The borrowability of structural categories

Yaron Matras

1. Introduction

The question of the "borrowability" of categories has often been equated with the presence or absence of constraints that rule out the borrowing of certain kinds of structures (cf. Campbell 1993, Thomason 2001 and elsewhere). I use the term here in a different sense. Borrowability is taken to mean the likelihood of a structural category to be affected by contact-induced change of some kind or other (whether matter- or pattern-replication; see Matras and Sakel 2007). From a strictly structure-oriented point of view, one might interpret this as the "ease" with which a category can be re-shaped through contact. I am not quite happy with this formulation, either, since it leaves open the source of the process and its motivation. Nor is the issue resolved by re-stating the obvious, namely by claiming that there is a link between the sociolinguistic norms of a speech community, the intensity of cultural contacts, and the outcomes of structural processes of change (cf. Thomason 2001, following Thomason and Kaufman 1988); for such a statement does not account for the fact that the borrowing of some categories requires more intensive contact than that of others. In other words, it fails to explain the hierarchical relationship between individual positions on the borrowing continuum.

When we speak of "ease" of borrowing, we are referring implicitly at least to the communicative behaviour of speakers in a bilingual setting and to changes in that behaviour that have a long-lasting effect on the shape of the language that they use. What interests us in this connection is the likelihood that, in respect of a particular structure which serves a particular function in language processing, speakers might give up the separation of two sub-components within their linguistic repertoire – the two "languages" – and begin to employ the structure in question regardless of the choice of language. Bilingual speakers of English and German, for example, take for granted that the concepts COMPUTER, DOWNLOAD, and INTERNET are common to both sets of communicative interactions in which they normally engage: those where the chosen language of conversation is English, and those where it is German. Bilingual speakers of Domari and Arabic (see Matras,

this volume) are fully at ease with the fact that the entire system of clause combining and connectors is shared by their two languages; a shift in the interaction setting will lead them to switch into another "language",[1] and this will affect the selection of various structures – vocabulary, inflections, anaphora and deixis, and so on – but it will not affect strategies of clause-combining, which remain the same in all settings (i.e. for both languages). And for speakers of Macedonian and the local Turkish dialects spoken in that country (see Matras and Tufan, this volume) the way of organizing information in copula sentences is identical regardless of the language that is being spoken. This, in essence, is the core of the diachronic process that we call "borrowing". With "borrowability", then, we mean the likelihood that speakers will give up the separation between their "languages" – the mental demarcation line that divides their overall repertoire of linguistic structures – in respect of a particular function-bearing structure (a "category").[2]

2. Borrowing hierarchies

Essentially two kinds of generalizations have been proposed concerning the borrowing of grammatical categories. Those of the first kind relate to the frequency with which a category may be affected by contact-induced change. Generalizations of the second kind suggest an implicational relationship between the borrowing of individual categories: the borrowing of one category is understood to be a pre-condition for the borrowing of another.

The majority of observations on grammatical borrowing belong to the first group (cf. Haugen 1950, Heath 1984, Thomason and Kaufman 1988, van Hout and Muysken 1994, Stolz and Stolz 1996 and 1997, Winford 2003, Aikhenvald 2006). Some statements are based on casual impressions only, while others report the results of counting exercises performed on a corpus. An issue that merits attention is the distinction between the counting of tokens and the counting of types (cf. also van Hout and Muysken 1994: 42–43): can we consider nouns to be more borrowable than, for instance, conditional particles, simply because nominal tokens occur in a corpus more frequently than conditional particles? Surely, token frequency will tell us how often a borrowed form is *used*, but it will not necessarily reveal how likely it is to be borrowed? Counting types, in turn, raises problems of its own: Can we conclude on the basis of type-frequency that adjectives, for instance, are easier to borrow than conditional particles, considering that the first constitute an open class of nearly unlimited types (for the purposes of any practical

comparison), while a language is likely to have just a very restricted inventory of conditional particles (if at all more than just one)? Such issues make it difficult to compare frequency-based hierarchies drawn from conversational corpora.[3]

This kind of dilemma does not present itself when comparing the grammatical (and lexical) systems of different languages in a sample (whether a structured one, or a casual one). Here, we are interested in the number of languages within the sample in which category X has been re-shaped as a result of contact. The more languages show borrowing affecting a certain category, the higher the frequency of borrowing for that particular category in the sample. We might then say that this category is "more likely" to be borrowed, relative to other categories.

Implicational hierarchies entail frequency hierarchies,[4] but go a significant step further in suggesting a constraint on the occurrence of borrowing with any lower-ranking category. The usual format of the statement is: Y is not borrowed unless X is borrowed as well (cf. Moravcsik 1978, Stolz 1996, Matras 1998 and 2002, Field 2002, Elšík and Matras 2006). The postulation of implicational borrowing hierarchies thus goes beyond the assumption that categories have different susceptibility to contact-induced change. It suggests that the process of contact-induced change follows, to some extent at least, a predictable pathway, with one stage leading as a pre-requisite to another. Moravcsik (1978) had attempted in this way to link the borrowing of non-nouns to nouns ("no non-nouns are borrowed unless nouns are also borrowed"), inflectional to derivational morphology ("no inflectional morphology is borrowed unless derivational morphology is also borrowed"), and so on, resulting in a web of inter-dependencies among various structural types (cf. Field 2002 for a statement on agglutinative > inflectional morphology). Matras (1998, 2002) and Elšík and Matras (2006) concentrate their observations on inter-dependencies of what they consider "values" of the same category, i.e. members of a shared structural paradigm. Matras (1998) postulates a borrowing hierarchy 'but' > 'or' > 'and' with respect to coordinating conjunctions, and Matras (2002) and Elšík and Matras (2006) suggest 'necessity' > 'ability' > 'volition' with respect to expressions of modality, and many more.[5]

Frequency-based hierarchies and implicational hierarchies may complement one another. Stolz and Stolz (1996) and Ross (2001), both relying on frequency-oriented observations rather than strict implicational hierarchies, conclude that contact-induced change begins at the level of the organization of discourse, proceeds to the organization of the paragraph, utterance, and

sentence levels, and only then reaches the levels of the phrase and word. The postulation of this kind of implicational hierarchy rests in such cases on accumulated observations of a series of frequency hierarchies.

Borrowing hierarchies thus provide us with an opportunity to gain insights into the factors that prompt speakers to allow their language systems to converge around a particular structure. Explanations of borrowing generally take one of three directions: (1) The degree of borrowing is related to the intensity of exposure to the contact language, (2) The outcome of language contact is a product of the structural similarities and differences (congruence) among the languages concerned, and (3) Borrowability is a product of inherent semantic-pragmatic or structural properties of the affected categories. Issues such as prestige and domain-specialization of the languages typically fall under 1, while conjectures about functional "gaps" as motivating factors fall under type 2.

Our interest in the present context is in explanations of type 3. This interest derives from the realization that structures and paradigm values often behave in an asymmetric manner when it comes to contact-related change. Under the "prestige" or "intensity of contact" effect there is no a priori reason why 'but' should be more vulnerable and prone to borrowing than 'and'; in many cases structural congruence does not provide an answer to this hierarchical relationship, either (cf. Matras 1998). Where category or paradigm values consistently show unequal or asymmetrical behaviour in contact situations, factors promoting borrowing must be sought in the inherent properties that they possess. In trying to explain the borrowability of categories we must therefore return to our initial assumption that ease of borrowing reflects the ease with which speakers are willing to give up the separation of two "language systems" and allow them to converge or to fuse around a particular linguistic function. The question that we ask is therefore: What is it that makes one category (or category value) a more attractive candidate for "system conflation" than another?

Elšík and Matras (2006: 370ff.), following Matras (1998), argue that borrowing is motivated by cognitive pressure on the speaker to reduce the mental processing load by allowing the structural manifestation of certain mental processing operations in the two languages to merge. The need to do so arises especially around operations that gauge the presentation of propositional content to hearer expectations, for example connectivity and modality. In these domains, merger of the structures targets in the first instance those conceptual domains where the speaker's epistemic authority is in question, and the potential for tension at the interaction level is therefore greatest. This

occurs for example around the expression of condition, contrast, participant-external force, or other, more general conceptual complexity. This accounts for the direction of the hierarchy, which prioritizes category values such as "contrast" and "external force".

We should at this stage clarify the notion of "category". One of the factors impeding straightforward comparability among hierarchies postulated in the literature is the vagueness with which category labels are used. Some researchers have, for example, identified a category of "function words"; others speak of "particles", and others still make reference to a class of "pronouns". As "function words" we can classify anything from interjections and fillers on to definite articles and demonstratives – categories that show enormous variation in respect of their contact behaviour. "Particles" can include markers of modality, connectivity, aspect, and more, while "pronouns" are used with reference to such functionally diverse entities as anaphoric or third-person pronouns, indefinites, and participant deixis.

If we suspect that there is a link between category status and borrowing, then we must assume that "category" represents a functional notion, rather than just a constituent slot or a wholesale cover-term. Categories are understood here to be operational devices that trigger mental processing activities in communicative interaction: nouns name objects, interjections direct attention to emotive evaluations of the speech situation, connectors establish links between the processing of individual propositions, word order serves as a map to organize information at the utterance level, and so on. Their concrete representation in a given language is through a structural form, which may or may not have cross-linguistic equivalents. Our agenda is to accommodate the borrowing behaviour of categories within an explanatory model, one that accounts for the link between the processing function which the category triggers, and the degree to which speakers allow its structural representation to converge or fuse among the two (or more) components of their linguistic repertoire.

To be sure, different explanatory models may be appropriate for different structural components of language. There is no doubt that the borrowing of institutional terminology from a language that is dominant in the public or acrolectal domain is not a result of mere tension at the level of processing the speech interaction, but rather an attempt to extend the referential world of the "dominant" language into interactions in which the "minority" or "weaker" language is used. The borrowing of phonemes may, in turn, be simply instrumental in serving the authentic integration of loanwords without "distorting" them, by adjusting the phoneme system to accommodate them. Nevertheless,

these too are conversation-functional factors that motivate speakers to allow their "linguistic systems" to conflate around certain structures and categories.

Finally, we must briefly comment on the relevance of exceptions to postulated hierarchies. In a discussion that focuses on absolute constraints on borrowing and sets out to test their validity, the discovery of counterexamples can have a sensational effect in dismantling earlier claims. We aim here at taking what Aikhenvald (2006: 26) describes as the more "positive" route: understanding a variety of factors and preferences that facilitate structural diffusion among languages. In this spirit, it would be naïve and counterproductive to ignore tendencies that are followed by a substantial group of languages within the sample only because they are not followed by all, or indeed because they might be contradicted in one or two instances. Where there appears to be a motivation behind trends, one that is beyond pure coincidence, then these trends deserve our attention. Quite often, it is the counterexample that can be explained as resulting from a local, language-particular constraint that impedes the realization of common patterns in a particular instance.

The following sections mirror the organization of the language-oriented chapters in this collection, and are devoted to an evaluative overview of selected patterns arising from the discussion of the 27 sample languages.

3. Phonology

Phonology in particular is an area in which borrowings are traditionally considered to fill so-called "structural gaps", facilitated especially when borrowing does not entail changes to the actual phonemic system but merely to allophonic distribution (cf. Winford 2003: 55–56). The notion of "gap" is vague, given that languages have long been considered in descriptive linguistics to constitute autonomous, functional systems. We should therefore perhaps amend the definition to focus on bilingual speakers' quest for harmony among the two (or more) systems that constitute their linguistic repertoire; absence of harmony as a result of absence of a phoneme in one of those systems is presumably what is meant by a "gap". There is a functional motivation favouring consistency in the types and points of articulation as well as the distribution rules of allophonic variation, regardless of the speech situation in which language users find themselves, and hence pressure toward convergence of the two phonological "systems". At the same time, social norms and

awareness of identity and loyalty toward the group associated with the home language will counteract levelling. The process of phonological borrowing is the outcome of compromises between these two pressures.

Our sample shows three different types of change: (1) Incorporation of phonemes from a contact language in loanwords, (2) adjustment in the articulation of a phoneme following the model of the contact language, and (3) incorporation of a borrowed phoneme into the system of inherited words (substituting an inherited phoneme in some words, though not necessarily in the system as a whole). Changes of the second type may lead to simplification of the system, or to its enrichment through new distinctions, or they may simply alter the nature of certain phonemes, leaving the complexity of the system as a whole intact. In the sample, changes of the first type typically add to the phoneme inventory, as do in most cases changes of the third type. On the whole, then, our observation is that language contact in the cases under scrutiny here typically leads to an enrichment of the phonological system.

Another general observation is that contact-related change is more likely to affect consonants than vowels; indeed, we may even be able to postulate an implicational hierarchy of contact-related change:

(1) adoption of new consonants > adoption of new vowels

The reason behind this hierarchy is, however, likely to be rather trivial: It is a product of the fact that consonant inventories are generally larger, and so the potential for lack of overlap between consonant systems in contact is higher, resulting in greater pressure to adjust the consonant system. In fact, the hierarchy under (1) need not at all suggest that contact induced change in phonology *begins* with consonants, and it is not impossible that vowels are as prone, or even more prone to change in situations where there is no significant difference among the languages in the inventory of consonants.

Almost all languages in the sample incorporate loanwords along with at least some of their original phonemes, which are new to the recipient system. Examples are Macedonian Turkish /ts/ with Macedonian loans, the Vietnamese sounds /ʃ/, /f/. /v/, and /z/ with Chinese loans, the Domari pharyngeals /ħ/ and /ʕ/ used in Arabic loans, and Imbabura Quichua /b/, /d/, /g/, /ʊ/, /ʒ/, and vowels /e/ and /o/ in Spanish loans. This indicates that for speakers of the languages in question, the integration of lexical loans in an "authentic" manner, i.e. one that closely replicates their original use in the contact language, takes precedence over the preservation of the coherent phoneme structures of the recipient language. The system of the recipient language is adjusted in order

to accommodate loans in an unmodified form. Of the sample languages, only Mosetén and Biak appear not to extend their phoneme system to accommodate loanwords, while the overlap between the phoneme systems of Kriol and Jaminjung does not require Jaminjung speakers to make any special effort in order to accommodate Kriol loans.

The second type of process is the convergence of articulation modes and positions, which is often, as Winford (2006) suggests, a process affecting allophonic variation. There are numerous cases involving the introduction of allophonic variation, among them the interchange of /dʒ/ to /ʒ/ in Domari based on the Arabic model, of /q/ and /ɣ/ in Khuzistani Arabic based on the Persian model, and of /l/ and /r/ in certain positions in North-eastern Neo-Aramaic, based on the Kurdish model. In some cases, variation leads to a shift in articulation, as in the weakening of /h/ in Macedonian Turkish or the replacement of /h/ by /ɦ/ in some varieties of Yiddish, of pretone /o/ by /a/ in others, and a more general shift in this language from dental /l/ to velarized /ɬ/.

From a language user's perspective, all these are instances of harmonization of articulatory patterns, aiming to ease the burden of having to maintain complete separation of two distinct systems in different settings of conversational interaction. In some cases, phonemic distinctions in the recipient language are even given up in order to enable harmonization. Some varieties of Purepecha for instance lose the opposition between retroflex /ɽ/ and flap /r/, as well as between central /ɨ/ and front /i/, resulting in a vowel system that matches that of Spanish. In other cases the system becomes more complex, as with the introduction of pharyngealized consonants into Domari from Arabic and into Tasawaq from Tuareg, of long vowels into Rumungro from Hungarian, and of palatalization of stops into Kildin Saami and Yiddish from Slavic contact languages.

Informal observations lead us to believe that prosody is a domain of phonology that is particularly prone to contact. This can be the result of two interconnected factors. The first is the peripheral role that prosody has in conveying meaning, and the fact that it is a form of expression of emotive modes, operating at the speech act and utterance level, rather than the word level. This allows speakers to mentally disconnect prosody more easily from the matter or shape of words associated with a particular language, making it prone to change and modification in contact situations. The second factor may be the proven neurophysiological separation between prosody and other aspects of speech production, making prosody more difficult to control. Both factors may contribute to the fact that foreign "accents" are most persistent in the area of prosody.

We have in our sample little data on prosody, however. Several pairs of languages are reported to share prosodic features with their current contact languages, confirming the above hypothesis: They include Domari, Nahuatl, Rapanui, Rumungro, and Indonesian, and to some extent at least also Kurdish, some varieties of Yiddish, Hup, and Kildin Saami. This is already more than can be said about any other area of phonology, where we find borrowings and convergent tendencies, but no wholesale convergence.[6] Hence, we might carefully postulate the following frequency-based hierarchy of likelihood of complete convergence in the phonological system:

(2) prosodic features > segmental phonological features

Once again we need to emphasize that this hierarchy does not suggest that segmental phonological features are unlikely to be borrowed unless prosodic features are also borrowed; it merely reflects the tendencies toward full-scale convergence of the systems. In fact, it does not seem possible at this stage to point to any position within the phonological system (e.g. certain articulatory modes or positions, marked features, etc.) as being particularly prone to contact-induced change. It seems that the details of phonological change are entirely a product of the relations among the two systems – or congruence – and any statistics of change are likely to simply reflect the mere likelihood of the two phoneme systems in contact to share certain phonemes, and to differ with respect of others. The one additional generalization that we can make is that the borrowing of phonemes begins with the borrowing of lexical items that contain them:

(3) phonological features in loanwords > independent phonological features

Concluding this section, it seems that there are two alternative strategies that multilinguals can pursue in respect of phonology, taking for granted that language contact will lead at least to a transfer of lexical items from one language to another. The first is to maintain the complete integrity of the recipient language system by adjusting the phonology of any borrowed word to match that of the recipient system. It would appear that this strategy would be facilitated by widespread monolingualism in the recipient language, and the confinement of bilingualism to just a small or peripheral group of intermediaries. It is also possible that this strategy can be maintained for a while in situations where widespread bilingualism is a relatively new phenomenon,

or where speakers of the recipient language have no need to appear to have native-like or even good command of the contact (donor) language. In our sample, only Mosetén and Biak appear to adopt this kind of strategy.

The alternative, which is the route taken by most of our sample languages, is to maintain the authenticity of donor language items by adjusting the phonological system of the recipient language to accommodate phonological features of the donor language. This would seem to be facilitated by widespread bilingualism and the need for speakers of the recipient language to gain the approval of the donor language community. Authenticity in the pronunciation of loanwords is a token of the social value attributed to the donor language, and is emblematic of the social immersion into the donor language culture. The result is the incorporation of phonological features from the donor language into the recipient language. These will at first accompany loanwords from the donor language. With increased multilingualism and the need to operate regularly in two linguistic environments, it is advantageous for speakers of the recipient language – we assume that this is normally the language that occupies the weaker socio-political position, or a minority language – to allow major components of their phonological system to converge with that of the dominant, donor language, and so to rid themselves of the burden to maintain a separation of their two speech modes. This tendency of course competes with the sense of loyalty toward the group-language, which may favour maintaining an old system and limit innovation. The outcome is often a compromise in the form of an adjustment of certain aspects of the phonological system in favour of common patterns, or the adoption of some features but not others from the donor language, as we saw above.

4. Typology

A number of languages show signs of movement between morphological types: There are changes from polysynthetic to less polysynthetic structuring in Nahuatl and Imbabura Quichua, Otomi, and Guaraní, from an agglutinative type to a more isolating type in Indonesian, and from an agglutinative to a more analytic type in Purepecha. An increase in reliance on reduplication is found in (agglutinative) Likpe under the influence of (isolating) Ewe. On the other hand there is some acquisition of agglutination in Hup, and in some traits perhaps also in Kurmanji Kurdish as well as in Rumungro: In Kurmanji, the agglutination of case markers benefits from the presence of

inherited enclitic case markers, which historically form part of a circumposition.[7] In Rumungro, the adoption of agglutinative prefixes is a by-product of the almost wholesale adoption of indefinite markers, superlative markers, and a few other morpheme classes that are high on the "relevance" scale and so easily borrowable on functional grounds.

It is noteworthy that none of these developments seems to follow any predictable structural path, and the only common denominator is an accommodation to the patterns of a socially dominant contact language. In all cases, the drift begins in individual constructions such as adjective comparison or case marking, and it is yet to be seen whether it will continue to spread.

Frequencies and the evaluation of general trends are not applicable to our sample in the domain of typology, as some language pairs happen to belong to similar types (consider Domari and Arabic, Vietnamese and Chinese, Yiddish and Slavic languages), while others, among them those named above, show typological clashes with their contact languages. It is interesting that Macedonian Turkish maintains strong morphological agglutination despite considerable re-structuring in the domain of clause organization; indeed, the loss of converbal morphology in favour of grammaticalized conjunctions might be considered a small, yet not insignificant step in the direction of a typological drift. On this basis, we might conclude that typological drift begins at the clause level. In Khuzistani Arabic, the reinterpretation of the construct state ending and definite article as equivalents of the Persian *ezāfe* attribution marker does not constitute a drift in morphological type as such, but it does expose the path taken toward morphological re-analysis, here too in a possessive construction at the phrase level.

Perhaps one of the more outstanding typological shifts reported in the sample is the change in alignment in North-eastern Neo-Aramaic, modelled on the Kurdish ergative construction. An interesting aspect of the construction is that it does not mirror the complete ergative formation – in the absence of, for example, nominal case in Aramaic – but chooses instead only a number of pivotal features which it reconstructs with inherited means (see also Matras and Sakel 2007). Alignment is contact-sensitive elsewhere, too: We find an expansion of ergativity in Urban Manange in contact with Nepali, and the incipient loss of ergativity in Kurmanji in contact with Turkish.

In conclusion, although we do not have cases in our sample that display far-reaching changes in overall morphological typology, it is very evident that morphological type is certainly not immune from contact-induced change. It would seem fair to state, at least cautiously, that there is by and large an opportunist motivation for typological drift: It is subordinate to pressure to-

ward convergence in a particular salient construction, such as the possessive construction, clause linking, or indefiniteness marking, or else it is triggered by accidental similarities in the shape and position of markers with similar meaning.

5. Nominal structures

This is a diverse and complex domain, containing many different sub-categories, and it is perhaps not a surprise that only two languages in the sample, Jamingjung and Biak, are reported to show no contact influence at all under this heading. A prominent sub-domain of nominal structure is case, but here it is noteworthy that no borrowing of bound case markers is attested in the sample. The closest evidence of contact influence on case markers is the reliance in some contact varieties of Kurmanji on postposed markers as enclitics rather than as components of circumpositions. There are also indications of meaning extension of case markers, such as the ablative becoming genitive under Dutch influence in Indonesian, or the loss of a distinction between comitative and instrumental in Imbabura Quichua.

Adpositions on the other hand do show matter-replication; attested cases include Indonesian *sama* 'with' and *guna* 'for the purpose of' from Sanskrit, Spanish *de* in Guaraní, numerous Arabic prepositions in Domari, Spanish *por* and *para* in Purepecha, and Tuareg *ámmàs* 'inside' and *àláqqàm* 'behind' in Tasawaq. The preposition indicating 'between' is the most frequently borrowed, examples being Indonesian *antara* from Sanskrit, Spanish *entre* in Guaraní and other languages, and Arabic *bēn* in Domari. This gives some vague evidence in support of the hierarchy proposed by Elšík and Matras (2006) for the borrowing of local relations expressions in Romani dialects:

(4) peripheral local relations > core local relations

"Core" relations ('in', 'at', 'on') are borrowed less frequently than "peripheral" relations ('between', 'around', 'opposite'), and this finds some support in the appearance of 'between' as the most frequent borrowing in the sample.

Developments affecting gender marking include a shift in unmarked gender from feminine to masculine in Mosetén, the loss of neuter gender in NE Yiddish (as in the contact languages Lithuanian and Latvian), and the incipient system of nominal classifiers in Hup (classifying inanimates by shape, and animates by gender), adopted from Tukano. Definitely the most extensive

development in this domain is the borrowing of Chinese classifiers into Vietnamese. There are, in addition, some marginal phenomena such as the loss of gender in pronouns (in Rumungro as well as in North-eastern Neo-Aramaic). Our sample gives us the impression that gender in the narrow sense (of a two- or three-gender system) is more stable in contact situations than more differentiated classifier systems, where influence might be more extensive.

Nominal possession is a domain in which contact phenomena are fairly widespread. The most common change to possessive constructions is a modification of word order, often drawing on existing flexibility and enhancing the frequency of a more peripheral pattern to match that of the contact language. Examples of contact phenomena in possession are found in Domari, Macedonian Turkish, Rumungro, Khuzistani Arabic, Guaraní, and North-eastern Neo-Aramaic, concerning mainly the order, and in some cases the distribution of morphs and their meaning; incipient cliticization of the postposed pronominal possessor in Kurmanji; frequent postpositive possessors in Rapanui; use of a preposition bearing the same meaning as in the contact language in Indonesian; and use of a borrowed preposition in Guaraní.

Definiteness is known for its areal diffusion, but in the sample we have only few cases of contact developments in this domain. Rapanui illustrates that re-arrangement of definiteness rules may occur when definite markers exist in both systems prior to contact. Khuzistani Arabic shows selective retreat of definiteness marking in some constructions, in contact with a language with no overt marking of definiteness. Interestingly, North-eastern Neo-Aramaic borrows the Kurdish definite article *-ak-*.

Nominal morphology is most frequently replicated in the case of plural markers, which are often maintained in loan nouns, either as productive markers of plurality, as in Yiddish, Rumungro, and Tasawaq, or in conjunction with a native expression of plurality, as in Domari and Quichua. In Likpe we find replication of a specific plurality marking pattern. Plural markers may be said to occupy a position in between derivation and inflection markers. On the one hand, they are potentially linked to the expression of plural agreement elsewhere in the sentence, and so they operate at the level of the sentence rather than just the word. On the other hand, they indicate clear semantic opposition to singulars at the word level. Morphological plural marking thus meets the criteria for semantic transparency which is so often noted as a factor facilitating morphological borrowing (Moravcsik 1978, Matras 1998, Field 2002, Winford 2003). The direct borrowing of derivational morphemes is attested throughout the sample. Macedonian Turkish, Yiddish, Quichua, Purepecha, and Rumungro all borrow diminutive suffixes (with Kildin

Saami re-organizing its diminutive derivation based on a Russian model), and Quichua and Yiddish borrow agentive suffixes (Quichua *-dur* from Spanish, Yiddish *-nik* from Russian).

The overall pattern leaves us with a picture that is not incongruent with that reported on in contact linguistics so far. The most widespread changes are in the possessive construction. They affect the nominal phrase at the syntactic or morphosyntactic level, having to do primarily with the position of the possessor and possessed object, and partly with the arrangement of possessive morphology. This is in line with predictions that phrase-level borrowing will be more intense than word-level borrowing. Borrowing of bound markers favours in particular plural markers, diminutive and agentive derivational markers, and classifiers (but not gender markers), confirming that semantic transparency facilitates borrowing. Adpositions are more borrowable than bound case markers (borrowing of which is not attested in the sample), with *between* being the most borrowable in our sample, confirming the tendency of borrowing to favour peripheral relations, and so for the process of convergence to begin with remote, cognitively less accessible or conceptually more complex domains. In other domains, such as the distribution of case, definiteness, or gender assignment, languages may develop similarities, often by extending or limiting distributional rules. However, bound case and gender markers remain on the whole among the most stable features in the nominal domain, resisting especially direct replication of matter.

6. Verbal structures

Little attention has been granted in the literature to borrowing of features belonging to the domain of verbs (on the borrowing of lexical verbs see below); reports on the borrowing of TMA markers are quite rare. It is useful to consider the categories one by one. In the domain of tense, we see contact-induced similarities in the organization of the future tense in several languages: It is lacking in Domari and Arabic, it is suffixed in Hup and Tukanoan, and it shows a similar periphrastic structure in Kildin Saami and Russian. To this we might add the similar organization of the prospective aspect in K'abeena and Amharic.

Contact phenomena appear to be somewhat more frequent in aspect and aktionsart, where we find matter replication as well as shared patterns. Domari uses an Arabic habitual auxiliary *kān*, and Nahuatl introduces a progressive based on the Spanish model. We find aspectual use of the borrowed comple-

tive *ya* from Spanish in Guaraní, and *sudah* from Sanskrit in Indonesian, as well as similar expressions of experiential perfect in K'abeena and Amharic, while Likpe adopts a periphrastic present progressive similar to Ewe, and North-eastern Neo-Aramaic adopts a present progressive particle-turned-prefix based on a Kurdish model.[8] Yiddish re-organizes its verbal prefixes to replicate so-called Slavic Aspect distinctions (essentially, grammaticalized aktionsart), and Rumungro shows iteratives that are calqued on Hungarian.

Mood is similarly somewhat richer in contact developments, though differences in the structural organization of mood and modality make an exhaustive comparison somewhat difficult. Noteworthy are similarities in the use of the subjunctive in North-eastern Neo-Aramaic and Kurdish, and Domari and Arabic (cf. Matras and Sakel 2007: 843). Vietnamese prohibitive and conditional markers are borrowed from Chinese, while Kurmanji borrows the Turkish conditional marker *-ise*. Modality shows the most widespread contact phenomena, especially as regards matter replication. Almost half of the sample languages show matter replication of modality markers, such as Turkish *gerek* 'must' in Kurmanji, Arabic *lāzim* 'must' in Domari, Spanish *tiene que* 'must' in Rapanui, Arabic *mungkin* 'can' in Indonesian, Spanish *pudi* 'can' in Quichua, Malay *harus* 'must' in Biak, and more. The most common are markers of obligation (i.e. expressing external forces), followed in turn by necessity, possibility, ability, and desire. This hierarchy is almost always implicational, the only exception being Domari (which borrows markers for all meanings except ability):[9]

(5) obligation > necessity > possibility > ability > desire

The hierarchy proceeds from the most intensive external force, to the most participant-internal dimension. It is identical to the hierarchy identified by Elšík and Matras (2006) for the borrowing of modality markers in Romani dialects: necessity > ability > (inability) > volition. The more abstract theme in this hierarchy might be described as the degree of "speaker control", low speaker control correlating with high borrowability.

As far as Domari is concerned, its minor deviation from the implicational hierarchy can be explained by the fact that all the modality markers that it borrows from Arabic are impersonal expressions, or non-verbs. Even *bidd-* 'want', is a nominal, and its person inflection in Arabic follows the paradigm reserved for nominals, i.e "my-wish" etc. (and this inflection is carried over into Domari as well). The Arabic expression for ability, however, *'a-qdar-*, is an inflected verb, and although Arabic verbs are borrowable in Domari,

it competes with an inherited Domari verb *sak-*, which prevails. Thus, it is the formal inconsistency in the system of the donor language which in this case imposes a constraint that breaks the hierarchy.

Returning now to a general view of TMA and modality, we have seen the high density of (matter) borrowing in the domain of modality, in some cases also in mood, frequent matter and pattern replications in the area of aspect and aktionsart, and few cases of pattern replication in tense, all involving the future. This picture lends itself to an interpretation in terms of the hierarchy in (6), which depicts the likelihood of the respective categories to be affected by contact:

(6) modality > aspect/aktiosnart > future tense > (other tenses)

By and large, this hierarchy reflects both frequency, and implicational relationships. There is one case in the sample – Kildin Saami – where there appears to be contact influence in the arrangement of tense, but not in modality. Yiddish might be considered a case for contact influence on aktionsart, but it similarly lacks borrowed modal verbs from its Slavic contact languages, though to some extent this might be explained by the presence in this domain of Hebrew loans acquired through literary tradition.

The internal rationale of the hierarchy – which, once again resembles the findings for Romani dialects (cf. Matras 2002) – leads us to postulate again that external circumstances that limit the degree of speaker control – mood and modality in general – are the most contact-sensitive. They are followed by a qualification of the internal structure of the event – aspect and aktionsart – these too being beyond the immediate control of the speaker. Only then do we find contact influence in tense, the most intimate relationship between the event and the speaker's own perspective, though it is noteworthy that in our sample this is limited to the future tense, which identifies the event as being least stable and secure from the speaker's perspective. The overall theme is therefore once again the speaker's epistemic authority; its absence or weakening correlates with high borrowability.

Existential and possession verbs are affected by contact in several of the sample languages. In Domari, the Arabic copula is adopted in its function as a periphrastic expression of the habitual aspect, and it co-exists with the Domari enclitic copula. But Rapanui uses Tahitan and Spanish forms as copula, and both Indonesian and North-eastern Neo-Aramaic are reported to have developed copula forms through imitation of forms in their contact languages. The borrowing of 'have' is on the whole marginal. Spanish *tengo* is

employed in Rapanui, while in some other languages, such as Hup, we find calquing of constructions.

Contact phenomena in the area of voice and valency are almost exclusively pattern-oriented, and usually involve an increase in frequency distribution of an existing option: analytic reflexives in Quichua, the periphrastic passive in Purepecha, a causative marker derived from the verb 'to do' in Manange (copying the function of the Nepali affix *-aau*), and productive use of the inherited morphological causative in Rumungro. Vietnamese is exceptional in directly borrowing reflexive and passive markers from Chinese. Recent contact-induced grammaticalizations lead to the emergence of causative, passive, and reflexive markers in Hup, and to a reflexive in Likpe. It is difficult to further sub-divide this domain into category components, or to draw clear-cut connections with other components of the verb, other than to say that derivation generally appears much more contact-susceptible than tense, and perhaps at the same level as aspect. Both are in a sense statements about the internal organization of events, not directly connected to the speaker's position, but with no explicit evaluative statement concerning the truth-relevance or factuality of the event (as in mood or modality), either.

Moravcsik (1975) had drawn attention to the frequent use of incorporation strategies to accommodate borrowed verbs, and this typological discussion of verb borrowability has recently been revived (cf. Wichmann and Wohlgemuth, forthc.). There appears to be a near-consensus view that the borrowing of verbs is not, of course, impossible, but made more cumbersome in some languages due to the widespread tendency of verbs to be morphologically more complex (see Winford 2003: 52). In our sample, direct borrowing of verbs, without any formal adaptation, is found in Vietnamese, where there is no morphology in either the recipient or the donor language Chinese; in Likpe, where isolating Ewe contributes verb roots into an agglutinative structure; but also in many other languages, including Tasawaq, Quichua, Otomi, Guaraní, Hup, and K'abeena. There is thus obviously no universal constraint on the integration of borrowed verbs.[10]

Nonetheless, several languages in the sample prefer to apply an overt, morphological accommodation strategy when incorporating verbs of foreign origin into the lexicon. A favourite strategy is the use of so-called "light" verbs. Macedonian Turkish, Kurmanji, Domari, and Khuzistani Arabic belong to a larger isogloss covering the Caucasian–Mideastern–South Asian area, where mainly two light verbs are used, each combining with a root of nominal form of the borrowed verb. The distinction between the two light verbs is usually one of valency, and they usually derive from or are identical

to the lexical verbs for 'to make/ to do' and 'to be/ to become'. In Domari, the carrier verbs are semi-grammaticalized, and full forms *-kar-* 'to do' and '*-hr-* 'to become' co-exist with the abbreviated integration markers *-k-* and *-Ø-*. A somewhat similar strategy is found in Mosetén, where one of the verb adaptation markers is also a valency-augment, in Yiddish, where Hebrew-derived verbs are accompanied by *zayn* 'to be', and in Jaminjung, where loans have coverb status and are always used in combination with a native inflected verb. A series of other languages use a verbalizing augment which is otherwise employed to derive verbs from non-verbs as an integration marker: Mosetén, Nahuatl, Indonesian, Guaraní, Biak, Purepecha, Manange, Rumungro, and Yaqui. The borrowed verb itself usually appears in either the root form, an infinitive form, or an unmarked inflected form, quite often – for Spanish verbs, especially – the third-person singular present.

While no constraint on the borrowability of verbs can be upheld, it is nevertheless evident that a large number of languages require greater grammatical effort in integrating verbs than for the integration of nouns. The borrowing hierarchy

(7) nouns > verbs

expresses the grammatical "ease" or simplicity with which elements belonging to these two word classes can be integrated. Why is it that verbs require greater morphological integration effort, and what does this greater effort represent? On the one hand we find a pair of morphologically isolating languages like Vietnamese and Chinese, with no morphological complexity surrounding verbs in either the recipient or the donor language, and where verbs are integrated in a straightforward manner, just like nouns and other parts of speech. It is difficult, however, to attribute the need for explicit loan-verb adaptation markers even in other languages to the morphological complexity of verbs alone. Recall that most languages tend to integrate simple forms of the verb, such as the root, the infinitive, or an unmarked form. There does not seem to be, in those languages, any difficulty in stripping the target verb to its bare lexical essentials, before transposing it into the host morphology.[11]

I suggest instead that the difficulty lies in the conceptual complexity of the verb, and the fact that when borrowed and integrated, the verb is expected to perform two operations: The first is to serve as a referential lexical item – a content word, not dissimilar to a noun, adjective, or descriptive adverb. The second is to initiate the predication and so to serve as the principal anchor point for the entire proposition of the utterance. This latter function consti-

tutes its "verbness". It appears that borrowing of verbs is motivated by a similar need for modifying the inventory of lexical-referential expressions as the borrowing of nouns (and no doubt various specific semantic motivations could be postulated for groups of lexical content words). Speakers thus allow the lexical component of the verb to "cross" the mental demarcation boundary between languages, i.e. they license themselves to employ the same action/event signifier in any speech interaction. The bare lexical stem, however, is not always sufficient in order to assume the role of predication-initiator. A great number of languages therefore require this additional, crucial function to be explicitly marked out in the verbal expression; in other words, they need to transform the strictly "lexical" depiction of an action/event into a predicate. This is achieved through explicit marking of its "verbness".

A more detailed study is required in order to ascertain the conditions under which languages require some form of loan-verb integration. A first, banal observation is that a pre-requisite for the employment of loan-verb adaptation markers is the availability in the recipient language of a morphological procedure to derive verbs from non-verbs. Whereas an isolating language like Vietnamese may rely to a considerable extent on the pragmatics of morpheme juxtaposition as a way of (indirectly) marking out word classes, flectional languages will often require an additional means of identifying derived verbs.

This, of course, only explains a part of the story. It is clear that the similarities in loan-verb adaptation strategies found among the languages of the Caucasian–Mideastern–South Asian area are as much areally motivated (i.e. through contact and imitation among the languages) as they are functionally motivated by the respective morphological structures of these diverse languages. Moreover, not all languages that possess verb derivation strategies employ them with loan-verbs. And finally, languages are known to have changed their loan-verb adaptation strategy over time, without adopting any significant changes to their morphological typology. Thus, Romani appears to have shared the "light verb", valency-marking strategy of loan-verb adaptation during its early, Byzantine period, with languages of the Caucasus–Anatolia–South Asia "area" (see Matras 2002). It then transferred the function of marking out loan-verbs to a set of Greek-derived aspectual markers. Finally, in some contemporary dialects of Romani, such as Sinti, Kaale, and Vlax (spoken in or around Germany, Finland, and Romania/Hungary, respectively), loan-verb integration markers are being reduced altogether. For the time being, our principal conclusion can be that the hierarchy depicted under (7) applies for those cases where integration of a lexical item requires morphological support through derivational means.

7. Other parts of speech

The present section gives an overview of those elements often summarized together as "function words" or "unbound grammatical lexemes". Lumping them together in one section is a matter of convenience, and follows, as did the previous sections, the structure of the Language Convergence and Borrowing questionnaire, which forms the basis of the descriptive chapter. By discussing them under a shared heading, I am not suggesting at this stage that they share properties that motivate borrowing, nor that their behaviour in contact situations should, for any reason, be uniform or even similar at all.

7.1. Numerals

Numerals, in fact, are often considered low on the borrowing scale. This may derive from an assumption that all languages have some form of quantification. Although it is now known that not all languages possess systems for counting discrete entities, it is not necessarily the clash of systems of quantification that provides the motivation for the borrowing of numerals. Several types of borrowing involving numerals can be identified. Pattern replication appears in some languages: A decimal system is reported to have been adopted as a result of contact in Mosetén, Indonesian adopts a Javanese tag-lexeme indicating "teens" and re-organizes its earlier system of numeral juxtaposition above 10 accordingly, Hup adopts the Tukano quintenary system for numerals between 5 and 20, and a combination lexemes ('ten-and-one') replace single lexemes ('eleven') in some varieties of Kurmanji, replicating the Turkish arrangement. This rather small group suggests the following implicational hierarchy representing the likelihood of pattern-replication in numerals, which is yet to be confirmed by a larger sample:

(8) over 10 > below 10

More than two-thirds of the sample languages show some form of direct matter-replication of numerals. This includes most of the languages that reorganize their pattern of numerals, and which often employ borrowed numerals alongside the re-modelled "internal" or inherited system. In some cases, numeral replication is subject to sociolinguistic constraints, with contact-language numerals used as the preferred system for formal purposes such as citing dates and addresses and performing even simple mathematical tasks such

as counting (as opposed to the casual use of numerals as attributes), in transactions involving money, in names for coins of banknotes, or in some cases in the citation of grades. Such contextual splits are described for the use of Turkish numerals in Kurmanji, for the use of Hokkien numerals in colloquial Indonesian, for dates in Rumungro, and for Chinese numerals in Vietnamese, leading us to postulate a sociolinguistic hierarchy for the likelihood of use of borrowed numerals:

(9) more formal contexts > less formal contexts

This hierarchy reflects the fact that numerals enter languages through the dominance of the second language in formal and business transactions, and through education and other forms of institutional discourse. In many of the sample languages, especially those in post-colonial contexts, knowledge of the indigenous system of numerals is reported to be in decline, and the younger generation shows a clear preference for borrowed numerals. The adoption of borrowings in such situations clearly favours higher numerals over lower numerals, allowing us to postulate the following implicational hierarchy for the borrowing of cardinal numerals:

(10) higher numerals 1000, 100 > above 20 > above 10 > above 5 > below 5

This hierarchy appears related to some of the hierarchies postulated above, where borrowing is facilitated around conceptual complexity and inaccessibility. At the same time, higher frequency in casual language use of lower numerals is clearly a factor supporting the retention of native forms. Languages for which speakers are reported to be using native forms for lower numerals under 5 – either primarily, or alongside borrowings – but mainly or exclusively borrowings for numerals above 5 include Domari, some Kurdish speakers, Jaminjung (which has no native numerals above 3), Tasawaq, Otomi, Guaraní, Purepecha, Yaqui, Kildin Saami, and we might add Rumungro, where 7–9 are Greek loans into Early Romani, and 6 has been argued to be an early loan from Dardic (Indo-Iranian frontier languages). For some languages, such as Mosetén, Quichua, Nahuatl, and Biak, the cut-off point tends to be 10, while Hup shows a split at 20. Higher numerals show an independent susceptibility to borrowing. Tasawaq for instance borrows its lower numerals from Arabic, but its word for 100 from Tuareg; Rumungro generally shows Greek borrowings, but 1000 is Hungarian; and Vietnamese uses a Chinese word for 10,000.

Somewhat paradoxically, "0" ranks closer to the higher numerals 100, 1000, and so on: the only K'abeena borrowing is *zeeruta*, from Italian via Amharic, while in Rumungro it derives, like the higher numerals, from Hungarian. This is not surprising, and shows that cognitive complexity in the counting system operates in respect of the ability to easily identify and appreciate a quantity. This is hindered the greater the quantity, but it is similarly hindered in the absence of any quantity at all. An additional factor that no doubt plays a role is the relative formality of the term "zero", which is associated with mathematical and other formal notations and transactions, but not with everyday, casual expression of "nothingness".

Sample languages that do not borrow numerals are Macedonian Turkish, Khuzistani Arabic, Yiddish, and Manange (though incipient influence of Nepali on the numeral system is reported). In all but the latter, we can attribute the stability of numerals to a firm tradition of native-language education, media, and literacy, if not widespread among all speakers, then at least firmly anchored in the community and its history. This confirms once again that the borrowing of numerals is motivated not necessarily by "gaps" in the system of counting, but by a much more general accommodation to the language of formal institutions and the public domain in the way of conceptualizing and expressing formal transactions surrounding quantification.

The hierarchies presented in (9) and (10) are fully in line with the observations described for Romani dialects by Matras (2002) and by Elšík and Matras (2006), which lends support to their validity as universal indicators. A tentative case can be made for the following hierarchy of the likelihood of borrowing of ordinal numbers:

(11) lower ordinals > higher ordinals

This hierarchy is presented by Elšík and Matras for Romani. In the present sample it is confirmed by Kildin Saami and Rumungro, which use borrowed ordinals for 'first',[12] and Western Neo-Aramaic, which uses Arabic ordinals for '1–10', while Domari, Otomi and Purepecha generally rely on borrowed ordinals. Note that English (not part of the sample) is an exception to the hierarchy, having borrowed *second* but not *first* from Romance.

The ordinal 'first' is often a separate word, quite often suppletive to the rest of the ordinal paradigm. In some languages, this is also true of 'second'. This structural conspicuousness could be a factor promoting borrowing. In the assessment of Elšík and Matras (2006) the high borrowability of lower ordinals is a direct factor of this universal tendency to prioritize the ordinal 'first'

through lexical suppletion, which in turn is an expression of its cognitive saliency. Borrowing therefore simply follows the same path; in other words, the search for a renewable (=suppletive) item as an outstanding marker of the pragmatic saliency of "firstness" exploits the bilingual situation in recruiting an item from the contact language. We might therefore, in respect of ordinal numbers, postulate the following hierarchy of borrowability

(12) exclusivity > inclusivity

where "exclusivity" is taken to mean the separation of a single concept, entity, topic or state of affairs from a larger set – here in relation to the order of prominence, such as temporal sequence, or the attention granted to the object. Conceptually, this hierarchy is well in line with the contact-susceptibility of such properties as condition (see above), privative 'without, instead of, except for' (cf. Elšík and Matras 2006), contrast, phasal change ('already'), restriction ('only'), and the superlative (see below), all of which denote a broken chain of expectations, singling out our delimiting one entity from a presuppositional set.

7.2. Pronominal forms

The borrowing behaviour of so-called "pronouns" illustrates how limited a wholesale structural approach to category borrowing can be, and how it is the functionality of categories that motivates borrowing. Only Indonesian shows borrowing of personal pronouns, from Sanskrit and Javanese, into a system of highly differentiated, lexicalized terms of address and reference. Other contact-developments in pronouns are limited to the organization of the system of reference: Imbabura Quecha is reported to have developed a polite form of the second-person pronoun *kikin* on the basis of Spanish *Usted*, and in some heavily Hispanicized varieties of Guaraní, the inherited distinction between inclusive and exclusive is dissolved. Borrowing of other deictic and anaphoric forms is limited. Spanish *la* is used as an anaphor in Guaraní, the Arabic resumptive pronoun *iyyā-* is used in Domari in relative clauses, and Rumungro borrows the Hungarian deictic prefixes *am-* and *ugyan-* which are combined with Romani deictic stems.

Reflexive pronouns are borrowed in Tasawaq, Western Neo-Aramaic, and Rumungro. Yiddish shows an extension of reflexivity based on a Slavic model, and Indonesian calques a reflexive apparently on a Sanskrit model.

Reciprocal pronouns are borrowed in Domari and Western Neo-Aramaic, and Rumungro calques a Hungarian model. What motivates the borrowing of reflexive and reciprocal forms? Unlike straightforward deictic and anaphoric reference devices, reflexives and reciprocals may be said to constitute an extension of the derivational system of the verb, contributing to the layout of actors involved in and affected by an event. They are thus part of a construction that revolves around the verb's "actionality".

Interrogatives are borrowed into several languages. Those that stand out as more highly borrowable are the interrogatives for quantity ('how much'), borrowed into Domari, Otomi, and Manange, and time ('when'), borrowed into Domari, Indonesian, Quichua, and Rumungro. Borrowing also affects indefinites. Domari, Otomi, and Rumungro borrow all or almost all of their indefinite expressions, while Tasawaq and Purepecha borrow time indefinities, Guaraní the indefinites for person and thing, and Yiddish borrows indefinite markers. Although there is no direct, predictable link to other categories, borrowing in the domains of interrogatives and indefinites appears to be a more "advanced" stage of borrowing among Other Parts of Speech.

7.3. Connectors/conjunctions

The grammatical category that is by far the most susceptible to borrowing is that of connectors (see already Matras 1998). All languages in the sample borrow connectors, and the general picture confirms the implicational hierarchy postulated as universal in Matras (1998), and confirmed by Elšík and Matras (2006) for Romani (and recently by Stolz 2007 for a number of languages in contact with Italian):

(13) but > or > and

Sample languages that borrow all three connectors include Domari, Mosetén, Nahuatl, Kurmanji, Rapanui, Indonesian, Quichua, Otomi, Guaraní, Kildin Saami, and Western Neo-Aramaic; languages that borrow only 'but' and 'or' are Tasawaq, Purepecha, Vietnamese, Rumungro, K'abeena, and Likpe. No languages borrow 'and' without also borrowing 'but' and 'or'.

There are, however, a few languages that deviate slightly from the expected pattern. Macedonian Turkish borrows Macedonian *i* 'and' as well as *ili* 'or' and *a* 'or, whereas', but retains Turkish *ama* 'but'; however, the latter is identical to Macedonian *ama*, which is a Turkish borrowing (cf. Matras 2004).

Jaminjung uses the borrowed contrastive marker Kriol *ani* 'only' alongside its native *bugu*, while only borrowed forms are used for addition and disjunction; but this is due to the absence of any native connectors for addition or disjunction to compete with the borrowings. In Manange, Nepali *ani* 'and then' can also be used for clause coordination (cf. Stolz's 2007 discussion of Italian *allora*). Neither of these cases necessarily contradicts the hierarchy in (13). Biak, however, is reported to use the Indonesian disjunction marker *atau* 'or' and less frequently the addition marker *dan* 'and', but no mention is made of a borrowed contrastive marker. Hup borrows *ọu* 'or' from Tukano. The source is ultimately Portuguese, and is reported to have diffused widely in the area. Aikhenvald (2002), too, reports on Tariana *ou* 'or', with no borrowing of other Portuguese connectors. Similarly, Yaqui appears to borrow only Spanish *o* 'or'. The fact that counterexamples can be found does not invalidate the overall observation that contrast is a semantic-pragmatic feature that facilitates borrowing, nor of course that clause-combining is an operational domain that is prone to contact-related change. Most likely, certain constraints of a structural and perhaps also a cultural nature (conventions on structuring discourse and expressing overt contrast) override the universal tendency in some cases. Noteworthy is the cluster of Amazonian languages within which Portuguese *ou* diffuses, often via secondary sources only. Borrowing in the domain of coordinating conjunctions is missing only in Yiddish and Khuzistani Arabic.

Subordinating conjunctions are similarly a frequent target of borrowings. Among the complementizers, borrowing is almost entirely restricted to those that introduce factual clauses, which are borrowed in Domari, Khuzistani Arabic, Rumungro, Western Neo-Aramaic, and Likpe. Although modality has been shown to be contact-prone, at the level of the organization of the complex sentence factual complements show greater event independence and so greater effort is needed in order to process the connection between the two clauses (see already Givón 1990, Dixon 1995, 2006). Factuality at this level is thus quite in line, as a factor promoting borrowing, with contrast, limitation, exemption and other properties that we have so far encountered at the top of borrowability hiererachies. Also well in line with these tendencies is the high presence, among borrowed conjunctions that introduce adverbial clauses, of those that mark concessive relations (borrowed in Yiddish, Tasawaq, Indonesian, Quechua, Guaraní, Domari, and Likpe), causal relations (Domari, Mosetén, Nahuatl, Kurmanji, Rapanui, Jaminjung, Tasawaq, and numerous others that calque causal subordinators), and purpose (e.g. Domari, Mosotén, Nahuatl). High on the borrowing scale are also conditional subordinators

(borrowed in Domari, Mosetén, Indonesian, Quichua, and Guaraní), while the borrowing of temporal subordinators is often linked to that of conjunctions expressing purpose and cause. We can therefore postulate the following tentative borrowing hierarchies:

(14) concessive, conditional, causal, purpose > other subordinators
(15) factual complementizers > non-factual complementizers

Concession, of course, is tightly linked to contrast and unexpectedness. Condition is an expression of modality. Cause and purpose are both efforts to link independent events, as are factual complementizers, while cause constitutes in addition an explicit argumentation effort at the interactional level. The hierarchies (14)–(15) thus supply us with a series of semantic-pragmatic properties that are borrowing-prone.

7.4. Particles

Not many languages in the sample borrow phasal adverbs, but those that do show a clear implicational hierarchy

(16) yet, already > still > (no longer)

confirming that observed in Matras (1998) for Romani, as well as in van der Auwera (1998) for a sample of European languages. While Rumungro and Domari show borrowings in all positions, Jaminjung borrows 'yet', Guaraní only borrows 'already', while Otomi also has 'still'. The semantic opposition involved is one of change vs. continuation, in the first instance. While 'no longer' essentially expresses change, its position on the hierarchy is partly influenced by its tendency to be composed of several structural elements. It is therefore the first two positions on the (left of the) hierarchy that are the most meaningful, and which continue the theme of contact-susceptibility of contrast and discontinuity of pre-suppositional expectations. Another particle that shows frequent borrowing is 'again' (Domari, Mosetén, Kurdish, Jaminjung, Indonesian, Otomi), expressing an unexpected repetition of events.

Half the languages in the sample borrow focus particles, giving the implicational hierarchy

(17) only > too > (even)

once again in line with observations on Romani and other languages in Matras (1998). In fact, the particles 'only' and 'too' usually go together, but 'only' can be considered higher on the hierarchy, since Indonesian and Western Neo-Aramaic have borrowed 'only' but not 'too'. Khuzistani Arabic borrows Persian 'too' (*hem*), but since Persian 'only' is itself an Arabic loan (*faqat*) it is not identifiable as a borrowing in Khuzistani Arabic. The hierarchy in (17) indicates that restriction facilitates borrowing, while the proneness of focus particles (and indeed of phasal adverbs and repetition adverbs) to borrowing shows the vulnerability of the system of processing states of affairs and attitudes that are high on the relevance scale and that assess information in direct relation to existing hearer-sided presuppositions and expectations.

Fully consistent with this observation is the overwhelming tendency of the languages in the sample to borrow discourse markers (once again, cf. Matras 1998). There are only two languages that do not show borrowing of discourse markers: Biak and Vietnamese. There is no obvious explanation for the absence of borrowed discourse markers in these languages, except perhaps the fact that using native intonation and modal particles is considered a central characteristic of talking Biak and so an important identity marker,[13] and that Chinese influence on Vietnamese was transmitted to a considerable extent via the formal and literary language, rather than via oral discourse. We also find less extensive borrowing of discourse markers in languages with a tradition of native literacy: Yiddish, Khuzistani Arabic, Macedonian Turkish, and Indonesian. Limited borrowing in this domain is also typical of Manange and Hup, while Domari, Mosetén, Jaminjung, Guaraní, and Purepecha show the most extensive use of borrowed discourse markers. It is noteworthy that Tasawaq has its fillers and discourse markers from Hausa, a contemporary "pragmatically dominant" language (cf. Matras 1998). On the whole, the following hierarchy (from Matras 1998) of borrowability, both frequency-based and at least in most cases implicational, could be upheld:

(18) discourse markers > other particles

Question and answer particles must also be considered in this connection. The former are not a universal phenomenon anyway, and it is not surprising that they are limited, in our case, to Macedonian Turkish, which borrows its question particle from Albanian, interestingly replacing a native Turkish question particle. The borrowing of the positive answer particle 'yes' is more common. It is often employed as a turn-taking particle rather than just as a signal of agreement with content, and it perhaps for this reason that it is more

vulnerable than its negative counterpart 'no'. Rumungro and Mosetén borrow 'yes' (from Hungarian and Spanish respectively), while both 'yes' and 'no' are borrowed from Tuareg in Tasawaq, from Arabic in Domari, and from Spanish in Guaraní. The small sub-sample suggests an implicational hierarchy for the borrowing of answer particles which agrees with that postulated for Romani dialects by Elšík and Matras (2006: 343):

(19) positive > negative

7.5. Grammatical vocabulary

A notable gap in the borrowing inventory surrounds expressions of place deixis, which in our sample appears fully resistant to borrowing. Borrowed expressions of place are largely limited to the place indefinites 'nowhere' and 'anywhere', in which the borrowed component may either be the entire expression or just the indefinite marker (Domari, Otomi, Rumungro). The Yiddish presentative *ot* from Slavic is used in pointing only, and although a deixis of sorts, it is arguably more a verbalized gesture and hence more closely related to discourse markers than a member of a deictic paradigm.

Borrowed time expressions encompass both indefinites ('always', 'never') and deixis ('now', 'then'). Here, both classes are subject to borrowing and their behaviour appears to be linked in the following implicational hierarchy:

(20) always > never > now, then

In line with the high frequency of borrowing around indefinites, 'always' is the most frequently borrowed, followed by 'never'. The relevance properties of indefinites – as operators that process hearer-sided presuppositions – explains their higher position on the hierarchy (at least, this is well in line with other hierarchies discussed so far). The deictic expressions 'now' and 'then' are usually linked, and therefore occupy a single position on the hierarchy; occasionally they are borrowed independently of one another: Tasawaq borrows just 'now' from Tuareg, while Rumungro borrows only 'then' from Hungarian. Note that the Romani sample (Matras 2002) shows a clear hierarchy which favours the borrowing of 'then', beginning most often in its sequential rather than remote-deictic function, while 'now' is rarely borrowed.

Terms for days of the week are generally borrowed from the language of education, or the formal-official language: We find borrowings in Nahuatl, Jaminjung, Indonesian (ultimately from Arabic), Quichua, Otomi Guaraní, Biak, Purepecha, Hup (from Portuguese), Rumungro, K'abeena, Kildin Saami, and Western Neo-Aramaic. Borrowing of expressions for times of day ('morning', 'noon', etc.) is usually linked to days of the week, the sample showing an implicational hierarchy:

(21) days of week > times of day

This hierarchy can be nicely connected to the role of institutional administration and the language of commerce: Between the two categories, it is days of the week which tend to be the property of the public domain, governing the schedule of activities of individuals in relation to institutions, more so than times of the day, which have a greater autonomous role within the private domain. Once again we find that languages that do not have a strong tradition of relying on external languages for affairs of the public domain – Yiddish, Macedonian Turkish, Khuzistani Arabic – show no borrowings in these fields. Manange once again occupies a similar position, showing here too resistance toward borrowing and behaving much like a language with a tradition of literacy and institutional discourse.

For adjective comparison, our sample clearly confirms the hierarchy postulated by Elšík and Matras (2006) for Romani dialects (omitting the least borrowable value "positive", which is the default form of the adjective and is not usually accompanied by any overt derivational marker):

(22) superlative > comparative

Formal means of constructing both the superlative and the comparative are borrowed in Domari, Otomi, and Vietnamese, as well as Western Neo-Aramaic. Languages in which borrowing is limited to the superlative include Indonesian, with a Javanese particle *paling*, Rumungro, which borrows the Hungarian superlative prefix *leg,* and Kildin Saami, which has the Russian the superlative marker *same*. Yiddish, also confirming the hierarchy, shows an interesting case of hierarchical distribution of matter- and pattern-replication: While the comparative shows pattern-replication – *greyser fun mir* 'bigger *from* me' – the superlative shows matter borrowing, replicating the Russian marker *same*: *same groys* 'biggest'.

8. Constituent order and syntax

Contact-induced change in word order is generally not common in our sample. Evidence for change in basic word order is found in Rapanui, with a tendency toward change in Quichua and Otomi and sometimes in Khuzistani Arabic (here we are dealing mainly with relaxation of discourse constraints). It is plausible that in their earlier history Hup, Rumungro, and Domari showed different word-order patterns, though concrete evidence is lacking. The most common change in word order appears to affect possessive constructions. Examples are Domari, Macedonian Turkish, Rapanui, Quichua, Rumungro, and Likpe (see above, under Nominal structures). This is understandable, given the fact that a change in the position of possessor and possessed does not affect the position of the verb and so it leaves the organization of the predication intact. The position of adjectives is affected by contact in Domari, sometimes in Quichua, and in urban Manange.

Relative clauses change their position relative to the head in Macedonian Turkish and in Nahuatl, as well as, arguably, in Domari, if we compare the language with attested Indo-Aryan languages. The position of the copula appears more vulnerable to change than the position of lexical verbs, as seen in Macedonian Turkish, Rumungro (tendency only), and North-eastern Neo-Aramaic. From this, we might propose the following tentative hierarchy, based partly on frequency and partly on pure prediction, for the likelihood of word order to be affected by contact-related change:

(23) nominal constituents (possessor, adjective) > copula predications > verbal predications

Note that the hierarchy is sensitive to the presence of a lexical verb as initiator of the predication – a factor which impedes borrowing. Thus the most borrowable are structures that do not involve a full predication, or at least not a verbal one; these are followed by non-lexical predications, while predications that contain full lexical verbs appear last.

In the area of clause structure, one of the frequent changes observed is the emergence of copula clauses: Otomi borrows the Spanish copula *ta*, Hup uses a possible Tukano loan as a copula, and Indonesian creates copulas on a Sanskrit model. This makes sense, if one considers that non-universality of (present-tense) copula predications, and the fact that a clash of systems, and so pressure toward convergence, is more likely to occur here than in other clause types.

As far as syntax-relevant grammatical vocabulary is concerned, negation particles are borrowed in Domari, Quichua, Guaraní, Biak, Hup. Connectors and conjunctions were dealt with above (under Other Parts of Speech), and it was seen that most languages in the sample show some kind of contact influence on clause combining strategies. This is true especially of coordination, where almost all languages are affected, some by mere borrowing of connectors, others by changing connectivity strategy, as in the case of Rapanui (from serialization to connectors), or Macedonian Turkish (from converbs to connectors introducing finite clauses). A new type of adverbial subordination is attested only in Macedonian Turkish, while relative clauses are re-structured in Macedonian Turkish, Domari, Yiddish (to some extent), and Tasawaq. Like connectors, relative particles appear to be borrowing-prone, and we find loans in Nahuatl, Rapanui, Domari, Otomi, Guaraní, Indonesian, and Kildin Saami.

On the whole, then, the sample languages do not offer an extreme wealth of data on contact-induced change in clause or sentence structure. In particular, colonial languages and languages of administration do not seem to have the effect on syntactic structures as they may have on other domains of structure, in particular grammatical vocabulary. Those cases where we do find far-reaching changes in syntactic typology tend to be languages in a prolonged situation of stable multilingualism, as in the case of Macedonian Turkish or Domari, confirming Thomason and Kaufman's (1988) prediction on the link between prolonged and intense cultural contact, and significant typological disruption.

9. Lexicon

All languages borrow lexical elements. Based on authors' responses to the questionnaire, we can make a statement about the likelihood of a certain word-class element to be affected by contact. This hierarchy is not implicational, as there is no evidence to suggest that borrowing in a lower-ranking category necessarily entails borrowing in higher-ranking categories. Rather, it is based on frequency; (24) shows how frequently selected word classes occur among the list of word classes affected by contact in the sample:

(24) nouns, conjunctions > verbs > discourse markers > adjectives > interjections > adverbs > other particles, adpositions > numerals > pronouns > derivational affixes > inflectional affixes

In some salient features, this hierarchy resembles other hierarchies proposed in the literature.[14] Thus, nouns appear at the top of the list, unbound grammatical vocabulary is rather high on the list, and bound morphology is low, with derivational morphology outranking inflectional morphology. Note however some differences to the hierarchy proposed by Muysken (1981) and others: Conjunctions and discourse markers occupy a high position, and outrank some of the lexical categories. Numerals outrank pronouns and derivational morphology in this particular sample, much of it composed of languages in contact with a colonial language. No differentiation is made between coordinating and subordinating conjunctions in our discussion, but it is clear that the position even of subordinating conjunctions is far higher on the hierarchy that that assigned to it in Muysken's data.

We must, however, treat the meaningfulness of such a full hierarchy with caution. We have seen evidence that the presence of numerals on the list can be biased by the type of contact situations that are selected, and the presence or absence of a tradition of literacy in the recipient languages. The presence of pronouns on the list is largely a product of the function of terms of address and their degree of lexicalization, or alternatively of the clash of systems that distinguish exclusive/inclusive reference, and those that don't. These and more are coincidental circumstances that influence the contact situation around particular categories, and which will promote or demote certain categories on the hierarchy, depending on the degree of presence of languages in the sample that answer to certain sociolinguistic and structural criteria. Thus, the more reliable hierarchies are those that provide a picture of the susceptibility of category values to borrowing, while the comparison among categories is not entirely free of arbitrary factors.

10. Concluding remarks

Noteworthy is the extent of borrowing across the different languages of the sample. If we take as an indicator 36 prominent categories representing various aspects of structure – from phonology, through to morphology, unbound grammatical vocabulary, lexicon, and syntax[15] – and assign scores to languages based on the number of categories that show some kind of contact influence, then the scores range from 31 (Rumungro, Guaraní), to just 6 (Yaqui) and 7 (Biak) (see Table 1).

Table 1 shows the overall "borrowing score" for each language. It also shows the borrowing score among 11 categories representing Other Parts of

Speech (OPS),[16] an indicator of mostly unbound grammatical lexemes, and the proportion of OPS among the categories affected by contact. The majority of languages range within a 0.05 distance from the proportional 0.305 share of OPS among the category total, indicating that there is, on the whole, a rather predictable share of OPS among borrowed categories. Note however that some languages show a disproportionately high level of borrowing of OPS: Yaqui, Biak, Likpe, Mosetén, Jaminjung score between 0.4–0.5, meaning that OPS account for up to 50 percent of borrowed categories. Note that all these languages tend to have overall borrowing scores of between 6–15, that is, on the lower side of the borrowing range. This might be interpreted as indirect

Table 1. So-called "borrowing scores", and the proportion of Other Parts of Speech among categories affected by contact.

Language	Total score	OPS score	OPS/Total ratio
Sampled	36	11	0.305
Yaqui	6	3	0.500
Biak	7	3	0.428
Manange	10	3	0.333
K'abeena	10	3	0.333
Likpe	10	4	0.400
Mosetén	11	5	0.455
Mac. Turkish	12	3	0.250
Rapanui	13	3	0.230
Khuz. Arabic	13	4	0.307
Kildin Saami	15	5	0.333
Jaminjung	15	7	0.467
Vietnamese	17	6	0.352
Nahuatl	18	5	0.278
Tasawaq	18	7	0.389
Purepecha	19	7	0.368
Western Neo-Aramaic	20	6	0.300
Hup	21	5	0.238
Kurmanji Kurdish	21	6	0.286
Yiddish	22	7	0.318
Domari	24	9	0.375
Otomi	25	9	0.360
Quichua	26	8	0.307
Indonesian	26	9	0.346
Guaraní	31	10	0.323
Rumungro	31	10	0.323

evidence that borrowing begins with OPS, before continuing to other categories. The lowest OPS scores, for Rapanui, Hup, and Macedonian Turkish (0.23–0.25), represent languages that undergo structural changes as a result of contact but in which borrowed OPS are under-represented. These figures generally confirm the predictions and observations that unbound grammatical morphemes are high on the borrowability scale compared to other categories; though they also allow us to conclude that contact influence is rarely limited to them, and that it is not impossible for a language to display even a certain amount of resistance toward borrowing of OPS.

It is also interesting to note that a number of categories occupy an entirely peripheral position in the borrowing behaviour of languages in this sample. They include bound case markers, bound tense markers, bound person markers as well as in most cases unbound person markers (deixis and anaphora; exceptions being reciprocal and reflexive pronouns, and "lexicalized" pronouns as in Indonesian), demonstratives and expressions of place deixis. In the following discussion it will hopefully become clear that the absence of borrowing in these domains is not taken to mean that constraints exclude them from being borrowed. Rather, our focus is on those categories that do show a more salient and frequent tendency to be affected by borrowing, and our agenda is to explain why speakers are motivated to borrow forms and structures in those categories. The absence of borrowing among other categories may be left to be interpreted as just that: the absence of any particular motivation to converge the two systems around these particular categories.

Typological features are neither excluded nor even rarely affected by borrowing in the sample. A number of languages undergo considerable typological convergence: Macedonian Turkish, Domari, North-eastern Neo-Aramaic, and Rumungro. Ongoing shifts in morphological typology can be detected in a number of other languages, too. Although statistically, unbound grammatical morphemes are more likely to be borrowed than typological features, there is no direct interdependency between any specific value or category that falls within these respective groups of structures. Implicational hierarchies of the kind postulated above only apply among the values of the same category. But since the likelihood of borrowing is different for different categories, there may be a quasi-implicational relationships across categories in different structural domains. Thus, since connectors are frequently borrowed, but re-structuring within the TMA domain is rare, we might expect a language that shows contact-induced re-structuring in the domain of TMA to show borrowed connectors as well. Such an expectation is based on the higher borrowing frequency of connectors.

Despite the lack of any direct functional link between the borrowing of connectors and the restructuring of TMA categories (or any other example of contact influence), the challenging question remains, *why* certain categories are more susceptible to change in situations of language contact, than others. It is here that borrowing hierarchies, especially the implicational relations among paradigm values of the same structural category, can shed some light. The fact that borrowing within such categories often follows a non-arbitrary, predictable course, suggests that semantic-pragmatic features that distinguish among category values participate in motivating borrowing. The relations among borrowed values can thus help us illuminate the motivation behind borrowing, and so help us make sense of the different degrees of susceptibility of categories to the borrowing process.

Let us, for this purpose, review the hierarchies. A first set of hierarchies might be grouped together based on a general notion of frequency, referential meaning, and usage context of the borrowed structural material. This group is rather diverse. The more frequent adoption of new consonants over new vowels (1) is conditioned by the mere diversity of consonants and the fact that they tend to outnumber vowels in each of the sample languages. The greater likelihood that phonemes be adopted as part of loanwords than as independent phonological features (3), is similarly a practical issue relating to the need to accommodate loanwords.

Borrowing as a utilitarian enrichment of means of expression also belongs here. The greater likelihood of borrowing of nouns over other parts of speech (as expressed in 24) is a product of the likelihood of nouns to express new concepts and to name objects and institutions (see already Weinreich 1953: 37). The high borrowability of lower ordinals (12) is connected to the fact that they mark exclusivity by assigning lower figures a special lexical item rather than a grammatical derivational procedure, with borrowing supplementing the procedure of lexical creativity. Certain usage contexts may favour borrowing, if there is a close association with the contact language in certain domains. Thus, borrowed numerals are more likely to be used in formal than informal contexts (9), higher numerals as well as mathematical "zero" are more likely candidates for borrowing, being reserved primarily to more formal-institutional contexts (10), and names of days of the week are more likely to be borrowed than times of the day (21).

An additional theme, which groups together another bundle of hierarchies, may be defined as accessibility, cognitive complexity, and expectedness. Low accessibility and/or high complexity correlate with the borrowing susceptibility of peripheral as opposed to core local relations (4), of higher

numerals (10) (reinforcing the usage-based motivation cited above), of independent (factual) embedded events over dependent (non-factual) ones (15), and of linked independent events (purpose clauses, causal clauses) over linked dependent ones (adverbial subordinations) (14). Low expectedness can be brought in connection with some of those, and in particular with the contact-susceptible properties of contrast (13), concessive subordination (14), phasal change (16), restrictive focus (17), and superlative (22). In all of these cases, the speaker's assertive authority is potentially reduced as a result of the speaker venturing into propositional domains involving a degree of uncertainty or unexpectedness.

Not unrelated are the properties around which external circumstances reduce the speaker's confidence and control even more overtly. These include the high borrowing susceptibility of conditionality over other subordinations (14), of participant-external modality over participant-internal modals (5), of modality itself over aspect and tense, as well as of aspect/aktionsart (the internal structure of the event, independent of the speaker's perspective) over tense, and of the future over other tenses (6), as well as of indefinites (which rely on a presuppositional domain) over deictics (which rely on the speaker's own orientation perspective)[17] (20).

Finally, we find a set of hierarchies that operate at the level of the interaction, where those structures are more borrowing-prone that are more tightly connected to the emotive level of the discourse or speech act rather than to the content level of the word or phrase. Such is the case with prosodic features over segmental phonological features (2), discourse markers over other particles (18), connectors over other parts of speech (as in 24), causal argumentation over other forms of subordination (14), and even the positive answer particle over the general (i.e. also content-bound) negative particle (19).

What do these three themes – accessibility/expectedness, external dependency, and interaction-level operations – have in common, and why are they especially susceptible to contact-related change? In order to answer this question, we must return to our hypothesis about what constitutes "borrowing" in the first place (cf. Section 1). Borrowing, we had said, is a strategic compromise which bilinguals adopted in conversation and which has become socially acceptable. Social acceptability is a pre-condition for change, since language is the collective, socio-cultural product and asset of a community. But there is no reason to assume that social attitudes should in any way prejudice contrastive connectors over additive connectors, or temporal indefinites over time deixis. The reason for the hierarchical arrangement of

categories in respect of their borrowing susceptibility has to do with the first part of our definition of borrowing, namely the part that describes borrowing as a strategic compromise adopted in conversation. It is here that speakers are naturally inclined to prioritize when handling the control mechanism that selects certain ("language-particular") structures in certain sets of interactions.

Maintaining the demarcation boundary between repertoire components (or "languages") is a burden on the mental processing of language in conversation, and yet it is a social requirement. Compromise is sought when the tension assumes its most extreme forms: when the burden of controlling the language selection mechanism coincides with other sources of tension in the interaction itself. Such tension emerges when the speaker's assertive authority is at stake and a special effort is needed in order to win over the hearer's confidence: When expressing unexpected chains of arguments, when contradicting or challenging presuppositions, when assuming responsibility for propositional content that lies beyond the domain of secure knowledge, or when directly intervening with hearer-sided processing by monitoring and directing turns and speech acts (e.g. through prosody or discourse markers). Since the conversational tension around such processing tasks cannot itself be reduced, bilingual speakers' only alternative is to eliminate the need to distinguish between sub-components of their linguistic repertoire – or "languages" – and to unify the structures that trigger the appropriate processing operations. The result is a fusion of the two systems of structures around the relevant functions (see Matras 1998).

The trigger for borrowing around these kinds of structures – those that cluster around the themes defined above as "accessibility", "expectedness", "interaction-level", and so on – is thus neither social acceptability, nor prestige, nor gaps in the recipient language. Rather, it is the need to reduce the cognitive load when handling a complex linguistic repertoire. Social acceptability is merely an accompanying condition for spontaneous innovations to become anchored in the long-term speech behaviour of the community.

In this respect, the susceptibility of a great number of grammatical categories to borrowing is pre-determined by their language processing function, and therefore universal. One of the most striking findings of the present investigation is the fact that so many hierarchies that were identified for the cross-linguistic sample presented in this volume, were a perfect or near-perfect match to those identified by Matras (2002) and by Elšík and Matras (2006) for the sample of Romani dialects in contact with a variety of different languages, and, to the extent that material for comparison was available,

also with hierarchies proposed for other samples. This clearly supports their universal predictive power. Moreover, the fact that a sample consisting of multiple recipient languages shows virtually the same results as a sample with a controlled recipient language (Romani) suggests that the structure of the recipient language plays only a secondary, perhaps even just a peripheral role in determining preferences of borrowing. The primary role is played by the functionality of the categories and the extent of bilingual pressure, i.e. the extent to which bilinguals need to make frequent decisions on language choice.

Note that this sharpens the focus of what Thomason and Kaufman (1988) had referred to somewhat more bluntly as the intensity of cultural contact, helping us to move toward a more specific characterization of relevant patterns of communicative interaction. To be sure, the structures of the languages involved, especially the recipient language, may play a certain role in the borrowing process. But this role must be seen primarily as an imposition of constraints on what is essentially a universal process, motivated by cognitive features of language processing. Such constraints might include the presence of a competing structure on one side of the paradigm (as with the Jaminjung contrastive marker); or the availability of literacy as a factor strengthening the coherence of the recipient system and thus reinforcing demarcation boundaries and helping to resist borrowing; or indeed the presence of social attitudes that block language mixing. On the other hand, the fact that languages like Malay employ a plethora of lexical means, and not just deictic and anaphoric expressions, to refer to participants creates a motivation for renewal of this inventory of expressions and so also for borrowing in the domain of (so-called) personal pronouns, which is not typically found in languages that rely on participant deixis and anaphora.

The principal conclusion that must be drawn from the above observations is that different borrowing motivations apply to different functional categories. With some, the motivation is lexical enrichment. With others, it is the fusion of elements of formal discourse with the language that dominates formal discourse, while in a series of categories the motivation is a reduction in the tension surrounding certain language processing tasks. Though neither gaps nor social prestige are primary motivators for borrowing, both are indirectly involved, as the process of "borrowing" can be defined as a license to speakers to dismantle the mental demarcation boundaries that separate their individual "languages" and, around a particular selection of categories, to make full use of their entire repertoire of linguistic structures and forms irrespective of the setting of the communicative interaction.

Acknowledgement

This chapter was written during a research visit to the Research Centre for Linguistic Typology at La Trobe University, Melbourne, made possible through a Distinguished Fellowship at the Institute for Advanced Study, La Trobe University, and an International Linkage Fellowship of the Australian Research Council.

Notes

1. I use the term "language" in quotes in this context since it is not obvious that multilingual speakers process language in the form of separate systems; it is safer to assume that multilingual speakers have an overall repertoire of linguistic forms, to which constraints are attached concerning the situations and conversational constellations in which those forms may be used, for various purposes. The notion of a separation of "linguistic systems" on the part of the language user is therefore somewhat of an abstraction.
2. We assume that borrowing always begins with at least some degree of bilingualism, however rudimentary, or at least with an exposure to settings of communicative interaction that require the selection of a separate inventory of forms and structures. Once a certain behaviour pattern is adopted by those speakers who interact in a variety of settings – and so have access to two (or more) "languages" – new forms and structures may diffuse into the speech patterns of monolinguals as well, or may survive the historical decline of widespread bilingualism. Such latter process may strengthen our ability to identify borrowings, but it is not a pre-requisite for borrowing.
3. On the problem of establishing "borrowability" on the basis of corpus frequency, see already Weinreich's (1953: 35–36) critical remarks.
4. Since the occurrence in a corpus of a low-ranking category presupposes that of the higher-ranking category, occurrences of the higher-ranking category will always outnumber those of the lower one.
5. Throughout I use the notation "greater than" (>) to denote the value that is more likely to be affected by contact induced change (in a frequency-based hierarchy), and which in an implicational hierarchy constitutes a pre-requisite for the borrowing of any item specified to the right of it and marked "lesser than".
6. The one exception being Kriol, which has a phonological system that is very similar to Jaminjung.
7. In both Romani and Domari, genetically related material (deriving from Indo-Iranian postposed adverbial specifiers) undergoes a similar development toward agglutinative case markers (cf. Matras 2002).
8. The model is in fact areal, and is also shared by Persian and Western Armenian, and to some extent by Levantine Arabic as well.

9. Rapanui uses Spanish *tiene que* which seems to express both necessity and obligation; Imbabura Quichua has *pudi-* "can", which could well cover both ability and possibility.
10. We also have no evidence to uphold the (frequency) hierarchy proposed by Wichmann and Wohlgemuth (forthc.) (note that prominence of strategies is arranged from left to right): light verbs < indirect insertions < direct insertion < paradigm transfer. But we have no grounds on which to challenge this hierarchy, either.
11. See Bakker (1997), however, on constraints that prevent the isolation of the Algonkian verb to a bare stem, in the context of Cree/French contact (albeit in connection with the formation of the mixed language Michif, not with borrowing in the conventional sense).
12. Like all Romani dialects, Rumungro too uses Greek derivation markers for form ordinals from cardinals, but the hierarchy applies to the borrowing of ordinal word forms.
13. Wilco van den Heuvel, p.c.
14. Compare with integrated hierarchy presented by Muysken (1981), repeated by Winford (2003): 51: nouns > adjectives > verbs > prepositions > coordinating conjunctions > quantifiers > determiners > free pronouns > clitic pronouns > subordinating conjunctions
15. The full list is: Consonants, vowels, morphological typology, alignment type, local relations, classifiers/gender, possession, plurality, definiteness, diminution/augmentation, nominalization, case marking, tense categories, tense marking, aspect categories, aspect marking, aktionsart categories, aktionsart marking, mood categories, modal verbs, voice and valency, numerals, personal pronouns, demonstratives, indefinites, interrogatives, connectors, subordinating conjunctions, phasal adverbs, focus particles, discourse markers, time deixis, adjective comparison, constituent order, syntax, basic cultural vocabulary.
16. These are: Numerals, personal pronouns, demonstratives, indefinites, interrogatives, connectors, subordinating conjunctions, phasal adverbs, focus particles, discourse markers, time deixis, adjective comparison.
17. More precisely, indefinites can be said to engage the hearer more actively in supplementing an imaginary knowledge domain in which the missing context can be situated: consider an indefinite expression such as 'anywhere', where it is up to the hearer to construct an image of possible locations that satisfy vague contextual criteria. With deixis, on the other hand, the speaker is confident that speaker and hearer share a very particular perspective. Thus, 'here' leaves no room for ambiguity, or for hearer-sided creativity.

References

Aikhenvald, Alexandra Y. and R. M. W. Dixon (eds.)

2001 *Areal Diffusion and Genetic Inheritance*. Oxford: Oxford University Press.

2006 *Grammars in Contact. A Cross-linguistic Typology*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y.

2002 *Language Contact in Amazonia*. Oxford: Oxford University Press.

2006 Grammars in contact: A cross-linguistic perspective. In: Alexandra Y. Aikhenvald and R. M. W. Dixon (eds.), *Grammars in Contact. A Cross-Linguistic Typology*, 1–66. Oxford: Oxford University Press.

Bakker, Peter

1997 *A Language of Our Own: The Genesis of Michif, the Mixed Cree–French Language of the Canadian Métis*. Oxford: Oxford Univer sity Press.

Campbell, Lyle

1993 On proposed universals of grammatical borrowing. In: Henk Aertsen and Robert Jeffers (eds.), *Historical Linguistics 1989: Papers from the 9th International Conference on Historical Linguistics*, 91–109. Amsterdam: John Benjamins.

Dixon, R. M. W.

1995 Complement clauses and complementation strategies. In: F. R. Palmer (ed.), *Grammar and Meaning: Essays in Honour of Sir John Lyons*, 175–220. Cambridge: Cambridge University Press.

2006 Complement clause types and complementation strategies in typological perspective. In: R. M. W. Dixon and Alexandra Aikhenvald (eds.) *Complementation. A Cross-Linguistic Typology*, 1–47. Oxford: Oxford University Press.

Elšík, Viktor, and Yaron Matras

2006 *Markedness and Language Change: The Romani Sample*. Berlin: Mouton de Gruyter.

Field, Fredric

2002 *Linguistic Borrowing in Bilingual Contexts*. Amsterdam: Benjamins.

Givón, T.

1990 *Syntax. A Functional-Typological Introduction*. Amsterdam: Benjamins.

Haugen, Einar

1950 The analysis of linguistic borrowing. In *Language* 26 (2): 210–231.

Heath, Jeffrey

1984 Language contact and language change. In *Annual Review of Anthropology* 13: 367–384.

Matras, Yaron
1998 Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36 (2): 281–331.
2002 *Romani: A Linguistic Introduction*. Cambridge: Cambridge University Press.
2004 Layers of convergent syntax in Macedonian Turkish. *Mediterranean Language Review* 15: 63–86.
Matras, Yaron and Jeanette Sakel
2007 Investigating the mechanisms of pattern-replication in language convergence. *Studies in Language* 31 (4): 829–865.
Moravcsik, Edith
1975 Verb borrowing. *Wiener Linguistische Gazette* 8: 3–30.
1978 Language contact. In: Joseph H. Greenberg, Charles A. Ferguson and Edith A. Moravscik (eds.), *Universals of Human Language*, Vol. 1: 93–122. Stanford: Stanford University Press.
Muysken, Pieter
1981 Halfway between Quechua and Spanish: The case for relexification. In: Arnold Highfield and Albert Valdman (eds.), *Historicity and Variation in Creole Studies*, 52–78. Ann Arbor: Karoma.
Ross, Malcolm
2001 Contact-induced change in Oceanic languages in north-west Melanesia. In: Alexandra Aikhenvald and R. M. W. Dixon (eds.), *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, 134–166. Oxford: Oxford University Press
Stolz, Thomas
1996 Grammatical Hispanisms in Amerindian and Austronesian languages. The other kind of Transpacific isoglosses. *Amerindia* 21: 137–160.
2007 Allora. On the recurrence of function-word borrowing in contact situations with Italian as donor language. In: Jochen Rehbein, Christiane Hohenstein and Luaks Pietsch (eds.), *Connectivity in grammar and discourse*, 75–99. Amsterdam: Benjamins.
Stolz, Christel, and Thomas Stolz
1996 Funktionswortentlehnung in Mesoamerika, Spanisch-Amerindischer Sprachkontakt. In *Sprachtypologie und Universalienforschung* 49 (1): 86–123.
1997 Universelle Hispanismen? Von Manila über Lima bis Mexiko und zurück: Muster bei der Entlehnung spanischer Funktionswörter in die indigenen Sprachen Amerikas und Austronesiens. *Orbis* 39 (1): 1–77.
Thomason, Sarah G.
2001 *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.

Thomason, Sarah G., and Terrence Kaufman
1988 *Language Contact, Creolization and Genetic Linguistics.* Berkeley: University of California Press
van der Auwera, Johan
1998 Phasal adverbials in the languages of Europe. In: Johan van der Auwera with Dónall P. Ó Baoill (ed.), *Adverbial Constructions in the Languages of Europe*, 25–145. Berlin: Mouton de Gruyter.
van Hout, Roeland, and Pieter Muysken
1994 Modelling lexical borrowability. *Language Variation and Change* 6: 39–62.
Weinreich, Uriel
1953 [1968] *Languages in contact.* The Hague: Mouton.
Wichmann, Søren, and Jan Wohlgemuth
Forthc. Loan verbs in a typological perspective. In: Thomas Stolz, Dik Bakker and Rosa Salas Palomo (eds.), *Aspects of Language Contact.* Berlin/New York: Mouton de Gruyter.
Winford, Donald
2003 *An Introduction to Contact Linguistics.* Oxford: Blackwell.

Grammatical Borrowing in Cross-Linguistic Perspective

Edited by
Yaron Matras
Jeanette Sakel

Mouton de Gruyter
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

Library of Congress Cataloging-in-Publication Data

Grammatical borrowing in cross-linguistic perspective / edited by Yaron Matras, Jeanette Sakel.
p. cm. – (Empirical approaches to language typology ; 38)
Includes bibliographical references and index.
ISBN 978-3-11-019628-3 (cloth : alk. paper)
1. Language and languages – Foreign elements. 2. Grammar, Comparative and general. I. Matras, Yaron, 1963– II. Sakel, Jeanette, 1973–
P324.G73 2007
410–dc22
2007042917

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at http://dnb.d-nb.de.

ISBN 978-3-11-019628-3
ISSN 0933-761X