The structural impact of morphological borrowing and convergence

A dissertation submitted to The University of Manchester for the degree of Master of Arts in the Faculty of Humanities

2014

Deepthi Gopal
School of Arts, Languages and Cultures

Contents

Abstract 3

Declaration & Intellectual Property Statement 4

Dedication & Acknowledgements 5

1. Introduction 6

2. Theoretical and historical preliminaries 8

3. Examples and case studies 12

    3.1 Faithfulness to the source 12

        3.1.1 Straightforward matter copying 12

        3.1.2 *Recombination and reinterpretation* 19

        3.1.3 Related languages and phonological analogy 22

        3.1.4 Pattern replication 24

    3.2 Disturbances to the morphological system 28

        3.2.1 Undisturbed systems 29

        3.2.2 New functions consistent with existing patterns 30

        3.2.3 New functional categories 31

        3.2.4 Innovations: new affixation and structure 33

4. Conclusions and outlook 41

5. References 43

Words: 14,111

Abstract

This work provides a selection of novel examples of morphological convergence and borrowing, with a particular focus on languages in the Austronesian, Austro-Asiatic, Dravidian, and Tibeto-Burman families. We investigate the systemic effects and typological features of morphological borrowing, with a focus on the 'recognisability' of a borrowed element from the perspective of the source language, on the interaction of morphology and other strata of the grammar, and on the ability of a given loaned element to trigger large-scale functional or structural alterations within the recipient language; we additionally describe contact effects on some non-affixational morphological processes. We observe that the borrowing of the category of numerical classifiers patterns strongly with the lexical copying of cardinal numbers; that the current morphological shape of the Chamic languages is likely to be best understood via the interface of prosody and morphology; that it is possible for relatively 'minor' derivational borrowings to nonetheless promote significant systemic alterations, and that reduplicative processes may be of interest in the broader context of contact linguistics. We also note the importance and underexamination of interface issues in discussions of contact and convergence.

Declaration

No portion of the work referred to in the dissertation has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Intellectual Property Statement

i. The author of this dissertation (including any appendices and/or schedules to this dissertation) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this dissertation, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has entered into. This page must form part of any such copies made.

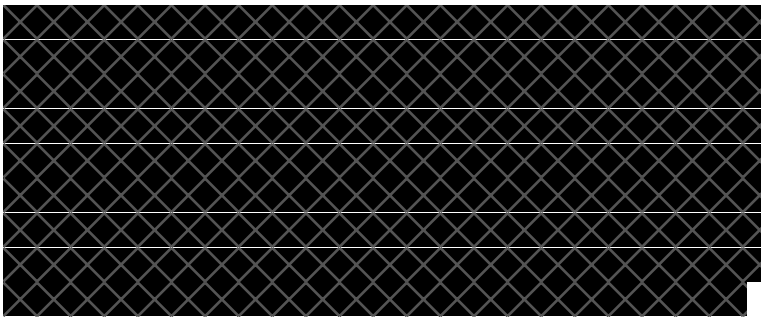
iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the dissertation, for example graphs and tables ("Reproductions"), which may be described in this dissertation, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this dissertation, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see *http://documents.manchester.ac.uk/display.aspx?DocID=487*), in any relevant Dissertation restriction declarations deposited in the University Library, The University Library's regulations (see *http://www.manchester.ac.uk/library/aboutus/regulations*) and in The University's Guidance for the Presentation of Dissertations.

Dedication

This dissertation is dedicated to my grandfather, Professor R. P. Pakshirajan, who is responsible for all my successes and forgiving of all my failures.

Acknowledgements



1. Introduction.

How does language contact affect the organisation of morphology? Although it is readily apparent that languages have the capacity to influence each other in a variety of domains, the outcome of any given process of 'language contact' can often be difficult to predict and explain: Attempts to systematise and find structure within the diverse effects of language contact often focus on the issue of its origin, motivation, and recognisability: what causes a language to borrow some form or function from another? How do we, after the fact, identify these occurrences and account for them?

A large portion of the discussion of borrowed morphology, in particular, focuses on enumerating the types of morphemes with a general tendency towards being borrowed, evaluating the factors that enable or restrict borrowing, and investigating the aspect of the source morphological element being borrowed (form, function, or some combination); or, ultimately, presenting universal borrowability hierarchies of various features. While these aspects are all of interest, we particularly decenter the perceived 'rarity' or frequency of morphological borrowings – especially given the inherent selection bias involved in choosing the most analytically interesting cases – focusing instead on those situations in which relevant contact effects can be demonstrated with some clarity: then, how do the borrowed morphological elements relate to the morphology of the recipient (and possibly donor) languages?

In this work, we aim to evaluate the *impact* of various situations of morphological borrowing on the underlying structures of the recipient languages, and additionally to call back to the 'recognisability' of borrowed morphology within the donor: to what extent is faithfulness as a constraint or requirement involved in language contact processes? The existing literature on morphological borrowing alone spans an extensive range of empirical documentation, and it would be highly impractical to attempt to address this vast array of situations in one brief work. We focus, in our selection of data, on presenting a small selection of interesting examples which are comparatively infrequently referenced in the existing discussions of language contact, and place particular emphasis on the languages of South, South-East, and Central Asia: the Austronesian languages (Malay; Cham; Madurese), the Austro-Asiatic

languages (Aslian: Jahai, Semelai; Munda), the Dravidian languages (Kurukh, Malto, Kannada), the Turkic and Tungusic languages of Siberia (Yakut, Evenki), and the Tibeto-Burman languages of the Himalayan region (Kiranti; Tamangic).

2. Theoretical and historical preliminaries.

Perhaps the earliest analytical approach to the investigation of language contact is that of Weinreich (1953): influential both in his codification of the true comparative approach to language contact, and in the selection of the perspective of the individual speaker's repertoire as the ultimate locus of language contact. Within the specific context of this work, we note that much discussion of contact pre-Weinreich additionally proposed that inflectional borrowing must necessarily be entirely nonexistent (see e. g. Sapir 1921); this is, evidently, inaccurate. Weinreich did note the difficulty of borrowing both bound morphology relative to free morphemes, and inflectional morphology relative to derivational morphology (additionally Thomason & Kaufman, 1988; Matras, 2009); this type of scale informs a large portion of subsequent work on the likelihood and licensing of morphological loaning, in which 'ease' of borrowing is linked to the threshold of 'intensity' of contact that might be required to produce the corresponding change. (Our focus in this work is not on the derivational-inflectional scale itself, although it has some importance in the matter of the salience of particular morphology to the overall system.)

Perhaps most crucially, we briefly establish the distinction between *matter replication* MAT and *pattern replication* PAT, which are referenced throughout and expanded upon in our typological analyses. This follows ultimately the terminology proposed by Matras & Sakel (2004, 2007; Sakel 2007), in delineating the distinction between the borrowing of morphological material (with attendant phonological form) or MAT, and pattern, or PAT – consisting of e. g. the formal organisation and structure, the distribution, and the semantic/grammatical function of a particular element. This definition has several precursors: Weinreich's 'transfer of elements' versus 'interference' (without transfer) provide a rough correspondence, as does Johanson's 'global copying' versus 'partial copying' (the naive interpretation of this system presumes that all MAT-copying implies PAT-copying, which is frequently although not absolutely true). We note here that *pattern* does concatenate the elementary structure and the semantic function of an element; while we adhere to this terminology throughout, we reference the 'finer' distinctions wherever they are salient i. e. in a situation where the distinction between affixational structure and syntactic structure becomes relevant.

We will make significant reference to the concept of the 'morphological

system': this is deliberately quite loosely and informally defined, and intended simply to denote the collection of aspects which would be collected under *pattern* in our initial framework: thus, both the structural-affixational elements of morphology, and the functional aspects including the existence of a particular inflectional category or instance of semantic marking. (We use 'source' and 'recipient' language on occasion to refer to the languages involved in the contact situation under discussion, for sheer terminological convenience.)

The theoretical preliminary which we have not yet addressed involves the specific constraints on our discussion – what qualifies as an example of morphological borrowing by our definition? We must clearly distinguish the acceptance of borrowed morphology from the preservation of unproductive morphological features of donor languages within borrowed lexical items. As an illustration, Thai pairs such as ตรวจ *truat* 'investigate, examine'/ตำรวจ *tamruat* 'policeman', or เดิน *doen* 'walk'/ดำเนิน *damnoen* 'procession' clearly suggest a possible derivational process involving the infixation of *-aN-* (Matisoff 1996: 1594). However, in fact, these items are direct loans of Khmer ត្រួត *truot* 'control'/តម្រួត *domrut* 'policeman, official' and ដើរ *dae* 'walk'/ដំណើរ *damnae* 'trip, journey, procession' (reflecting a historically productive morphological strategy in Khmer), with the complex forms having been borrowed into Thai with Khmer infix intact.[1] We cannot then say that the Khmer derivational process responsible for these items has been borrowed into Thai: although Thai forms displaying this pattern are numerous (Huffman 1986 provides a more complete list), they may all be traced directly to Khmer, and there is no evidence that the *-aN-* infix is used with native Thai lexicon, or even with Khmer loans outside the typical domains expected in Khmer. We may contrast this case to the example of the Turkish agentive suffix *-çi* in Iraqi Arabic (Matras 2009: 209): having been introduced in loans such as agšamçi 'night-watchman' (Turkish *akşam* 'evening' + *-çi*), it is extended through the lexicon and applied both to existing Arabic words, as in *ke:fçi* 'party-goer' (*ke:f* 'fun'), and to new

[1] Supporting this conclusion, several examples exist in which Thai has the Khmer *derived* form with *-aN-*, but not the base or other related forms: Khmer បាក់ *bak* 'break'/ផ្នែក *phnek* 'section' > Thai แผนก *phaenak* 'division'; Khmer ស្នាក់ *snak* 'stay'/សមនាក់ *samnak* 'abode' > Thai สำนัก *samnak* 'residence, bureau, office'. (partial adaptation of Huffman 1986: 202-203.)

loans (*go:lçi* ‘goalkeeper’). The difference in behaviour from *-aN-* in Thai is clear; in terms of process, we might describe Iraqi Arabic as having performed a morphological analysis of the original Turkish loans – in order to identify the affix and incorporate it into the repertoire – which has no parallel in the Thai case. Thus, we reiterate ultimately that the true borrowing of morphology requires the borrowed element or pattern to diffuse beyond the original context in which it entered the lexicon of the recipient language[2].

Another possible ambiguity arises in the fact that not all seemingly imported morphemes are necessarily borrowed on the ‘tier’ of morphology, or borrowed already grammaticalised; as an illustration of this, we investigate the status of novel bound morphemes in a recipient language B whose material derives from a source language A, but from free lexemes in A rather than from ‘similar’ morphological elements. In Burmese, *əkʰa* from Mon ‘time’ occurs as a derivational suffix in e. g. *mo-əkʰa* ‘rainy season’: what is it that separates a borrowing of this type from the copying of bound morphology? Several examples arise in Meitei, a Tibeto-Burman language of India; we briefly outline the situations of the nominaliser *jat* and the verbal marker *doy*, which represent distinct mechanisms of morphological innovation in the language. *jat*, borrowed from Indo-Aryan ‘caste’ (occasionally generalised to ‘type’), acts as a well-integrated suffix within Meitei, in combination with native vocabulary, and serves to indicate certain types of evidentiality (Chelliah 1997: 156):

Meitei:

	ma	*əy-gi*	*ka*	*-də*	*tʰək*	*-ləm*	*-lə*	***-jat***	*-lə*
(1a)	3SG	I-GEN	room-LOC	smoke-EVD-PERF	**-TYPE**	-INT			

‘Could it be that he smoked in my room?’

	məsi	*pʰu*	*-ləbə*	***-jat***	*-ni*
(1b)	this	beat-HAVING	**-TYPE**	-COP	

‘This looks like it has been beaten.’

jat or a similar form cannot serve this function in any of the Indo-Aryan languages. What seems to have occurred in Meitei is the borrowing of *jat* into the lexicon with a meaning of ‘type’ or ‘class’, *followed* subsequently by its grammaticalisation as an affix – with Indo-Aryan form, but an ultimately novel morphological character. *doy*, which marks intentionality in Meitei, has an Indo-Aryan origin which is less

[2] On the level of the speaker, it may require some act of processing and analysis which is itself likely predicated on a knowledge of the ‘source’ language and its morphological structure – we discuss this in chapter 4.

immediately apparent: it is ultimately a contraction of *dərkar oy*, Hindi *dərkar* ‘necessary’ + *oy* ‘be’.

Meitei:
nəŋ cət-pə ***dərkar oy***
(2a) you go-NOM **necessity is**
‘You must go.’ (Chelliah 1997: 170)

*əy-gi phi-du han-nə ləy-həw-**doy**-ni-ko*
(2b) I-GEN cloth-DDET first-ADV buy-START-**INTENT**-COP-TAG
‘I plan to buy some cloth for myself, OK?’ (Chelliah 1997: 233)

The content of *doy* is then the synthesis of a borrowed Indo-Aryan lexeme and an existing Meitei one. We therefore see that both *jat* and *doy* contain Indo-Aryan material, whether faithfully replicated or reinterpreted – do these fall under the purview of morphological borrowing in our analysis?

We suggest here that it is necessary to draw the distinction between the *act* of borrowing and the *result* thereof; the act of borrowing ends once direct engagement with the donor language ends. Comparing the acquisition of these affixes in Meitei with the case of *-çi* in Iraqi Arabic, all three processes have ultimately produced novel morphemes in the recipient language; however, only the third involves the explicit identification and reproduction of a morphological element of the source. The ‘act’ of borrowing for *-çi* involves both the assimilation of lexical items, and the generalisation of morphological information found within in order to reproduce the suffix; for *-jat,* only the first takes place. While the Meitei affixes described provide an interesting example of contact-induced change, we do not see here the active process of analysis that we describe in the example of Iraqi Arabic above: the grammaticalisation of *-jat* is not procedurally distinct from the grammaticalisation of any other suffix whose source is within the native Meitei lexicon. For this reason, we restrict our sample to instances of borrowing that involve some degree of implicit awareness of the morphological structure of the source.

3. Examples and case studies

In this section, we enumerate and, if necessary, justify various cases of morphological transfer and convergence, broadly evaluated within two functional schemes:

1. The ultimate relationship between the donor prototype and the 'output' morphological element in the recipient language.
2. The effect of the act of borrowing on the morphological system of the recipient language.

3.1 Faithfulness to the source.

We begin by outlining some examples of morphological borrowing, roughly organised on a scale from 'the most faithful to donor language' to 'most radically altered relative to donor language' – ignoring modifications performed on a level external to the recipient language's morphonology (e. g. the application of assimilatory rules of phonology, which do not change the underlying nature of the borrowing), this continuum can itself be investigated in terms of the degree of material and functional/structural copying seen. Broadly, this measures the 'recognisability' of the use of borrowed morphology in a language: would speakers of the source language recognise any aspects of the recipient language's eventual output?

3.1.1 Straightforward matter copying.

In a classification predicated on fidelity, it seems natural to identify those cases in which a concrete morpheme is transferred from donor to recipient language, with form and function 'pattern' as minimally modified as possible – such borrowings are then the *most* faithful to the donor, retaining both the matter, or phonological shape, and the pattern of the source material. We do not attempt to provide an exhaustive inventory of such copies here, but rather describe those

cases which are less well-examined in the contact-specific literature, or which illuminate some interesting theoretical point.

The copying of nominal derivational morphology seems to be the commonest form of morphological borrowing, perhaps in keeping with the general tendency of languages to permit the propagation of diverse functional devices in this domain (ultimately, due to the need for a wide set of subtly-differentiated referential meanings and semantic categorisations within the set of nouns – Matras 2009: 210). More generally, there is a significant cross-linguistic preference, within the borrowing of bound morphology, for derivational borrowings over inflectional borrowings (Weinreich 1953), as in chapter 2; it seems sufficient to briefly describe only a small sampling of derivational copies to avoid redundancy – as there seems to be little procedural variation possible in the borrowing of derivational morphology. In Ilocano (Iloko), as in many other Philippine languages, Spanish nominal suffixes are productive along with the Spanish infinitive in *-ar/er/ir*: the locative *-eria*, in *labandéria* 'laundromat' or *pansiteria* 'noodle restaurant'; the agentive *-ero/-era* (displaying a pattern of gender marking also imported from Spanish, and re-examined in 3.2) in *karaykayero* 'person who rakes' or *partera* 'midwife', reflecting an apparent cross-linguistic tendency towards the copying of similar markers (Rubino 2005: 346). In the Balkan Turkic language Karaim (Csato 2012: 373), derivational affixes are copied from Slavic languages: e. g. *-ski* deriving adjectives as in *karayski* 'Karaim', *rabbanski* 'Jewish'; *-ka/-tsa* marking the feminine/diminutive as in *dost-tsa* 'female friend'.

The Chamic branch of the Austronesian languages is notable for its significant and sustained early historical contact with various Mon-Khmer languages; all modern Chamic languages continue to be spoken in areas with high populations of Mon-Khmer speakers, save only Acehnese (in Indonesia) and Tsat (on Hainan). Thurgood (1999: 239) reconstructs the 'negative imperative' **bɛʔ* to Proto-Chamic (Acehnese *bɛʔ*, Jarai and Bahnar *beʔ*), and notes that it does not appear elsewhere in Austronesian (Roglai has now replaced this element with *ɗəŋ*, borrowed from Vietnamese). The reconstructed Proto-Chamic instrumental infix in **-ən-* (e. g. Cham *dak* 'to pile' > *da-n-ak* 'a pile') is also a Mon-Khmer borrowing, with clear parallels to the Khmer derivational process mentioned in chapter 2.

Several examples of morpheme transfer arise in the Dravidian languages, both in derivation and in inflection. Kannada and Telugu derivational morphology show significant borrowings from Indo-Aryan, enumerated in detail in section 3.2.5; these are the most significant morphological copies evident in the 'literary' Dravidian languages. Intra-Dravidian borrowings are attested, although not exhaustively studied: in modern Telugu (Dravidian: South Dravidian II), the conditional is formed with *-ite:/-te:/-ʈe: –aɳ* 'say'/*–aɳ-ʈe:-ni* 'if you said', *ce:si-* 'do'/*ce:si-te:-n* 'if you had done' (Krishnamurti 2003: 337). This does not reflect a wider Dravidian pattern; in fact, no conditional forms in modern Dravidian languages can be definitively traced back to any Proto-Dravidian form, and it would appear that the genesis of conditionality-marking in the various Dravidian languages is due to several individual developments post-dating most major splits in the family (compare *-al/-el* in Tamil and Malayalam, *-are* in Kannada, *-is* in Pengo, *-iŋa* in Konda, and *-ek* in Gondi). In various Dravidian languages classified as Central Dravidian and occupying minoritarian positions in majority Telugu-speaking areas, however, we find a small concentration of markers highly reminiscent of those in Telugu: consider Kolami *kak-te:* 'if (one) does', and Naiki *si-t-an-te* 'if I gave'.

The northern portion of the Dravidian family displays a particularly wide range of inflectional copies from Indo-Aryan. In Malto, spoken primarily in Indo-Aryan-majority areas of East India (Orissa, Bengal, Jharkhand), the genitive *-ki* is borrowed from Hindi/Bengali and ultimately from Persian: thus *male-ki* 'of the man', *maqe-ki ʈeɖuð* 'the boy's hand' (Krishnamurti 2003: 234). In both Malto and Kurukh (a closely-related language of the same region), Krishnamurti (2003:237) describes a set of non-native instrumental markers: Kurukh *-tri:* e.g. *eŋg-tri:* 'by me', Malto *-t/-et/-it*, e.g. *maler-it* 'by men', *male-t* 'by the man', *ki:r-et* 'because of hunger'. These seem ultimately to arise from the forms common through the Munda languages spoken in the area (in the case of Kurukh, via the common Dravidian *t/tt* > *tr* sound change applied to the historic alveolar): Juang *aro-te* he-INS 'by him', Ho *ɖaɳɖa-te* stick INS 'with the stick', *am-ete* you-INS 'by you', Mundari *hake-te* axe-INS 'with the axe' (Anderson 2006: 22-23). (This ultimately reflects part of the complex contact situation in Eastern India: in Santhali, one of the larger Munda languages of India, a significant series of enclitic postpositions is borrowed from Indo-Aryan (Anderson 2006: 25): *–são* 'with', *–hɛn* 'near', *–ləgit* 'for'.) In Kurukh, the perfective is formed in *-a:r*, seemingly derived from Hindi *-kar*

(e. g. *so:-* ‘to sleep’, *so:kar* ‘having slept’): thus *ci-* ‘to give’/*ci-a:r* ‘having given’.[3] Gardani (2012: 83) suggests that both Kurukh and Kharia (a Munda language of the same area) borrow the feminine nominal marker *-i* from Hindi: thus Kharia *cəngna*/*cəngni* ‘rooster’/‘hen’, Kurukh *a:las*/*a:li* ‘boy’/‘girl’; this indeed seems to be the case for Kharia, but we may question the robustness of this conclusion for Kurukh given the tendency towards *-i* as a marker of the femi nine across the Dravidian languages (Krishnamurti 2003: 213-215); indeed, *-i* is reconstructed for Proto-Dravidian. However, the Kurukh use of *-a:/-i:* to mark adjectival gender, with both loans (*alga:/algi:* ‘redundant’) and native items (*otxa:/utxi:*) is more convincing. Along similar lines, Malto has e. g. *bobe*/*bobi* ‘stupid’, *lela*/*leli* ‘foolish’.

The Indo-Aryan languages of eastern India have classifier systems consisting largely of native morphological material (although it is likely that the underlying structure represents a borrowing from Tibeto-Burman): the two most universal morphemes are *-(go)ta*/*to* (indicating some generic category of count nouns) and *-jan* (indicating humans), e. g. Bengali *nɔe-**ʈa** balish* ‘nine pillows’, *ɔnek-**jɔn** manuʃ* ‘one man’. The non-Indo-Aryan (i. e. Dravidian and Munda) languages of the immediate area almost universally copy these: thus among the Dravidian languages we find Kurukh *mu:nd **oʈa:** ekho:* ‘three cows’ and Malto *ti:ni **jen** maler* ‘three men’ (Krishnamurti 2003: 404). In the Dravidian languages Kolami, Parji, and Kui, borrowed Indo-Aryan classifiers seem to be restricted to loaned numerals (Emeneau 1956: 116–118): Kolami uses the Marathi form *zen* ‘human’ with numbers over five, all borrowed from Marathi; Parji uses *jan* and *gota* with the numerals over six, all of Halbi origin; Kui uses *zana* and *gotta* for the Oriya numerals from three onwards. No such restriction is seen in Kurukh or Malto, in which Indo-Aryan classifiers may apply to all Dravidian numerals. In Malto, a more complex system is seen and includes several additional Indo-Aryan items: *danra* (< Modern Indo-Aryan **dand* ‘stick’) for long objects, *pata* for flat objects (< MIA **pat* ‘board’). The Munda language Kharia (Peterson 2010: 195) uses, in addition to various native morphemes, *j(h)an* with human reference (presumably from Sadri): *tin jhan lebu=ki* ‘three people’.

Many of the Tibeto-Burman languages of Nepal show significant borrowings from the Indo-Aryan language Nepali, their major contact language. In

[3] A similar Indo-Aryan form, Sadri *-ker*, is calqued into the Munda language Kharia as *=kon* from Kharia *ikon* ‘make, do’. (Peterson 2010: 325)

many of the Kiranti languages (Ebert 2003a: 514), which do not natively mark undergoers for case, the Nepali dative *-lai* (*timilai* 'to you') is copied. In Camling, Belhare, and possibly in others (cf Ebert), this borrowed suffix is permitted only to apply to humans; no such restriction appears in Nepali, or in other Kiranti languages such as Puma.

Camling (Ebert 2003b: 536):

	Ram-mo	*m-nicho-wa*	*a-woini-**lai***	*bhe-wa*	*ap-u.*
(3a)	Ram-GEN	2sPOSS-sibling-ERG	1sPOSS-friend-DAT	arrow-INST	aim-3P
	'Ram's brother aimed at my friend with an arrow.'				

Puma (Bickel & Gaenszle 2005):

	*gai-**lai*** *ghasa* *itd-oŋ*	*khim-**lai*** *copt-u-ŋ*
(3b)	cow-DAT grass give-1sS.PST	house-DAT see-3P-1sA
	'I gave grass to the cow.'	'I see the house.'

Gurung also has a dative in *-lai*; however (Noonan 2008: 95), the Tamang-Gurung languages inherit a dative in **-(l/r)a* – it is thus ultimately unclear whether the Gurung dative is an inherited form or a later borrowing[4]. The status of other borrowed morphemes in Tamang-Gurung, however, is more readily apparent. Gurung, Chantyal, and Tamang all borrow the Nepali clitic *-bʱənda* 'than' (etymologically 'saying') for the comparative[5]:

Nepali:

	ne:pa:l	*bʱuʈa:n-**bʱənda***	*ʈʰulo*	*tsə*
(4a)	Nepal	Bhutan-than	big	3SG.PRES
	'Nepal is bigger than Bhutan.'			

Tamang (Mazaudon 2003: 309):

	*ŋa=**banda***	*ti:-tiŋ-la*	*puriŋ*
(4b)	I	than one-year-GEN	younger sister
	'She is one year younger than me.'		

In Chantyal, the comitative is typically marked with one of the related forms *-səŋ, -səŋə, -səŋgə* (Noonan 2003a: 319)*,* apparently copies of Nepali *-səŋga* (*ra:m=le si:ta**=səŋgə*** *bɪha: gər-yo* 'Ram married Sita.'). However, Chantyal appears to

[4] We revisit this case in chapter 4.

[5] Mazaudon also gives for Tamang the form *pi-ma*, a calque of the 'saying' meaning of *bʱənda.* A related say-derived comparative form, *pi-le*, occurs in Manange (Hildebrandt 2004: 102), which we may attribute to a similar pattern copy.

maintain a native comitative in *-ru,* with no significant distributional differences noted in the literature: the possible functional motivation for this borrowing is then unclear.

Several morphological elements are copied from Malay into the various Aslian (Mon-Khmer) languages of the Malaysian peninsula[6] [7]; we bear in mind the possible interest in any differences between the Aslian languages in 'uptake' of Malay elements. In Semelai (Kruspe 2004: 82-83, 153-5) and Jahai (Burenhult 2005: 96), the Malay 'middle voice' prefix *ber-* (*bər-*) is copied as *b(r)-.* In standard Malay, *bər-,* typically a progressive, provides various senses of utilisation, possession, and habituality: with verbs, *bercukur* 'to shave oneself' (*cukur* 'to shave'), *bertenun* 'to be a weaver' (*tenun* 'to weave'); with nouns, produce or emit (*berbunga* 'to flower', *bertelur* 'to lay eggs'), to use or possess (*berbasikal* 'to bicycle'), to work at or as (*berkedai* 'to work at a shop'; *berkuli* 'to work as a labourer); with adjectives, *berduka* 'to be sorrowful' (*duka* 'sorrow') (Benjamin 1993: 372). In Jahai, *b-* largely functions as a straightforward progressive, e. g. *cara?* 'to talk'/*b-cara?* 'to be talking'. In Semelai, the use of *br-* is outlined below, and broadly reflects the range of meaning attested in the Malay case:

	da?	*br-wɒj*	*kəh*	*sma?*	*b-misaj*	*b-jaŋɔt*
(5)	NEG	HAVE-knife	3SG	person	HAVE-moustache	HAVE-beard
	'He didn't have a knife.'			'a bearded, moustached person'		

	b-baju?	*b-ladzar*	*br-dɔl*	*b-kayuh*
(6)	USE-clothes	USE-sail	USE-house	USE-paddle
	'to wear clothes'	'to sail'	'to house oneself'	'to paddle'

Several narrower functions of the Malay *ber-* do remain unattested in Semelai: 'to emit NP', 'to address as NP'. In both Jahai and Semelai, *b(r)-* appears in a process of numeral-to-verb derivation (Burenhult 2005: 110) unattested in

---

[6] Although we focus on Jahai and Semelai here, Matisoff (2003: 40) provides an inventory of borrowed Malay morphology in Temiar: *-m-* 'subjectivity', *-a-* 'middle voice', *bar-* 'progressive', *ma-* 'object orientation', *ter-* 'causative'. The similarity to the set of borrowings into Jahai and Semelai is clear!

[7] Kruspe (2004: 82) includes the prefixed causative *p-* in the list of affixes borrowed into Semelai from Malay. This is perhaps the least convincing example: in fact, *pə-* as a causative occurs across many branches of Mon-Khmer, and the existence of a similar form in Burmese may even be due to Mon influence (Pan Hla 1989: 29). It is likely that *pə-* can be reconstructed for both Austronesian and Mon-Khmer; as Thurgood (1999: 242) does for Chamic, we may instead say that any contact effect served to reinforce its use rather than loan a similar element outright.

standard Malay, though attested in the Perak dialect (spoken in close proximity to all Jahai communities): Jahai *duwa?* 'two' > *b-duwa?* 'to be two', *tiga?* 'three' > *b-tiga?* 'to be three'.

In Semelai, the very productive Malay prefix *ter-* is copied as *tr-*, described as 'happenstance' (Kruspe 2004: 83). Malay *ter-* serves to mark a lack of intentionality: either in the passive (*Syarifah* ***ter****-pukul* 'Syarifah was hit') or an 'accidental' active (*Ridzuan* ***ter****-muntah* 'Ridzuan *accidentally* vomited'*).* The Semelai *tr-* largely fulfils the second function: *tr-ca* 'to happen to eat', *tr-yɔk* 'to happen to take'. *ter-* seems to have no presence in Jahai, surprising given the extensive parallels between Jahai and Semelai in the domain of affix borrowing, and given the widespread adoption of *ter-* into other Aslian languages. In order to explain this, we suggest that the borrowing of Malay *ter-* into Jahai is potentially blocked by its similarity to an existing causative in *tr-* (Malay *ter-* is typically rendered [tᵊr]): Jahai *tr-cɨp* 'to cause someone to walk' from *cɨp* 'walk', *tr-gɨm* 'to summon to a deliberation' from *gɨm* 'to deliberate'. (This element is Semelai *tar-*.) The productivity of a pre-existing element in Jahai with a superficially identical shape to Malay *ter-* but a substantially different semantic value may here prevent the incorporation of the latter into the Jahai inventory, thus giving rise to one instance in which the contact influence of Malay on Semelai and Jahai morphology is perceptibly different.

Both Jahai and Semelai display extensive systems of numeral classifiers, containing several material copies from Malay (itself a classifier-optional language):

Jahai (Burenhult 2005: 81)	Semelai (Kruspe 2004: 207)	Malay	Category
bidaŋ	*bidaŋ*	*bidang*	'broad piece' – large flat objects
bataŋ	–	*batang*	'stick/shaft/river' – oblong objects
kpiŋ	–	*keping*	'portion' – flat objects
biɟi?	*bjɛ?**	*biji*	'seed' – small objects, sometimes round
buwah	*bjɛ?**	*buah*	'fruit' – sizeable three-dimensional objects
?ikɔr	*?ikur***	*ekor*	'tail' – all animals
prdu?	–	*perdu*	'base of tree' – clustered objects
–	*hlay*	*helai*	'strand' – of hair, thin layers
–	*buku?*	*buku*	'lump' – earth, bread

Table 1. Loaned classifiers in Jahai and Semelai.

* – in Semelai, *bje?* is collapsed with *buah* and used as a genericised classifier for three-dimensional objects.

** – in Semelai, *?ikur* is extended, and may denote *any* animate entity, incl. humans.

The classifiers definitely borrowed into both Jahai and Semelai are *bidang, biji* and *ekor*; in Semelai, the Malay 'small object' classifier is extended over all three-dimensional objects (it is possible that a perceived phonological similarity promoted this).

3.1.2 Recombination and reinterpretation: modified matter copies

Within our initial conceptual framework, the transfer of an individual morpheme has two separate components – the replication of matter, and the replication of pattern. We therefore suggest that we cannot straightforwardly provide a one-dimensional scale from 'most faithful' to 'least faithful' copying; variability in borrowing might arise on either, or both, of two axes corresponding to the degrees of matter and pattern copy respectively. The examples described in section 3.1.1, although described in terms of matter copy, in fact involved the copying of both matter and pattern, and in later sections we will discuss pattern-only copies; is it then possible to find situations in which morphological matter is copied without pattern? We can at least describe MAT loans for which the corresponding pattern is significantly altered. In the case of classifiers in Semelai discussed in 3.1.1, we certainly find that a significant semantic shift has taken place in the scope of various Malay classifiers, such that the PAT-element corresponding to Semelai *?ikur* or *bjɛ?* is only partially similar to that in Malay. In a similar vein, a novel pattern (numeral-verb derivation) is associated with the borrowing of *br-* in both Semelai and Jahai.

A slightly different possible type of 'modified' matter transfer involves the synthesis of existing morphological material with matter copies acquired from the donor language: the resultant morpheme thus incorporates material from both donor and recipient: e. g. the indefinite article *njek* in Epirus Romani (Matras 2009: 217), seemingly a combination of the Romani *(j)ek* and the Albanian *një*.

Chantyal, a Tamangic (Tibeto-Burman) language of Nepal, uses a complex sequential converb *si-rə* which consists of the standard Tamangic sequential converb in *-si*[8], and the Nepali conjunction *-rə* (Noonan 2008); the sequential

[8] Compare Nar-Phu *-se* (Noonan 2003b: 351), Manange *-tse* (Hildebrandt 2004: 289).

converb performs all conjunction-related functions, as the overt conjunction of clauses is not used in Chantyal.

Chantyal (adapted from Noonan 2003a: 333)

(7) *Ram-sə gãw-ri ɦja-**si-rə** nɦaka-ye sya ca-i*
Ram-ERG village-LOC go-ANT-SEQ chicken-GEN meat eat-PERFECT
'Ram went to the village and ate the chicken.'

The genesis of this form involves a reinterpretation of the Nepali sequential converb in *-erə*: in Chantyal, *-erə* seems to have been analysed as consisting of the Nepali morphemes *-e* (the perfect particle) and *-rə*, followed by the equation of Nepali *-e* and historic Chantyal *-si* (no longer found in isolation).

In the Turfan dialect of Uyghur (Yakup 2005: 152), in significant contact with Mandarin Chinese, the nominal suffix *-ŋza* derives a certain set of honourific and/or "arrogant" kinship terms: *aka* 'brother'/*akaŋza* 'older brother', ajla 'sister'/*ajlaŋza* 'older sister', *ata* 'father'/*ataŋza* 'father'. In standard Uyghur, the (informal or intimate) second-person possessive is marked in *-(i)ŋ* (Engesaeth, Yakup & Dwyer 2009: 112), e. g. *somkaŋ* 'your bag' (*somka* + *-ŋ*), *ëtiŋ* 'your horse' (*at* + *-ŋ*). *za* then appears to be derived from what Yakup describes as 'the Chinese denominal nominal suffix *-zi*', used to produce second-person possessive forms within Xinjiang Chinese; we may identify this with 之 *zhi1*, the Mandarin particle most frequently indicating posession. *-ŋza* then derives from an apparently redundant combination of two possession-related morphemes, one native and one replicated from Chinese; it is unclear precisely how the resultant morpheme acquired its derivational/honourific function and lost the assumed historical function.

In Sakha, the *-ttar* plural sufffix (*oyu:n* 'shaman'/*oyu:ttar* 'shamans') combines a Mongolic plural *-t*/*d* (historic *d, e. g. Khamnigan *keegen*/*keege-d* 'child/children'; Kalmyk *colun*/*colu-d* 'stone/s') with the common Turkic marker *-lar* under the standard Sakha assimilatiory rules (Gardani 2012: 90). The evidence for and historical propagation of this partially-copied morpheme can be traced to the lexical borrowings which must have introduced it: in the Mongolic languages, *-d* typically operates only on nouns ending in *n* (occasionally *l* or *r*) thus we have e. g. (Khalkha) Mongolian *zocin* 'guest', *zocid* 'guests'. This rule is reflected in a set of Sakha nouns with broadly collective meanings in *-t* which still correspond to forms in *-n*: *tojot* 'authority'/*tojon* 'master', *χotut* 'ladies'/*χotun* 'lady' (examples

from Pakendorf 2007: 298) – this suggests to us that the Mongolic usage exists without the addition of Turkic *-lar* in Sakha, and thus was initially copied independently *prior* to synthesis. *-ttar* clearly applies to Turkic lexicon, as well as to more recent Russian loans: *uol* 'boy'/*uola-ttar* 'boys', *ystakä* 'glass'/*ystakä-ttar* 'glasses' (Gardani 2012: 90).

In the Učur dialect of Evenki (Tungusic), under Sakha influence, the Sakha 'presumptive'-'assertive' suffix *-tax* is copied, along with the entire paradigm of Sakha person marking, and is appended to the Evenki present-tense marker *-r(a)-* to produce a 'hypothetical' series in *-rdax-* (Malchukov 2006: 126; Pakendorf 2009: 98-105):

	Sakha *bar* 'to walk'	Učur Evenki *wa:* 'to kill'
1SG	*bar-daɣ-ɪm*	*wa:-r-daɣim*
2SG	*bar-daɣ-ɪŋ*	*wa:-r-daɣiŋ*
3SG	*bar-daɣ-a*	*wa:-r-daɣa*
1PL	*bar-dax-pɪt*	*wa:-r-dakput*
2PL	*bar-dax-xɪt*	*wa:-r-dakkit*
3PL	*bar-dax-tara*	*wa:-r-daktara*

Table 2. Complete paradigm for the hypothetical mood in Sakha and Evenki.

This case has several interesting aspects. The first is the copying of the entire Sakha personal paradigm: the morphemes thus superficially replicated have no independent status in Evenki, and do not arise elsewhere (the 'strings' *-dax* + (person) are morphologically complex in Sakha, but unanalysable in Evenki). The second, more thematically relevant, is the concatenation of the native Evenki *-r-* with the copied string; we might suggest in addition some analogy to the existing Sakha 'necessitative' in *-rda:x(-)* promoting this usage, especially given an apparent partial analogy in meaning within U. Evenki:

Učur Evenki (Malchukov 2006: 127):
(8) *suu gorolli**-r.dak-**kit*
you far.away.go-PROB-2PL
'You must have gone far away.'

While this is a tempting argument, *-rda:x* in fact requires the use of a slightly different personal paradigm, suggesting that its influence is likely not the *primary*

force behind this construction – the final situation in Učur Evenki must indeed be formed as previously described.[9]

It is worth noting at the close of our discussion of these combined forms that 'imperfect copying' is a slightly misleading description; in many of our cases, the material involved *is* indeed 'perfectly' replicated from the donor language, but reshaped at a later stage for whatever purpose. Our intent here is rather to point out that the alteration which takes place in generating this type of morpheme acts to distance the novel form from the source form, thus reducing *faithfulness* as we might understand it (in terms of 'distance' from the input).

3.1.3 *Contact between related languages and phonological analogy.*

Although in this work we primarily focus on the strictly morphological aspects of borrowing, morphology is not an isolated entity: the act of encoding morphology affects and is affected by the other aspects of language; as such, we provide this brief note on the possible role of phonological similarity in driving or inhibiting contact processes. In 3.1.2, we suggested that the existing Sakha necessitative affix in -rda:x might promote the Evenki use of the innovation -r-dax due to the combination of phonological similarity and related meaning; in 3.1.1, we discussed the possibility of a phonological resemblance to an unrelated function *inhibiting* the borrowing of Malay *ter-* into Jahai. This type of behaviour in contact situations is not unknown – we note in particular that in closely-related languages, functional changes promoted by phonological analogy can render it difficult to identify the nature and scope of a potential borrowing (Matras & Sakel 2007: 5); in addition, we may ask ourselves whether it is even meaningful to propose the 'borrowing' of forms which were near-identical to begin with.

In the Sebjan-Küöl dialect of Even (also Tungusic) in Northern Siberia, the same presumptive form is copied from Sakha, once again along with the entire personal paradigm; however, the implementation differs slightly (Pakendorf 2009). Rather than being attached to the present-tense marker or similar as in Učur Evenki, the copied Sakha morphemes are integrated using a connective *-j-*, of unclear provenance, which appears *exclusively* in the context of the combination of an Even root and a succeeding Sakha inflectional morpheme:

Sebjan-Küöl Even (Pakendorf 2009: 94-97):

	*a:ŋŋa-**j-dagim***	*hista-**daga***
(9)	spend.the.night-CONN-ASS.1SG	get.infected-ASS.3SG
	'spend the night'	'get infected'

[9] *hista* above is a Sakha loan, with which the connector *-j-* does not arise.

In the Chamic languages, the causative in **pə-* is near-universal[10] and occurs freely across the lexicon (Thurgood 1999: 242). This causative has no parallels in much of Malayic, and we note that **pə-* is also the form of the causative reconstructed for Proto-Mon-Khmer – is it then possible that the Chamic causative has been borrowed from Mon-Khmer? In fact, the presence of various reflexes of **pə-* in the Austronesian languages of Taiwan, the Philippines, and Borneo – all relatively free from Mon-Khmer influence – suggests that this prefix may be reconstructed for Austronesian, and has simply been lost in large sections of the family. Nonetheless, we suggest that the coincidental similarity might have served to reinforce the use and retention of *p-* causatives in Cham. A similar situation arises in Gurung (Noonan 2008: 95); the current dative form in *-lai* bears clear resemblance to the Nepali dative, also *-lai*, but the attestation of forms in *-la* and *-ra* in closely-related languages leaves us uncertain as to the precise status of contact influence in this case. Phonological equivalence may also arise in the Semelai generalisation of Malay *bidʒi > bjɛʔ* in order to cover functions that in Malay are split between *bidʒi* and *buah*: Semelai phonotactics would likely nativise Malay *buah* as *bwɛʔ*, which lends itself well to the possible analogy. More clear-cut is the status of the 'Batakisms' *ni-* and *mər-* which replaced Classical Malay *di-* and *bar-* in Old Malay (Mahdi 2005: 183-4): despite the clear genesis from identical Malayic forms, the phonological difference is sufficiently pronounced to be unambiguous.

It is possible to produce 'uncertain' borrowings in function as well as in form. In Tagalog, the circumfix *pag-...-an* denotes an object focus counterpart of actor focus verbs with mag-, for instance giving the pairs *mag-áral* 'to study' and *pag-arál-an* 'study (n.)'. This reflex of putative proto-Malayo-Polynesian *paR-X-an also appears in Malay as *pər-...-an*, with an allomorph *pel-...-an* appearing if there is a rhotic in the base it modifies, as in *bəl-ajar* 'to study' : *pəl-ajar-an* 'studying, lesson'. These circumfixes do not correspond exactly in meaning, with the Tagalog form carrying a more 'locational' sense and the Malay form acting primarily to describe processes; but, the above examples' isomorphism, and the status of *áral* as a loaned root in Tagalog from Malay *ajar*, suggest that an analogy has arisen due to Malay influence on Tagalog. (Blust 2013: 400).

---

[10] Save only in Tsat, where *all* initial syllables were dropped due to extensive contact with various Chinese.

A case which contains both these elements of ‘closeness’ but also much clearer evidence for functional borrowing is the situation of the Bengali plural markers. *-ra:* and *-gulo* are human and non-human plurals, respectively; the clear similarity to Dravidian *-*r* and *-*gal* suggests the transfer of matter from Dravidian during the acquisition of the human-nonhuman class distinction (otherwise absent in Indo-Aryan). However, it is argued (e. g. Chatterjee 1926) that Indo-Aryan derivations for these markers are more plausible; reconciling this fact with the sudden acquisition of a hitherto unknown functional distinction recalls us to the issue of phonological analogy.

3.1.4 *Pattern replication.*

We now turn to cases in which the formal substance, or matter, is *not* imported, but a contact-induced morphological change is nevertheless evident: in such a situation, what is copied is the pattern, i. e. aspects of distribution, semantics, and formal structure. Pattern replication is well-established in the existing literature on both language contact in general (Matras & Sakel 2007b) and morphological borrowing; we discuss a small selection of examples which are absent from the current literature or which lend themselves particularly well to our organisation.

From a functional perspective, this may take several ‘shapes’, which we discuss in greater depth through section 3.2: one possibility, and perhaps the clearest, is the introduction of a wholly new grammatical or semantic function either causing the extension/reinterpretation of an existing affix or generating a *novel* bound form from other portions of the repertoire; but it is also possible to examine situations of e. g. significant *structural* rearrangement in accordance with a foreign template, without any particular associated shift in function or semantic value. We also note that such borrowing may be seen to be ongoing in a real sense: in Jahai, the prefix in *b-,* itself originally borrowed from Malay, was initially incorporated with a strictly progressive meaning, e. g. *ʔimbus* ‘to ambush’/*b-ʔimbus* ‘to be ambushing’, and typically does not change the argument structure associated with its verbal forms. However, in a few exceptional cases (Burenhult 2005: 158), it participates in a valence-reducing operation and allows a passive-like construction:

Jahai (Burenhult 2005: 158)

wa=bk-ʔɛk ka=gtah
(10) IRR.3S=PROG-to.give SUBJ=rubber
'Rubber will be given.'

This seems ultimately to derive from another of the various functions of Malay *ber-*, not initially copied into Jahai during the original transfer of *b-*.

In the Mandarin Chinese variety of Xinjiang, in northwestern China, extensive plural marking occurs using the repurposed morpheme 们 *men5* (Baki 2012). 们 *men5* in standard Mandarin exclusively occurs in plural forms of pronouns, or in certain collective nouns referring to indefinite numbers of humans: 你-们 *ni3-**men5*** 'you.pl', 学生-们 *xue2sheng1-**men5*** 'students' – it cannot be used with a number, or apply to non-humans (Yip & Rimmington 2006: 13). In contrast, its function in the Xinjiang variety appears significantly extended:

Xinjiang spoken Mandarin (Baki 2012: 52):

	*shei2-**men5*** *lai2-le5?*	*ni3-men5 je4xie1 seng1ko3 - **men5***
	谁 们 来 了?	你 们 这些 牲口 们
(11)	who - PL come-PAST	2PL these animal - PL
	'Who has arrived?'	'You animals!'

Standard Mandarin (Baki 2012: 52):

	shei2 lai2-le5?	*ni3-men5 zhe4xie1 sheng1kou3*
	谁 来 了?	你 们 这些 牲口
(12)	who come-PAST	2PL these animal
	'Who has arrived?'	'You animals!'

In the examples of Xinjiang spoken Mandarin above, 们 *-men5* is seen to be permitted with interrogatives (谁 *shei2* 'who') and with non-human plurals (牲口 *sheng1kou3* 'animal'), neither of which may occur in standard Mandarin. The motivation for this shift is readily discernible in the application of standard Uyghur morphology:

Uyghur[11]:

	kim-ler kel-di?
(13)	who-PL come-PAST
	'Who has arrived?'

In Uyghur, as in the other Turkic languages, the suffix *-ler/lar* (conditioned by vowel harmony) compulsorily marks plurality on any type of noun; the pattern found in the Xinjiang variety of Mandarin is then a straightforward transfer of this usage, but makes use of a native – if dispreferred – morpheme rather than a copy

[11] Constructed by the author, modelled on Engesaeth, Yakup, Dwyer 2009.

of the Uyghur template. We may briefly note an example of partial pattern transfer in the opposite direction (i. e. from Mandarin to Uyghur), although in a less salient portion of the morphology: in the Turfan dialect (Yakup 2005: 174), the Mandarin use of 人 *ren2* 'person, people' to indicate origin – as in e. g. 北京人 *bei3 jing1 ren2* 'Beijinger/s', 本地人 *ben3 di4 ren2* 'local people' – is copied and implemented using the Uyghur *adam* 'man'. This replaces the historical suffix in *-liq/-lik* (e. g. Uzbek *Hindistonlik* 'Indian'): *qashqa-adimi* 'person from Kashgar', *bendi adam* 'local people' (with *bendi* being additionally copied from the Mandarin 本地).

A similar extension of an existing morpheme occurs in certain varieties of Even. In the Even dialects outside the Sakha/Tungusic contact zone, the existing sociative marker *-lda-* (Malchukov 2006:125–6) may only be used to express the sense of reciprocality when applied to 'symmetrical predicates': e. g. *bak-* 'to find'/*baka-lda-* 'to find one another'; *naa-* 'to collide'/*naa-lda-* 'to meet', but typically *not* e. g. *em-* 'to come'/**eme-lde-* 'to part'. No such situation arises in Sakha, in which the form *-(t)s-* marks both a sociative and a reciprocal function:

Sakha (Nedjalkov & Nedjalkov 2007):
*Kiniler [uulussa-nə] muostala-**s**-t-ələr*
(14a) they street-ACC pave-**SOC**-PAST-3PL
'They paved the street **together**.'

*Kiniler muostala-**s**-t-ələr*
(14b) they pave-**REC**-PAST-3PL
'They paved **each other.'**

(14b), if an odd locution, is perfectly grammatical in Sakha – we note that the form *-s-* indeed serves both functions. Under the influence of this usage, eastern dialects of Even no longer maintain the functional restriction otherwise expected, and extend the use of *-lda-* to both functions for which *-(t)s-* would be an appropriate choice in Yakut: thus *haa-* 'to know' > *haa-lda-* 'to know each other'.

The Sri Lankan variety of Malay (Bakker 2006: 139; Nordhoff 2009)[12] has undergone significant restructuring based presumably on a Tamil template. Standard Malay entirely lacks case marking, and virtually all its morphology is derivational; in order, therefore, to implement an extensive system of nominal inflection, Sri Lanka Malay grammaticalises a large series of markers whose material content is entirely 'native', but whose use is modelled on Tamil.

[12] There is some mismatch between the data provided in Bakker 2006 and the analysis in Nordhoff 2009.

	Sri Lanka Malay (Nordhoff 2009)	Tamil	Malay equivalent
NOM	*0*	*0*	–
ACC	*=yang* < Malay *yang* 'which'	*-a/-e*	–
DAT	=nang < Javanese *nang* (allative)/Malay *nya* (possessive)	*-(u)kku/-(u)kkaha*	–
LOC	*=ka* < Malay *dekat* 'near'	*-le/-la/-ile*	free forms *di, ka, dari*
ABL	*=dering* < Malay *dari* 'from' (Bakker 2006: *-ka dupa* < Malay *dekat* 'near' + ??)	*-iruntu*	free form *dari*
GEN	*=pe* < Malay *punya* 'of'	*-ɽa/-o:ʈai*	**optional** *-nya* or various pronominal clitics; *anak raja itu/ anak-nya raja itu* equally acceptable for 'the child of that raja'
COM	*=samma* < M. *sama* 'with'	*-o:ʈai*	free form *sama*

Table 3. Nominal case clitics in Tamil and in Sri Lanka Malay.

We briefly describe a few examples of the semantic convergence between the Sri Lanka Malay and Tamil systems.

In both (15a) and (15b), the subjects are marked with the dative case; the marking of semantic experiencers with the dative is a feature common to the languages of South Asia, and clearly acquired by Sri Lanka Malay due to its prevalence in surrounding languages.

Sri Lanka Malay (Nordhoff 2009: 569):
inni ***o:rang=nang*** *itthu thara-tha:u*
(15a) PROX **man=DAT** DIST NEG-know
'This man did not know that.'

Tamil:
ɪndɨ ***a:l=ukku*** *adɨ teri-ille*
(15b) PROX **man=DAT** DIST know-NEG
'This man did not know that.'

In Tamil, the accusative case is typically only used with animates, with which it is obligatory – inanimate objects do not require accusative marking, and its presence in such contexts indicates definiteness or specificity (thus *na: tanni kʊɋɪce:n* 'I drank water', but *na: tanni-**ye** kʊɋɪce:n* 'I drank *the* water'). Sri Lanka

Malay *=yang*, similarly, is selected for by animacy, topicality/definiteness, and singular reference (Nordhoff 2009: 330). We may note in particular for this example that the most interesting feature of the Sri Lanka Malay inflectional system seems not to be its specific correspondences to Tamil, but rather its complete restructuring relative to standard Malay. While standard Malay is primarily analytic morphologically, Sri Lanka Malay displays a significant tendency towards agglutination, acquires overt case-marking, a significant system of verbal markers, and plural-marking (accomplished via total reduplication in most Malay varieties).

3.2 Disturbances to the morphological system

In section 3.1, we have outlined some of the possible ways in which loaned elements in the morphology can be related to their ultimate source. This does not, however, cover several aspects of morphological borrowing that seem crucial to developing our understanding further; in particular, we have established no understanding of how the differences between the morphological systems of languages in contact affect their impact on one another; we also have relatively little discussion of the status of the various copies presented in 3.1 within the morphologies of the recipient language. Not all loans are created equal; as we suggest above, those which function in one very limited derivational context will often be less salient, in terms of their frequency of use and pervasiveness through a discourse, than e. g. a loan in some virtually omnipresent domain. (Consider the example of Uyghur *-ŋza* given in 3.1.2; while interesting from the perspective of the mechanism of borrowing, it is of little broader structural significance within Uyghur, being restricted to the derivation of a specialised and entirely optional set of lexical items.)

A criterion that we set on morphological borrowing is that the borrowed element must be productive and propagate throughout the language, rather than remaining restricted to the loaned portions of the lexicon. We might then set a similar criterion on *systemic* loans: to what extent do borrowed morphological templates propagate through the structure of the language? A possible partial typology of contact morphology might then be suggested based on the *disturbance* created by the instantiation of loaned material or pattern into the

recipient language, both in function and in affixational structure/broad morphological organisation. We classify examples of contact-induced morphological change from the least system-disturbing to the most radically system-disturbing – how 'foreign' is a particular borrowed element to the pre-contact state of the recipient language, and to what extent is the post-contact state of the recipient language modified to account for this? We aim additionally to articulate the differences in 'systems' modified: the hypothetical borrowing of a simple derivational suffix [A] might be quite unremarkable in its functional impact on the recipient language, but very remarkable indeed in a language whose morphology did not previously contain suffixes.

3.2.1. *The replacement of existing morphemes; undisturbed systems.*

Perhaps the most trivial change to a morphological system arises in the simple replacement of an existing morpheme. While this might be quite an interesting change from the perspective of the functional justification for morphological borrowing – what motivation can we suggest for the simple replacement of a pre-existing form? – it is, structurally, almost entirely unremarkable: in the idealised 'perfect' material replacement, a marker [Q] is supplanted by a distributionally, affixationally, and functionally identical marker [X], which from the point of view of the grammatical system is as invisible as the selection of a different lexical item. Examples already presented include the replacement of the genitive in Malto with Indo-Aryan *-ki* (3.1.1): given the presence of genitive marking throughout the Dravidian languages, and the ease of reconstruction, it is almost certain that *-ki* replaced some pre-existing, functionally and semantically equivalent morphological element in Malto. This is similarly true of e. g. the Nepali comitative in Chantyal, which exists alongside a native marker and does not differ in function or permitted distribution. We may class alongside such material copies recombinations such as the Chantyal sequential converb in 3.1.2: Nepali *-rə* is ultimately not used independently, but is combined with the existing Chantyal *-si*, creating a materially new sequential converb whose function and distribution remain unchanged. Similarly, the Sakha plural suffix *-ttar* does not represent any structural novelty in Sakha, and may be considered simply a minor surface modification of an existing marker.

3.2.2. *New functions consistent with existing patterns.*

We next consider the introduction of some new ‘minor’ function, new type of marking, or new structure into a language whose existing systems are largely consistent with this: for example, the grammaticalisation of a novel comitative-case marker in a language with a well-defined pre-existing case system. To an extent, we may include in this class cases in which an existing morpheme is extended over a new function: the object focus circumfix of Tagalog, mentioned previously in 3.1.3, provides a marginal example of this. Its use as a nominaliser appears to result from Malay influence, given that it also functions to change the focus of verbs as in the pair *mag-tiʔís* 'suffer'*, pag-tiʔis-án* ‘endure’; Tagalog has no pre-existing morphological strategy filling the function that the Malay pattern does, but ultimately this represents quite a small change to the situation of Tagalog.

In the Austronesian language Madurese, an irrealis suffix in *-a* (e. g. *Siti entar-**a** dha’ Sorbaja* ‘Siti will go to Surabaya’) is quite frequent (Davies 2010: 260), apparently copied from Javanese; similar material does not arise in the closely related Bali-Sasak languages, or in Malayic in general. Verbal affixes specifically signalling irrealis mood are unusual in Austronesian, although the marking of mood is not unusual; this is then an innovation (of unknown source) in Javanese, and a minor functional novelty in Madurese. We may class the Evenki hypothetical in *-rdax-* (3.1.2) similarly: the existence of morphological marking denoting mood is not new to Evenki, although the specific mood was previously unknown.

In Cantonese, the use of a numeral classifier without a numeral indicates definiteness, a usage ungrammatical in other Sinitic languages but commonplace in the Hmong-Mien and Tai languages:

Cantonese (Matthews 2007: 230)
(16a) *ngo5 wan2 dou2 [zek3 maau1]*
I find successfully CL cat
‘I found the cat.’

Hmong (Matthews 2007: 230)
(16b) *Huab-tais nrhiav tau [tus poj ntsuag]*
Emperor find get CL widow
‘The Emperor found the widow.’

It is notable that this function in Hmong, as in (16b), ultimately derives from a demonstrative [Classifier-Noun Demonstrative] (CL N DEM) construction, and only occurs without the demonstrative when anaphoric; precisely as in Cantonese.

Classifiers are additionally used in Hmong and Cantonese (*not* Zhuang/Tai or Mandarin) to express possession: the Cantonese *ngo5 zoeng1 toi4* [I-CL-table] is equivalent to the Hmong *kuv lub rooj* [I-CL-table], and the classifier is dropped in cases of inalienable possession in both languages. While Cantonese also has a Sinitic marker of possession in *ge3* (Mandarin *ta5*), the Hmong-like classifier constructions ultimately suggests a South-East Asian origin of classifier morphosyntax in Cantonese (Matthews 2007: 231)[13]. However, the development of this novel function for classifiers in Cantonese, while undoubtedly noticeable, does not create a new category in Cantonese, or promote any large-scale restructuring of the Cantonese grammatical system.

3.2.3. *Introduction of entirely new functional categories.*

We continue to introduce more 'alien' morphology to the system. In what situations do entirely new functional categories enter a language? A case with significant scope is the incorporation of numerical classifiers into the languages in which they can be proven to be borrowings (after Matras 2009: 216, we may note that classifiers are unusual enough in the world's languages to be areal in their overall distribution). Turfan Uyghur has adopted the use of classifiers from Mandarin, largely repurposing existing Uyghur material to accomplish this: **ikki müŋüz* 'two horns', instead *ikki tal müŋüz* 'two-CL horns' (Yakup 2007: 107); *danä* 'seed, grain' (< Persian) as in *on danä nävrä* 'ten-CL grandchildren' or *jetti danä kala* 'seven-CL cows'. *danä*, originally specific to small objects, has been extended by analogy to Mandarin *ge5* to be used for other objects as well as people (Yakup 2005: 173). This increased preference for overt quantifiers is a significant structural deviation in Uyghur. We have previously discussed the existence of classifiers in the Indo-Aryan and Dravidian languages of eastern India, a similar case of the evolution of this system based on a likely Tibeto-Burman template; the Tibeto-Burman language Chantyal, part of a Tibeto-Burman branch in which classifiers are otherwise unattested, has borrowed these

[13] Vietnamese, a language of the same area in heavy contact with both Cantonese and Tai languages, also uses a classifier construction to indicate definiteness (Simpson 2005: 813).

elements from Nepali. Classifiers exist in the Mongolic language Mangghuer, and are almost exclusively borrowed from Mandarin, along with the entire numeral system save 'one' and 'two': *=ge* 'piece'/*liang=ge* 'two=CL', *=zhuan* 'circle' (Slater 2003: 313); the system itself is foreign to Mongolic and a clear result, again, of the influence of Mandarin. Matisoff (2003) cites the borrowing of classifiers into Aslian languages (as discussed in 3.1 for Semelai and Jahai) as invariably a Malay-ism in Aslian, with the note that classifiers are typically optional and presumably a contact feature across Mon-Khmer (the influence of Chinese on the classifier system in Vietnamese is described in Alves 2007 and in Matras 2009).

This wholesale adoption of classifier systems as a result of language contact is quite striking, being a repeatedly-implemented category addition in several different language families. In contemplating the possible procedural and functional reasons for this adoption, we note that in almost all the cases discussed, numerals are also partially or wholly adapted from the contact languages. As we see in section 3.1, the Dravidian languages of Eastern India borrow numerals from Indo-Aryan, and their use of classifiers may be either restricted to the borrowed numerals (Kolami, Parji, Kui, Pengo) or permitted with native numerals (Kurukh, Malto); with the exception of Malto, which generates classifiers based on native lexicon, these classifiers invariably consist of Indo-Aryan material. The numerals in Korean are entirely borrowed. In Chantyal, no native numerals exist, and borrowed Nepali classifiers are used. Between the Aslian languages Semelai and Jahai, Jahai displays a much greater preference for Malay classifiers, borrowing a larger set thereof; in Semelai, the first seven cardinal numerals save 'two' are of Mon-Khmer origin (Kruspe 2004: 205), while in Jahai all numerals save only 'one' are borrowed from Malay. The situation in Vietnamese is slightly more complex, as native numerals and 'Sino-Vietnamese' loans coexist within the lexicon and are typically both available within a speaker's repertoire; nonetheless, we may suggest with some confidence that the *material* borrowing of classifiers patterns together with the borrowing of numerals.[14] In Uyghur and in the various Indo-Aryan languages with classifier systems, numerals are *not* borrowed from the presumed origin languages of the classifier construction, and we must seek a different explanation.

[14] In all these cases, the contact language responsible for classifiers is also the predominant language of the area, and likely to be the language in which everyday economic activity is conducted; the functional explanation for *numeral* borrowing (after Matras 2009: 201) is not hard to see.

Grammatical gender as a category can also be subject to morphological borrowing; however, the exact circumstances may be limited. In Tagalog, lexical material borrowed from Spanish seemingly retains grammatical gender and triggers agreement with modifiers:

(17a) Tagalog (Stolz 2012: 99)
Komik-o ang kuwento
funny-MASC DEF.TOP story
(17b)
Komik-a si Linda.
funny-FEM DEF.PROPER Linda
(17c)
Komik-o ang mga bata.
funny-MASC DEF.TOP PL child
(17d)
Komik-o si Fred.
funny-MASC DEF.PROPER Fred

komik ‘funny’ is seen to require agreement with animate head nouns, but takes the masculine ‘default’ with inanimates and plurals. Although this gender-marking seems to be exclusively restricted to the Spanish portion of the lexicon, it has palpably ‘progressed’ further in Tagalog than in e. g. Indonesian, in which Sanskrit nominal loans such as *putra* ‘son’/*putri* ‘daughter’ display fossilised gender-marking that extends no further. In Chamorro (Austronesian: Guam), marginal grammatical gender appears to have arisen as a result of extensive Spanish contact. Chamorro distinguishes grammatical gender in loaned modifiers and in nouns: *apbladot* 'male gossiper'/*apbladora* 'female gossiper'. Agreement in Spanish loans remains, however, limited to animates: *siuda* 'town', from Spanish *ciudad* (f.) 'city, town' does *not* trigger agreement – *bunitu na siuda* 'pretty town' despite the explicit *-a*. Feminine nouns in Chamorro that do *not* end in *-a*, by contrast, can trigger agreement, as in *mi-milagrosa na Bithen* 'abundantly miraculous Virgin', suggesting the category of gender in Chamorro is ultimately semantic, with animate female forms as the marked class (Stolz 2012).

3.2.4. *Innovations on the word level: new affixation and structure.*

We have from the beginning operated within the ‘form-function’ split derived from the distinction between matter- and pattern-replication. In order to encapsulate the behaviour of morphology, we further subdivide pattern in a similar form/

function manner. We have in 3.2.1-3 discussed the behaviour of the *function*, or various semantic properties, and we now turn to the *form*: issues of word-internal and word-external combination and relative frequency of use. At this end of our scale, we can further propose a small division in the behaviour of loaned morphology: *given* that a particular loaned element is 'alien' to the existing system, has it or has it not caused the propagation of a systemic change? The systemic change to which we refer can ultimately become a change in morphological typology, which we also briefly address here.

The morphology of the Dravidian languages is typically described as entirely suffixing; no prefixes can be reconstructed for Proto-Dravidian, and none are seen in the southernmost (least exposed to non-Dravidian languages) members of the family (Krishnamurti 2003: 28). In both Telugu and Kannada, however, a significant number of Indo-Aryan (Sanskrit) prefixes are borrowed, and are quite productive in nominal derivation across both the Dravidian and the Indo-Aryan components of the lexicon: thus Kannada *durbaɭake* 'misuse' (*dur-* + *baɭake* 'use'), *pratiobba* 'each other' (*prati-* 'opposite, counterpart, contrary' + *obba* 'one [human]'), *nirakʃarate* 'illiteracy' (*nir-* + *akʃara-te* 'literacy'); Telugu *atiya:se* (*ati-* 'excessive' + *a:se* 'desire'), *apanambike* 'mistrust' (*apa-* 'negative' + *nambike* 'trust'). (Notably, many of the Kannada items in which these prefixes are used with Indo-Aryan roots appear to have been derived within Kannada itself; no attestation for them exists within historic or current Indo-Aryan lexicons.) These prefixes represent a relatively straightforward example of matter replication from Indo-Aryan into Dravidian; what is here particularly interesting is the apparent introduction of a 'foreign' scheme of affixation into the morphological systems of Kannada and Telugu. Is it possible to then determine the extent to which prefixation as a viable affixational strategy has been 'nativised' within these languages? It is clear that the morphological systems of both Kannada and Telugu now permit the use of prefixation, due to the productivity of the borrowed prefixes in the formation of novel locutions; has this had consequences for the systems at large?

Although we begin with the premise that 'prefixes' are not permitted in the Dravidian language, *compounds* are; following Aronoff and Shridhar (1988), we note that the difference lies in the inherent bound nature of prefixes – forms which we describe as prefixes must necessarily be phonologically and morphologically tied to their 'heads'. In Kannada, we find a series of seemingly restricted,

phonologically-unique morphemes derived from the native lexicon: consider as an illustration *keN-* from *kempe* 'red' (the Kannada nasal is likely underlyingly placeless). We cannot find the form *keN* in isolation in Kannada; it may only occur in e. g. *kembal* 'red tooth' (*hallu* 'tooth'; Kannada [h] is underlying /p/), *kemmugilu* 'red sky' (*mugilu* 'sky'), *kemmannu* 'red mud' (*mannu* 'mud'). Without the leisure for a detailed phonological analysis, we nevertheless suggest that this appears to be an instance of productive prefixation in Kannada, using native material and a function which is not parallel to an attested Indo-Aryan one; there is no material or semantic borrowing from Indo-Aryan involved in the generation of this form, but we argue that there *is* an Indo-Aryan influence on the set of constraints that determines acceptable affixation in Kannada. Previously in section 3.2, we argued that the borrowing of an individual derivational morpheme does not, in itself, constitute a major systemic disturbance in the domain of function; if we accept the Indo-Aryan influence on the Kannada morphological template, however, then we must stress that their effect within 'function' and their effect within 'structure' is quite different.

Nothing similar is attested in Telugu; given the strong structural similarity between Kannada and Telugu, we cannot look to structural factors in order to explain the difference in adaptation of the use of prefixation, and must consider other reasoning[15] – however, irrespective of the ultimate motivation, we seem to find a difference in the extent to which prefixation has been 'nativised' in both languages. Prefixation using identifiably Dravidian material is only attested in one other Dravidian language: in Brahui, under extreme contact influence from Indo-Aryan and Iranian (Brahui is isolated from all other Dravidian languages by a significant distance), Proto-Dravidian **me:l* 'high' finds a reflex in *be:- – be:-harsing* 'to turn over' (Krishnamurti 2003: 137).

Although we have not discussed the *loss* of morphology to a significant extent, we find it relevant to note that almost the opposite arises in Sri Lanka Malay. As we have previously described, Sri Lanka Malay diverges from Standard Malay in marking case with overt and often compulsory suffixes; it also marks plural pronouns using *pada*, a predicative affix found in Javanese dialects of Malay, rather than the standard Malay use of total reduplication. Although Sri Lanka Malay maintains some standard Malay prefixes (Baker 2006: 146), the relative frequency of novel suffixes (a type of affixation existent, but mildly

[15] This is possibly an effect of the greater historical prestige of Indo-Aryan in Karnataka.

dispreferred in the Malayic languages), and general increase in affixation, leads to a comparative perceived 'lack' of prefixes – a part of the larger-scale morphological-typological change, from highly analytic and prepositioning to postpositioning-agglutinative, seen in Sri Lanka Malay.

The Munda languages almost universally display a high degree of divergence from the expected morphological typology of the Austroasiatic languages. In contrast to e. g. Khmer and other languages of the Mon-Khmer branch, Munda is primarily head-final and agglutinative, and heavily suffixing; in fact, Ebert (2003) suggests that the highly agglutinative morphology of the south-eastern Kiranti (Tibeto-Burman) languages is ultimately due to Munda influence. We contrast the following examples of Khmer and Sora, a Munda language of Orissa:

Sora (Donegan & Stampe 2004: 3)
(18) *anin dɔŋ-ɲɛn darəj-ən ə-tiy-ben idsɪm-tɛ ted*
3.SING OBJ-me rice -ART INF-give-INF want-3PR not

Khmer
koət ʔət caŋ ʔaoy baay kɲom
3.SING not want give rice me
'He doesn't want to give me the rice.

The Sora example is largely synthetic (Donegan & Stampe also provide polysynthetic forms, e. g. *ədməltiydariɲdɑe)* and thus incorporates a good deal of affixation; the Khmer example involves no inflection, consistent with the analytic and typically typologically South-East Asian nature of Khmer.

Our final note in this case revisits the issue of phonological analogy and the morphology-nonmorphology interaction. In the Chamic languages[16], as we have briefly mentioned, a significant degree of morphological restructuring has arisen due to Mon-Khmer contact; this includes the marked loss of suffixes in all languages of the family (Thurgood 1999: 238). In addition to the obvious implications for morphological typology, the Chamic loss of suffixes recalls the question of the phonological influence on morphological change: is it plausible that this affixational change is promoted by a clear prosodic change? Morphemes in the Mon-Khmer languages (Thurgood 1999: 61) are typically monosyllables or, more commonly, sesquisyllables: with a largely iambic stress pattern in which words consist of an unstressed 'presyllable' and a stressed *main* syllable (i. e.

[16] This includes the Acehnese language spoken on north Sumatra, which displays an areal affinity to the eastern Mainland South-East Asian linguistic area despite its distance from the Asian mainland (Thurgood 1999: 48).

compulsory final stress). It is quite likely that, historically, the predecessors of modern Chamic had stress systems analogous to modern-day Malayic, in which penultimate stress is the norm; the Chamic languages in the modern day follow the Mon-Khmer pattern without exception. The question that we ask here is then whether we might expect prosody to change 'faster' than morphology. If so, it is very likely that an intermediate state of ancestral Chamic involved a strict word-final stress pattern; then, the combination of the need to maintain a prosodic distinction between function and content (see e. g. Selkirk 1996) and an apparent universal tendency to prefer the placement of main stress on lexical items would seem to strongly promote the adoption of an entirely prefixing morphological system.

We substantiate our interpretation of Chamic and briefly describe an interesting case of contact-induced system shift: as per Kruspe (2004: 82), Semelai appears to contain *two* systems of morphological affixation, one indigenous 'non-concatenative' traditionally Mon-Khmer system, and one 'concatenative' borrowed from Malay. The distinction between them lies in the sensitivity of the affixation process to the prosodic nature of the root: the indigenous system attaches morphemic templates strictly to the left edge of the stressed final syllable (regardless of the shape of the root – thus both prefixation and infixation arise), while the borrowed system is characterised by adherence to the edge of the word. The relevance to the Chamic case lies in the strictness of the indigenous pattern: a prosodic shift in Chamic *combined* with the adoption of these Mon-Khmer-isms in morphological application would inevitably produce the exclusive prefixing which we seek to explain.

3.3 The borrowing of non-affixational morphological patterns.

The major portion of our discussion so far has focused on the status of individual borrowed morphemes, and their effects on the morphological system of a language; however, it is interesting to additionally consider the potential for transfer of other types of morphological template. The use of reduplication (which, in this work, we assume to be morphological, sidestepping a significant history of debate) is often referenced in the context of creolisation and pidginisation, but we suggest that there are instances in which it is most meaningfully considered in the same light as the borrowing of bound morphology: previous work exists on Iwaidja (Evans 2009) and Oroqen (Li & Whaley 2000). Consider the very familiar example of English 'shm'-reduplication, in which the repetition of a word with a prefixed *s(c)hm-* indicates irony or derision: *a fancy-schmancy hotel*. This almost certainly originates in Yiddish (see e. g. Nevins & Vaux 2003) and is well-integrated into the English morphology, despite the relative paucity of other reduplicative constructions in English; we can provide no analysis that does not treat this as a structural borrowing, carrying over a physical template, a phonological element (*shm-*), and a semantic value.

Reduplication, both total (whole-word copy, e. g. *luwak-luwak* 'civets') and partial (involving some change or deletion: e. g. *perlahan-lahan* or *sayur-mayur*) is highly productive in Malay/Indonesian. Total reduplication in particular is the major pluralisation strategy in Malay. We have previously described the replacement of this in Sri Lanka Malay; however, it also seems plausible that total reduplication has diffused outwards from Malay into languages in which it did not have a significant presence. In both Jahai and Semelai, reduplication seems to have been borrowed from Malay (it is not a particularly significant feature in Mon-Khmer; Khmer itself prefers synonym compounding for related semantic purposes), in Semelai almost exclusively in order to express verbal morphology, but in Jahai additionally marking plurals, a clear borrowing from Malay. In Madurese, an Austronesian language of Eastern Java and the island of Madura, the typical reduplicative plural does not form via 'total' reduplication, but rather via an unusual process of final-syllable copy: thus *ku-buku* 'books', *lem-malem* 'nights'. Javanese and Indonesian influence seem to be causing the gradual erosion of this usage in favour of total reduplication (Davies 2010: 130). We see a similar adoption of total reduplication, along with the related process of synonym compounding, in Manchu (Gorelova 2002: 380-383) under heavy influence from

Mandarin Chinese; neither process is a significant feature of other Tungusic languages.

Reduplication is one of the few morphological strategies available within Cantonese, as in many of the Sinitic languages. While many of the attested reduplicative patterns in Cantonese are of uncontroversially Sinitic origin, of particular interest to us is the widespread (particularly colloquially) use of 'expressive' ABB reduplication, which serves to add emphasis to adjectives – A here is an adjective or verb, and BB a reduplicated semantically-related syllable (Matthews & Yip 1994: 162-165).

	cau4-bang1-bang1	*hak1-mang1-mang1*	*jyun4-luk1-luk1*
(19)	smelly-poo-poo	black-dark-dark	round-wheel-wheel
	'very smelly'	'pitch-dark'	'very round'

In Zhuang, the largest Tai language of South-West China, this pattern is similarly productive, and serves a similar semantic function (Milliken 1998: 194; Matthews 2007):

	rang-ngau-ngau	*cok-maet-maet*
(20)	fragrant-sweet-sweet	strong-?-?
	'extremely fragrant'	'very strong'

A similar construction arises in Kam, another Tai language of the region (Milliken 1998):

	ton11-tok31-tok-31	*pʰaŋ35-ŋaŋ53-ŋaŋ53*	*toŋ53-tɐt55-tɐt55*
(21)	round-plate-plate	tall-up-up	crooked-twist-twist
	'round'	'tall'	'crooked'

Matthews suggests based on this evidence from Zhuang and Kam that the expressive ABB construction has an ultimately non-Sinitic origin (a 'South-East Asian areal feature') in Cantonese (2007: 228); it seems worth scrutinising this conclusion with more care.

The extent of this particular style of ABB reduplication in Mandarin seems to lack consensus; it is not described in traditional grammatical descriptions, or in some surveys of reduplication in Mandarin (e. g. Meng 2012), but this may simply be due to its status as a relatively colloquial usage. Matthews claims that it is a marginal pattern in Mandarin, and suggests that existing examples are likely due to diffusion from southern dialects (presumed to be in heavier contact with non-Sinitic languages). However, Wang (2010) collects a significant dataset suggesting

the existence of a similar pattern in Mandarin, of which we present a small sample[17]:

	辣酥酥	泪汪汪	乱糟糟
	la4-su1-su1	*lei4-wang1-wang1*	*luan4-tsao1-tsao1*
(22)	pungent-crisp-crisp 'very spicy'	tears-vast water-vast water 'tearful'	chaos-bad-bad 'very messy'

T'sou (1978) claims that expressive ABB reduplication arises in Mandarin, Min, and Burmese; slightly further afield, Lin (2011) provides examples from Dongshi Hakka (Taiwan).

If we are to accept the prevalence of ABB reduplication within Mandarin in particular, and more generally within the Sinitic languages, then some doubt is cast upon the argument that the Cantonese use of these forms is of external origin. However, this does not immediately imply that no borrowing has taken place; we might suggest instead that the forms in Zhuang and Kam – both Tai languages occupying relatively minoritarian positions in Southwestern China, with consequently heavy Sinitic influence – are derived from contact with Cantonese or other Sinitic languages. Within the Tai languages, a superficially similar pattern of ABB reduplication arises in Shan (e. g. *kʰo4 kʰik4 kʰik4* 'laughter'; Lengtai 2009: 90) but is restricted to the domain of onomatopoeia and serves no function of emphasis. No evidence of any similar pattern arises in Thai or Lao (Roffe 1975); across the Southwestern Tai languages including Shan, Ahom Tai, and Tai Phake (Phukan 2009), we see instead a tendency towards ABAC reduplication. The picture is complicated slightly by the fact that 'Zhuang' taken as a whole represents much of the internal diversity of the Tai languages; however, we may nonetheless suggest that this pattern is borrowed into Zhuang from Sinitic, rather than the other way around.

[17] Glosses are original and were not provided in Wang 2010.

4. Conclusions and outlook.

At the end of section 2, we described situations in which morphological borrowing may 'appear' to occur (fossilised infixation in Thai), and in which the genesis of new affixation in the recipient language occurs without the 'input' of the morphological systems in the donor (novel morphemes in Meitei). We also briefly revisit the situation of loan-verb morphology in the Philippine language Ilocano: borrowed roots are verbalised using the Spanish infinitive endings -ar, -er, and -ir, seemingly unconditioned by phonology, despite the fact that by the time of the first Spanish influence in the Philippines the suffixes -er and -ir were no longer productive in Spanish (Rubino : 345). What this suggests to us is that the copying of Spanish infinitival suffixes in Ilocano is sensitive to the apparent use of these forms in Spanish, but simultaneously *not* sensitive to their productive/unproductive status in the morphology. We might therefore surmise that any model of contact processes that we codify – although we do not intend to do so explicitly here – must account for this separation between morphological information and other strata of the grammar. In particular, if we imagine the genesis of a contact-induced change as occurring within the repertoire of a single speaker, then we note that in the transition from individual-speaker to large-scale change, the original speakers' grammatical awareness within both donor and recipient language is imperfectly transmitted. A somewhat related question which we have addressed insufficiently is the issue of the morphology-nonmorphology interaction as it pertains to the effects of borrowing. We propose and justify possible prosodic effects on the selection and retention of morphology in Chamic, and simpler phonological pressure promoting or constraining the copying of various suffixes; we also consider semantic aspects of the use of morphology, such as the role of experiencer-subjects in dictating case in Sri Lanka Malay. We have been unable to address the interaction of morphology and other systems in particular detail.

We do not find any particular reason to pay further attention to 'faithfulness' beyond the relatively obvious note that heavier contact may lead to higher incidences of particularly faithful copies; the concept does, however, illuminate the operation of the elements that we refer to as *matter* and *pattern*, given that in order to have a clear concept of fidelity we must necessarily determine, even if only conceptually, the 'distance' between one replica of a

particular morpheme and a hypothetical other. The effect of morphological borrowing on the systems within a language seems to be best described in terms of the *propagation* of features of an individual borrowing: we (re)visit the situation of prefixation and its origin in Kannada, and note the possibility for a relatively 'small' set of morphological borrowings to promote an unexpectedly significant change. We additionally provide evidence that the borrowing of numeral classifiers is heavily promoted by the borrowing of cardinal numbers, and further evidence that the status of non-affixational morphology is clearly influenced by language contact phenomena in much the same way as the morphology most frequently described, and that the borrowing of reduplication has potentially interesting implications.

An important element of this discussion has been the clearer delineation of subcategories of what we describe as 'pattern': two different cases of pattern-replication may differ from each other in their relationship to semantics, to categories within the source and target language, and to the underlying structure of and constraints on affixational morphology. The potential for variation across these distinct elements is as significant as the potential for variation in the level of overall matter- versus pattern- borrowing: a given case may involve little semantic and categorical change, but profound alteration in the physical-typological aspects of the morphology (e. g. Cham, to an extent Sri Lanka Malay), may produce a significant difference in available categories without any corresponding change in the affixation mechanism (e. g. most instances of the acquisition of gender-marking), or may vary in several of these 'variables' at a time. The perceived value to subdivision here is, however, somewhat at odds with our goal to streamline and schematise our understanding of contact phenomena; the reconciliation of these factors likely requires more sophisticated theoretical machinery than we establish here.

5. References.

Adamou, E. 2012. Verb morphologies in contact: evidence from the Balkan area. In: Morphologies in Contact, ed. M. Vanhove, T. Stolz, A. Urdze, H. Otsuka, p. 143-163. Berlin: Akademie Verlag.

Aikhenvald, A. 2003. Mechanisms of change in areal diffusion: new morphology and language contact. Journal of Linguistics , 39 (1). pp. 1-29.

Aronoff, M. and Sridhar, S. N. 1988. Prefixation in Kannada.

Anderson, G. D. S. 2006. The Munda Verb. Berlin: Mouton de Gruyter.

Andronov, M. S. 1980. The Brahui language. Moscow: Nauka.

Baki, A. 2012. Language contact between Uyghur and Chinese in Xinjiang, PRC: Uyghur elements in Xinjiang Putonghua. In: Int'l. J. Soc. Lang. 215, pp. 41–62.

Bakker, P. 2006. The Sri Lanka Sprachbund: The newcomers Portuguese and Malay. In: Linguistic areas. Convergence in Historical and Typological Perspective, ed. Y. Matras, A. McMahon, & N. Vincent, 135-159. Basingstoke: Palgrave Macmillan.

Benjamin, G. 1993. Grammar and polity: the cultural and political background to standard Malay. Working papers (National University of Singapore. Department of Sociology), no. 88.

Bickel, B. 1997. The Possessive of Experience in Belhare. In: Papers in Southeast Asian Linguistics No. 14: Tibeto-Burman Languages of the Himalayas, ed. David Bradley. 14: 135-155. Pacific Linguistics, the Australian National University.

Bickel, B. and Gaenszle, M. 2007. Generics as First Person Undergoers and the Political History of the Southern Kirant. Conference presentation: 7th Biannual Meeting of the Association for Linguistic Typology, Paris, September 27, 2007.

Blust, R. 2013. The Austronesian Languages. Asia-Pacific Linguistics Open Access Monographs. Canberra: Australian National University.

Burenhult, N. 2005. A grammar of Jahai. Pacific Linguistics 566. Canberra: Australia National University.

Butt, M. and Ahmed, T. 2011. The redevelopment of Indo-Aryan case systems from a lexical semantic perspective. Morphology 21:545-572.

Chelliah, S. 1997. A Grammar of Meithei. Berlin: Mouton de Gruyter.

Csato, E. 2012. On the sustainability of inflectional morphology. In: Copies versus Cognates in Bound Morphology, ed. L. Johanson & M. Robbeets, 371–380. Leiden: Brill.

Davies, W. 2010. A Grammar of Madurese. Berlin: Mouton de Gruyter.

Donegan, P. and D. Stampe. 2004. Rhythm and the synthetic drift of Munda. In: The Yearbook of South Asian Languages and Linguistics 2004, ed. Rajendra Singh, 3-36. Berlin: Mouton de Gruyter.

Droese, E. 1884. Introduction to the Malto Language. Agra: Secondra Orphanage Press.

Ebert, K. E. 2003. Kiranti languages: an overview. In: The Sino-Tibetan Languages, ed. LaPolla, R. & Thurgood, G. 336-352. London, New York: Routledge.

Ebert, K. E. 2003. Camling. In: The Sino-Tibetan Languages, ed. LaPolla, R. & Thurgood, G. 336-352. London, New York: Routledge.

Emeneau, M. B. 1956. India as a linguistic area. In: Language and Linguistic Area: Essays by Murray B. Emeneau, 1980, 105-126. Stanford: Stanford University Press.

Engesaeth, T., M. Yakup, and A. Dwyer. 2009. Teklimakandin Salam: hazirqi zaman Uyghur tili qollanmisi. / Greetings from the Teklimakan: a handbook of modern Uyghur. Lawrence: University of Kansas Scholarworks.

Evans, N. 2009. Doubled up all over again: borrowing, sound change and reduplication in Iwaidja. In: Morphology, Vol. 19, Issue 2, 159-176.

Gardani, F. 2012. Plural across inflection and derivation, fusion and agglutination. In: Copies versus Cognates in Bound Morphology, ed. L. Johanson & M. Robbeets, 71–99. Leiden: Brill.

Gerner, M. 2004. Expressives in Kam (Dong) : a study in sign typology (Part I). In: Cahiers de linguistique - Asie orientale, vol. 33 n°2, 159-202.

Gorelova, L. 2002. Manchu Grammar. Brill: Leiden–Boston–Köln.

Hildebrandt, K. H. 2004. A grammar and glossary of the Manange language. in Genetti, C.(ed.). Tibeto-Burman languages of Nepal: Manange and Sherpa. Pacific Linguistics.232-440

Huffman, F. E. 1986. Khmer loanwords in Thai. In: A Conference on Thai Studies In Honor of William J. Gedney, ed. R.J. Bickner et al.,199-210. Ann Arbor, Michigan: University of Michigan.

Krishnamurti, Bh. 2003. The Dravidian Languages. Cambridge: Cambridge University Press.

Lengtai, A. 2009. Shan phonology and morphology. PhD thesis. Bangkok: Mahidol University.

Li, F. and Whaley, L. J. 2000. Emphatic reduplication in Oroqen and its Altaic context. Linguistics. Vol. 38, Issue 2, 355-372.

Lin, H. 2011. Special tonal alternations in Dongshi Hakka reduplication. In: Taiwan Journal of Linguistics, Vol. 9.2, 1-66.

Mahdi, W. 2005. Old Malay. In: The Austronesian Languages of Asia and Madagascar, ed. Adelaar, K. A. & N. Limmelmann, 326-345. London, New York: Routledge.

Malchukov, A. 2006. Yakut interference in North-Tungusic languages. In: Turkic Languages in Contact, ed. H. Boeschoeten & L. Johanson, 122-138. Wiesbaden: Harrassowitz Verlag.

Masica, C. 1976. Defining a linguistic area: South Asia. Chicago: University of Chicago Press.

Masica, C. 1991. The Indo-Aryan languages. Cambridge: Cambridge University Press.

Matisoff, J. A. 1996. Contact-induced change, genetic relationship, and scales of comparison. The Fourth International Symposium on Language and Linguistics, Thailand. 1591-1612. Bangkok: Mahidol University.

Matisoff, J. A. 2003. Aslian: Mon-Khmer of the Malay Peninsula. In: Mon-Khmer Studies: A Journal of Southeast Asian Linguistics and Languages 33, 1-58.

Matras, Y. and Sakel, J. 2007. Investigating the mechanisms of pattern replication in language convergence. Studies in Language 31, 829-865.

Matras, Y. 2009. Language contact. Cambridge: Cambridge University Press.

Matthews, S. and Yip, V. 2003. Cantonese: A Comprehensive Grammar. London: Routledge.

Mazaudon, Martine. 2003. Tamang. In Thurgood, Graham and LaPolla, Randy J. (eds.), The Sino-Tibetan Languages, 291-314. London: Routledge.

Meng, C. 2012. A description of ideophonic words in Mandarin Chinese. PhD thesis. Leiden: Universiteit Leiden.

Nedjalkov, I. V. and Nedjalkov, V. P. 2007. Reciprocals, sociatives, comitatives, and assistives in Yakut. In: Reciprocal Constructions, ed. V. Nedjalkov. Amsterdam: John Benjamins.

Nevins, A. & Vaux, B. 2003. Metalinguistic, shmetalinguistic: the phonology of shmreduplication. Proceedings from the Annual Meeting of the Chicago Linguistic Society. 39 (1).

Noonan, M. 2003. Chantyal. In: The Sino-Tibetan Languages, ed. LaPolla, R. & Thurgood, G. 315-334. London, New York: Routledge.

Noonan, M. 2003. Nar-Phu. In: The Sino-Tibetan Languages, ed. LaPolla, R. & Thurgood, G. 336-352. London, New York: Routledge.

Noonan, M. 2008. Contact-induced change in the Himalayas: the case of the Tamangic languages. In: Language Contact and Contact Languages, ed. P. Siemund & N. Kintana, 81-107. Amsterdam: John Benjamins.

Pakendorf, B. 2007. Contact in the prehistory of the Sakha (Yakuts): Linguistic and genetic perspectives. PhD thesis. Leiden: Universiteit Leiden.

Pakendorf, B. 2009. Intensive contact and the copying of paradigms: en Even dialect in contact with Sakha (Yakut). Journal of Language Contact, Varia 2, 85-110.

Pan Hla, N. 1989. An introduction to Mon language. Center for Southeast Asian Studies, Kyoto University.

Phukan, J. N. 2009. An introductory primer and grammar of Ahom (Tai) language. Accessed online: http://www.vjf.cnrs.fr/brahmaputra/doc/Ahom_Primer.pdf

Roffe, G.E. 1975. Rhyme, reduplication, etc. in Lao. In: Studies in Tai Linguistics in Honor of William J. Gedney, ed. J.G. Harris and J.R. Chamberlain, 285-315. Central Institute of English Language.

Rubino, C. 2005. Iloko. In: The Austronesian Languages of Asia and Madagascar, ed. Adelaar, K. A. & N. Limmelmann, 326-345. London, New York: Routledge.

Selkirk, E. 1996. The prosodic Structure of Function Words. In: Optimality Theory in Phonology: A Reader, ed. J. McCarthy. Blackwell Publishing.

Simpson, A. 2005. Classifiers and DP structure in Southeast Asia. In: The Oxford Handbook of Comparative Syntax, ed. G. Cinque and R. Kayne, 806-838. Oxford University Press: Oxford.

Slater, K. W. 2003. Mangghuer. In: The Mongolic Languages, ed. J. Janhunen, 307-323. London, New York: Routledge.

Slomanson, P. 2011. Dravidian Features in the Sri Lanka Malay verb. In: Creoles, their substrates, and language typology, ed. C. Lefebvre, 383-411. Amsterdam: John Benjamins.

Stolz, T. 2012. Survival in a niche. On gender-copy in Chamorro. In: Morphologies in Contact, ed. M. Vanhove, T. Stolz, A. Urdze, H. Otsuka, p. 93-143. Berlin: Akademie Verlag.

T' Sou B.K. 1978. Sound symbolism and some socio– and historical linguistic implications of linguistic diversity in Sino-Tibetan languages. In: Cahiers de linguistique - Asie orientale, vol. 3 n°1, 67-76.

Thomason, S. and Kaufman, T. 1988. Language contact, creolization, and genetic linguistics. Berkeley/Los Angeles: University of California Press.

Thurgood, G. 1999. From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change: With an Appendix of Chamic Reconstructions and Loanwords. Honolulu: University of Hawai'i Press.

van Driem, G. 1987. A Grammar of Limbu. Berlin: Mouton de Gruyter.

Weinreich, Uriel. 1953. Languages in contact. The Hague: Mouton.

Wichmann, S. and Wohlgemuth, J. 2008. Loan verbs in a typological perspective. In: Aspects of language contact. New theoretical, methodological and empirical findings with special focus on Romancisation processes, ed. Stolz, T.; Bakker, D. & Salas Palomo, R., 89-121. Berlin/New York: Mouton de Gruyter.

Yakup, A. 2005. The Turfan dialect of Uyghur. Wiesbaden: Harrassowitz Verlag.

Yip, P.-C. and Rimmington, D. 2006. Chinese: an essential grammar. (Second edition) London, New York: Routledge.