their lexical forms. One result is that it is easy to translate from one language to the other, simply by replacing a morpheme in one language by its form in the other.

CIA seems to be much more common than CID. This may, however, be due to the fact that it is much easier to detect, because similarity is inherently less likely to occur by chance than dissimilarity.

2.2 Contact-Induced Differentiation

Because dissimilatory change is sometimes, but not always, hard to detect, many of the known cases of it arise because it is done deliberately and speakers report that they are doing it. Thomason (2007) gives two principal motivations for this kind of deliberate change: (a) a desire or need to increase the difference between one's own speech and someone else's, and (b) a desire or need to keep outsiders at a distance. However, the two recent studies already mentioned – François (2011) and Arnal (2011) – describe how this type of change may arise without "differentiation" per se being the primary motivation (see François 2011:229-30 in particular).

A situation that fits the first description is that found in one of the dialects of Lambayeque Quechua where speakers systematically distort their words in order to make their speech different from that of neighbouring dialects. One of the processes used involves the distortion of words by metathesis giving, for example: /yaw.ra/ from /yawar/, /-tqa/ from /taq/, /-psi/ from /pis/ and /kablata /from /kabalta/ (Thomason 2007:51). This kind of process clearly gives rise to a system with different phonotactics.

There is also anecdotal evidence that non-Castilian languages of the Iberian Peninsula have undergone deliberate differentiation. Wright (1998) reports that some late-medieval Portuguese avoided using words similar or identical to the corresponding Castilian words when a less similar synonym was available, while Vidal (1998) reports the same behaviour among the Catalan. More recently Arnal (2011) has described further differentiating change to Catalan lexical forms due to increased levels of Spanish/Catalan bilingualism among native Spanish speakers, following the establishment of Catalan as a co-official language in 1983. There have also been processes of differentiation at play in Galician, where purists have promoted alternatives to items shared with Castilian (Posner and Green, 1993; Beswick, 2007). These in turn are balanced by movements to assimilate Galician with Portuguese.

François (2011) describes the strong tendency for languages spoken in the Torres and Banks islands of northern Vanuatu to diverge in the forms of their words, resulting in a pattern where closely related languages that would be expected to have high levels of cognacy, instead exhibit highly distinctive vocabularies.

Perhaps the most extreme example of change aimed at increasing the difference in one's own speech is that of the Uisai dialect of Buin, a language spoken in Papua New Guinea on Bougainville island. Laycock (1982:34) reports that Uisai shows diametrically opposed noun categories to other dialects. The markers for category 1 in Uisai occur only with category 2 elsewhere, and vice-versa. In this particular parameter these dialects are significantly more different than would be expected by chance.

The desire to differentiate languages in this way doesn't necessarily imply hostility or antagonism. Laycock also reports an opinion from the Sepik region of Papua New Guinea: *it wouldn't be any*

good if we all talked the same, we like to know where people come from.

One of the reasons for the current work is to create the tools which might let us see whether these efforts to change languages, for social or political reasons, actually have a lasting effect on the vocabulary, or whether they are at best ephemeral (see eg. Thomason & Kaufman 1988, Ross 2007, Aikhenvald 2002; and François 2011, Arnal 2011 on differentiation).

3 Evolutionary Change in Relations

In this section, we explore the formal model that we will use to distinguish normal, independent change from contact-induced change. The first step is to model languages as a bundle of relations. Modelling language in this way is not new, but is rarely made explicit.

3.1 Language as a Bundle of Relations

Much language structure can be expressed as relations between different spaces. For example, the lexicon can be regarded as a relation between the space of meanings available in a language and the phonological forms of morphemes expressing that meaning. There can be meanings represented by multiple forms, such as **ready** and **prepared**, or forms with multiple meanings such as **fire** in the sense of **burning** or **terminating employment**.

Another language relation maps phonemes-in-contexts to phones that can realise them. Phonemic distinctions may collapse in some contexts, such as with the final devoicing of obstruents in Polish, so that distinct phonemes are realised with the same phone. Likewise, the same phoneme, even in the one context, may be realised by multiple phones; the Portuguese phoneme /ʁ/ is realised as [ʁ], [ʀ], [ɣ] or even [r], with multiple possible realisations even for the one speaker.

So both the lexicon and phonetic realisation can be modelled with relations.

3.2 Primitive Changes on Relations

If some important language structures are relational, an interesting question is what sort of evolutionary changes can effect these relations. This subsection explores a number of minimal changes which can effect relations. To the best of our knowledge, this is the first time that language changes have been characterised this way. The starting point is a simple relation between a domain and a range, as shown in figure 2.

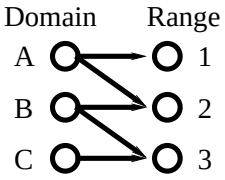

Figure 2: shows a relation from a small domain to a similarly-sized range.

The first kind of change is a global substitution, see figure 3. This is where a change of permutation or merger applies to elements of either the domain or the range. All of the pairs which contain the affected elements are modified, hence the name.

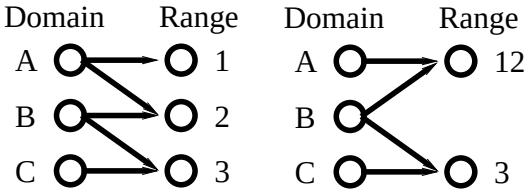

Figure 3: shows a global substitution: range elements 1 and 2 are merged, preserving all links. It is called a **global substitution** as every link with 1 or 2 in the range now has 12 as its range element.

Modifications of the phonetic relation can be of this kind. For example, when Gaelic – both Irish and Scottish – merged [ð] into [ɣ], the change affected both lexical /ð/ in closed class words, such as the preposition <dha>, /ðə/, *to*, as well as lexical /ð/ in open class words such as <duine>, /duɲə/, *person*. This was a global substitution.

More frequently met are small changes, we will call **local mutations**. These involve either the insertion of a single link, or the deletion of a single link.

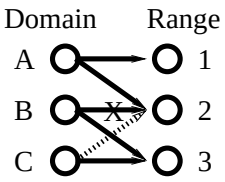

Figure 4: shows two separate local mutations in a relation: a deletion marked by an X on the link, and an insertion shown as a dotted arrow.

Gloabl changes can be expressed as local changes combined with relation composition.

The lexical relation associates meanings with the phonological forms, which may take the form phonemes in contexts. The phonemic map then projects these onto their phonetic realisations.

If a single link in the phoneme realisation map is dropped, then all lexical meanings expressed using that phoneme-in-context can no longer realise it with that phone. If a single link is added to the phonetic relation, then all lexical meanings expressed using that phoneme-in-context can now realise it with the new phone. This multiplier effect on changes means single sound changes can have a disproportionate effect on the similarity of cognate forms in two languages. Ellison and Kirby (2006) presented a similarity measure which bypasses this superficial difference: pairs of domain elements are compared for the similarity of the corresponding sets of range elements, and these similarity values are then compared cross-linguistically. This measure mitigates the effect of global substitutions.

The iterated application of local mutational changes to language structures is called **drift**. In traditional models of language history, it is the primary mechanism for explaining difference, while the shared parent language is the primary explanation of similarity.

3.3 Contact-induced change

So far, we have only looked at change arising in independent relations. Change, in language at least, is often the result of contact with the corresponding relational structure in another language. Figure 5 shows two relations between the same domain and range, superimposed. Later diagrams will use this same superimposed representation in describing contact-induced changes.

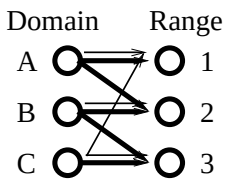

Figure 5: shows two relations simultaneously: the links from one are shown with thick arrows, those from the other with thin. Links common to both relations are doppels.

In considering contact-induced change, it is worth noting that the change need not be symmetrical between the languages involved. If one language is spoken by a dominant, larger population, it may see no reason to differentiate itself from the language of a smaller community. The smaller community may feel that language differentiation is a way to protect its identity. Whatever the reason, we shall call the relation undergoing differentiation the **assimilating** or **differentiating relation**, and the relation it is pushing away from, or pull towards, the **reference relation**.

Contact-induced assimilation or CIA can consist of the insertion of a new link into the relation, or the deletion of a link in the relation. As assimilation is about making the relations more similar, so insertion applies to create doppels where the reference relation has a link and the assimilating relation does not. Likewise assimilation applies to delete links where the reference relation does not have a link but the assimilating relation does. Examples of this kind of assimilation are shown in figure 6.

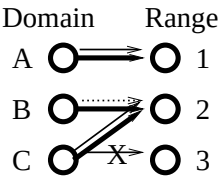

Figure 6: shows contact-induced assimilation (CIA) as an insertion shown as a dotted line and a deletion marked with an X. Existing links of the assimilating relation are shown thin, while those of the reference relation are shown thick. In CIA, links are more likely to be inserted to make a doppel, and deleted where no doppel exists.

The reverse is true in cases of contact-induced differentiation – see figure 7. The differentiating

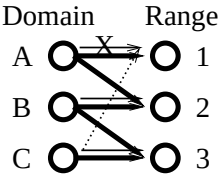

Figure 7: shows contact-induced differentiation (CID) in the form of an insertion shown as a dotted line and a deletion marked with an X. Existing links of the differentiating relation are shown thin, while those of the reference relation are shown thick. In CID, links are more likely to be deleted if they have a doppel, and inserted where they do not.

relation is more likely to delete a link which is half of a doppel than delete other links. Likewise, it is

more likely to create a link where there is none in the reference relation, rather than borrow a link from it.

4 When can CIA/CID be Inferred?

This paper addresses the question: how much data is required to distinguish cases of contact-induced change from similarity due to a common ancestor and differences due to drift? The question will be addressed in terms of relations and the types of changes covered in section 3.2 and section 3.3. To render the problem tractable, we need an additional assumption about the lexical relations: they have the form described in section 4.1.

4.1 RPOFs

We restrict lexical relations to RPOFs. An **RPOF** is a **reverse of a partial onto function**, in other words, a relation such that each element of the domain participates in at least one link, while each element in the range participates in at most one link. An example of such a relation appears in figure 8. If the lexical relation in a language is

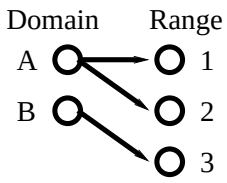

Figure 8: shows an RPOF relation. In RPOFs, each element of the domain has at least one link, while each element of the range has at most one link.

an RPOF, then each meaning is expressible with at least one morphemic form, and each potential form expresses exactly one meaning, or else is not used in the language. In other words, the language has no homophones.

This assumption is usually only mildly inaccurate. For some languages, however, such as Chinese, mono-syllabic morphemes are frequently homophonous. The analysis presented here may fail for languages of this kind.

The advantage of using RPOFs is that their structure can be summarised by a cardinality function – a partial function from natural numbers to natural numbers. This function associates with any cardinality of range subset the number of elements of the domain which associate with a range set of exactly that size. For example, the relation shown in figure 8 maps one input onto two outputs, while it maps the second input to a single output. Thus its cardinality function is $\{2 : 1, 1 : 1\}$. Such specifications completely characterise an RPOF relation upto permutation of either the domain or range.

One of the effects of assuming RPOF structure for the lexical relation is that we do not allow the sole link from any domain element to undergo deletion. This is because all domain elements must retain at least a single link. For the lexical relation, this has the fairly likely consequence that the sole morpheme representing a meaning is unlikely to be lost, while if there are multiple synonyms, one might fall out of use.

4.2 Pairs of RPOFs

When we are comparing RPOFs evolved from a common parent, we can characterise their relationship, upto permutation of the domain and range, by frequency counts over triples. The triples are numbers describing how many elements of the range a domain element links to: solely in relation 1, in both relations (ie, the number of doppels), and solely in relation 2. For each triple, we count the number of domain elements which have the correspondingly sized projections on the range. This kind of summarisation allows us to describe the similarity of two lexical relations with a few hundred numbers if we limit ourselves to, say, domain elements linking to at most 10 range elements in either relation.

4.3 Significance Testing

It easy to evaluate the posterior likelihood of a set of data associating a counting number with each triple, $D \in \mathbb{N}^{Triples}$, given a model $M \in Dist(Triples)$ in the form of a distribution over triples. The triple associated with each domain element is assumed to be the result of independent processes – in other words, we assume that the number of doppel and non-doppel forms associated with a meaning is independent of the numbers associated with other meanings.

$$P(D|M) = \prod_{t \in Triples} M(t)^{D(t)}$$

We can evaluate the likelihood of one model M_1 generating data at the frequencies produced by a second model M_2. The posterior probability of the data relative to the second model is shown

in equation (1), while the probability of generating that data from the model which did indeed generate it is shown in equation (2).

$$P(M_2|M_1) = \prod_{t \in Triples} M_1(t)^{M_2(t)} \quad (1)$$

$$P(M_2|M_2) = \prod_{t \in Triples} M_2(t)^{M_2(t)} \quad (2)$$

The likelihood ratio, i.e. the ratio of posterior likelihoods of M_2 and M_1, is shown in equation (3).

$$\frac{P(M_2|M_1)}{P(M_2|M_2)} = \prod_{t \in Triples} \frac{M_1(t)}{M_2(t)}^{M_2(t)} \quad (3)$$

This ratio expresses the amount of information we are likely to gain about which distribution is correct as a result of looking at a single data item. In terms of RPOF relations, this single data item is the triple of counts for relation-1-only, doppels, and relation-2-only associated with a meaning. If, as assumed above, the counts associated with each domain element are independent, then the likelihood ratio is raised to the power of the number N of items seen.

$$\frac{P(M_2|M_1)}{P(M_2|M_2)}^N = [\prod_{t \in Triples} \frac{M_1(t)}{M_2(t)}^{M_2(t)}]^N \quad (4)$$

To establish a chance prediction at $p < 0.05$, we merely need to know that $P(M_2|M_1) < P(M_2|M_2)$, and then determine the minimum level of N for which the ratio in equation (4) is less than $1/19$. This number of items generated from the target distribution would allow it to be distinguished from chance at a ratio of $19:1$.

Determining the correct value for N here is a general problem known as **power analysis**. For standard experimental designs and corresponding statistics, the power analysis can be found in many texts, such as that by (Bausell and Li, 2006), and many computing libraries such as the **pwr** library for power analysis in R (see `http://cran.r-project.org/web/packages/pwr/`). Where the model design is as complex as that described here, the power analysis must be constructed from first principles.

It is often easier to work with this quantity in informational rather than probabilistic form, where it takes the form shown in equation (5).

$$\begin{aligned} & -\log \frac{P(M_2|M_1)}{P(M_2|M_2)} \\ = & -\sum_{t \in Triples} M_2(t) \log \frac{M_1(t)}{M_2(t)} \end{aligned} \quad (5)$$

The quantity in equation (5) is the well-known **Kullback-Liebler divergence** $D_{KL}(M_2||M_1)$ of the two distributions, also known as the **discrimination information**. Significance is achieved when this value multiplied by the number of data items is greater than $log_2(19) = 4.2479$.

4.4 Models with and without Context-Induced Change

The construction of the no-CIA/CID and the with-CIA/CID distributions makes use of four parameters.

In the non-context model:

> **insertion** of a link combines the probability α of making a change at all for any given domain element, with the probability $\beta/(1+\beta)$ that the change will be the addition rather than deletion of a link, into a likelihood of adding a link per domain element of $\alpha\beta/(1+\beta)$.
>
> **deletion** of a link combines the probability α of making a change at all for any given domain element, with the probability $1.0/(1+\beta)$ that the change will be to a deletion, with the number m of links to select from for that domain element, so the probability of deleting any of those links is $\alpha/(m+m\beta)$.

In the case of CIA/CID, we only consider the impact of contact on deletion. The per-link probability of deletion $\alpha/(m+m\beta)$ is modified by a parameter γ indicating how strong the effects of contact are. Positive γ brings about CIA – with shared links less likely to be dropped than others, while negative γ develops CID – shared links are more likely to be dropped than others. The probability of dropping any given doppel link from a given range node is $(1-\gamma)z$, and of any unshared link is z where n_d is the number of doppel links from the domain element, and n_u the number of

unshared links in the differentiating relation, and z is given in equation (6).

$$z = \frac{\alpha}{((1-\gamma)n_d + n_u)(1+\beta)} \quad (6)$$

4.5 Simulation Results

The above model was used to generate distributions over triples for non-CIA/CID relation pairs, and relation pairs with additional CIA/CID processes. The number of iterations of the mutation process with or without CIA/CID was fixed at 100 in creating the generating distribution M_2. The parameter α was fixed at 0.1 and β at 0.5. The value for β was chosen to approximately reproduce the single-language distribution of range-set sizes for Castillian as computed from the Spanish wordnet. The bias parameter γ was varied from -0.5 to 0.5 in steps of 0.1. For each level of bias, a search was made over non-CIA/CID distributions at different depths from the common ancestor – this is the parameter N – until the distribution with the least K-L divergence from the generated distribution was found. This found distribution M_1 represents the null hypothesis, that the data arose without CIA/CID bias.

The number of data items needed to achieve significant recognition of the presence of CIA/CID bias is $4.2479/D_{KL}(M_1||M_2)$. The results for various levels of γ are shown in figure 9.

γ	N	S	D
-0.5	118	3128	0.091
-0.4	115	4364	0.096
-0.3	111	6839	0.101
-0.2	108	13800	0.107
-0.1	104	47378	0.114
0.1	95	30913	0.133
0.2	90	7331	0.145
0.3	85	2793	0.160
0.4	79	1278	0.178
0.5	72	654	0.203

Figure 9: Tabulation of number S of data items needed to achieve significance and the number of iterations N of the best non-CIA/CID model, and fraction of doppels remaining D, against CIA/CID bias parameter γ. Note that fewer data items are needed to recognise significant assimilatory bias (positive values for γ) than differentiating bias (negative values of γ) at the same strength.

5 Conclusion

This paper has looked at different ways that relations may evolve from a common parent structure. They may undergo local mutational changes, global substitutions, independent changes, or those triggered by contact with other relations. In one class of relations, with reasonable assumptions, it is clear that a large, but possible, amount of data needs to be adduced to ascertain that CIA and/or CID have occured, rather than just shared origin and independent drift.

In historical linguistics, this opens the door, for testing whether the impressionistic accounts of CID are reflected in the distributional properties of the languages concerned. It may also be possible to circumvent the onerous data requirements, by bringing in data from multiple independent relations within the language, such as those defining morphological structure and phonology, as well as the lexicon.

As mentioned in the introduction, this work is part of a larger programme by the authors to develop statistical tools able to show that CID has taken place, if it has. This work is partly driven by the need to account historically for the low cognacy but high structural similarity between nearby Australian languages. In the Daly River area, adjacent languages with very similar phonology, syntax and morphology show remarkably low cognacy counts, often around 8% (Harvey, 2008). One possible explanation for this is a powerful CID imperative acting over a short time depth to differentiate the vocabularies of the languages. The result presented in this paper suggests that with sufficient lexical data, direct statistical evidence could be found if this is indeed the correct explanation.

There are potential uses for this work beyond historical linguistics as well. The model might assist in some cases of plagiarism detection, for example, where two students worked together on an assignment, and then set out to deliberately differentiate them by altering vocabulary. Similar analysis of documents might reflect other reasons for reworking a text, such as to give it a new identity for a new setting.

References

A. Aikhenvald. 2002. *Language contact in Amazonia*. Oxford University Press, Oxford.

Antoni Arnal. 2011. Linguistic changes in the catalan spoken in catalonia under new contact conditions. *Journal of Language Contact*, 4:5–25.

R. Barker Bausell and Yu-Fang Li. 2006. *Power Analysis for Experimental Research: A Practical Guide for the Biological, Medical and Social Sciences*. Cambridge University Press, March.

Jaine E. Beswick. 2007. *Regional nationalism in Spain: language use and ethnic identity in Galicia*. Multilingual Matters.

Karl Brugmann. 1884. Zur frage nach den verwandtschaftverhltnissen der indogermanischen sprachen. *Internationale Zeitschrift fr allgemeine Sprachwissenschaft*, 1:226–56.

T. Mark Ellison and Simon Kirby. 2006. Measuring language divergence by intra-lexical comparison. In *ACL*, pages 273–80, Sydney.

Thomas Finkenstaedt and Dieter Wolff. 1973. *Ordered profusion: studies in dictionaries and the English lexicon*. C Winter.

Alexandre François. 2011. Social ecology and language history in the northern vanuatu linkage: a tale of divergence and convergence. *Journal of Historical Linguistics*, 1:175–246.

Mark Harvey. 2008. *Proto-Mirndi: a discontinuous language family in northern Australia*. Pacific Linguistics, Canberra.

Sir William Jones. 1786. The third anniversary discourse, delivered 2nd february, 1786: on the hindus. *Asiatick Researches*, 1:415–31.

Donald C. Laycock. 1982. Melanesian linguistic diversity: a melanesian choice? In R.J. May and H. Nelson, editors, *Melanesia: beyond diversity*, pages 33–38. Australian National University Press, Canberra.

Rebecca Posner and John N. Green. 1993. *Bilingualism and Linguistic Conflict in Romance*. Walter de Gruyter.

Malcolm D. Ross. 2006. Metatypy. In K. Brown, editor, *Encylcopedia of language and linguistics*. Elsevier, Oxford, 2nd ed edition.

Malcolm Ross. 2007. Calquing and metatypy. *Journal of Language Contact, Thema*, 1:116–43.

August Schleicher. 1861. *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*. Hermann Bhlau, Weimar.

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

Sarah Grey Thomason. 2007. Language contact and deliberate change. *Journal of Language Contact, Thema*, 1:41–62.

Carrasquer Vidal. 1998. Untitled post in 'Cladistic language concepts' thread, HISTLING mailing list, Oct.

Roger Wright. 1998. Untitled post in 'Cladistic language concepts' thread, HISTLING mailing list, Oct.