The full extent of 'Fusion':
A test case for connectivity and language contact

YARON MATRAS

1. Introduction: What is Fusion?

In earlier work I have argued for the susceptibility of connectivity devices to contact-induced change – borrowing of connectors, as well as re-arrangement of connectivity strategies (MATRAS 1996, 1998). My impressions, based on the case studies I examined, resembled some of the conclusions put forth in other studies which suggested that language contact progresses from the level of discourse, through the level of sentence grammar, to the clause and word level (STOLZ & STOLZ 1996), or from 'larger' to 'smaller' structural units (ROSS 2001). They are also in line with numerous observations on the contact-susceptibility of connectors and other sentence-peripheral, unbound and indeclinable function words, especially discourse markers (e.g. SALMONS 1990, DE ROOIJ 1996).

I have previously proposed the term 'Fusion' to capture the wholesale adoption of a category and its structural markers from a contact language, or the wholesale replacement of an inherited category through borrowings (MATRAS 1998, 2000). I argued for a communicative and cognitive motivation for Fusion: In communicative terms, Fusion is triggered in the first instance by the adoption of communicative patterns that are automatic and which resemble verbalised gestures, and whose function is to direct the hearer through the interaction (and specifically through the speaker's current turn) (cf. also the notion of 'convergence' in Salmons 1990). Fusion affecting long-term structural change is

therefore a result of the wholesale accommodation on the part of the speaker of Language A, to the structures that allow this speaker to exercise most effective control over hearer-sided processing operations while communicating in language B. It is for this reason that discourse markers and connectors occupy the top position on the hierarchical list of categories that are prone to Fusion. Fusion is directed toward a target language that is 'pragmatically dominant', meaning that the speaker's maximal effort and control are concentrated on producing adequate speech performance in this language. However, Fusion is unlikely to have a long-lasting effect unless the target language has social prestige. The sociolinguistic settings in which Fusion will lead to language change are therefore those in which B is the language of prestige and power, and effective accommodation to discourse-regulating structures in B is, on a regular basis, a key to successful negotiation of vital transactions. Since B is a prestige language, speakers of A are likely to react with tolerance, in the first instance, toward the occasional import of B-language structures into A-language, and might ultimately replace the entire category in A-language through counterpart structures from B-language.

Alongside the communicative and sociolinguistic sides, there is a cognitive dimension to Fusion. In cognitive terms, Fusion can be regarded as the non-separation of the two languages in a bilingual's repertoire for a particular class of functions or category. The task of monitoring and directing hearer participation while at the same time continuing to structure the utterance intensifies the pressure – or 'processing load' – on the speaker around structures that are employed for such monitoring-and-directing operations. The need to keep apart two linguistic systems adds to this load. Now, monitoring-and-directing operations appear to be managed separately from other linguistic operations – tending to function as automatic, situative gestures. At the point of insertion of a connector or discourse marker by a bilingual speaker two things can therefore be assumed to happen: a) the speaker directs the hearer and monitors the hearer's reaction while at the same time continuing to plan the utterance, which places considerable processing strain, and b) the speaker switches into the monitoring-and-directing mechanism. Fusion begins at the point at which the switch into the

monitoring-and-directing mechanism is wrongly interpreted by the processing apparatus as accompanied by a switch into the second language, triggering a bilingual speech production error. I have shown in previous work (MATRAS 1998, 2000) that Fusion can occur at the local level, at points at which switching is not to the speaker's communicative advantage, where it serves no conversational function, and where it entails loss of face rather than allow the speaker to gain prestige. More so than emblematic or discourse-functional switches (cf. POPLACK 1980, MASCHLER 1994), bilingual errors around connectors and discourse markers can explain the beginnings of long-term wholesale import of a category: Where they do not act against the speaker's prestige, such slips are tolerated. Ultimately a link is established between a switch into the monitoring-and-directing mechanism and a switch into B-language, and this becomes the norm. The two systems are then no longer separated for monitoring-and-directing operations.

Further evidence for this interpretation of Fusion comes from the cline of borrowability. This cline can be applied within a single morphosyntactic category, such as co-ordinating and subordinating conjunctions, or focus particles. Contrast (*but*, *however*, *although*, *except*), for instance, is more prone to Fusion than addition (*and*, *as well*). In other words, the stronger the clash with hearer expectations, the more intense the speaker's attempt to gain control over hearer-sided processing, the more likely it is for the speaker's processing effort to become channelled and for control over other mechanisms to become lax, triggering more frequently an 'erroneous', unintentional link between the switch into the monitoring-and-directing mechanism and the switch into B-language. Moreover, the cline can apply across categories: Thus connectors and discourse markers are more prone to borrowing than other function words. Finally, the cline may apply metaphorically, by associating the semantic representations of 'external forces' or even 'event independence' with the gesture of 'reaching out' to the hearer. Thus modal verbs expressing external pressures and obligations ('must') are more prone to borrowing than those expressing (internal) volition ('want'), and factual complementisers (such as Bulgarian *či* or Greek *oti*), are

more likely to be borrowed than non-factual ones (Bulgarian *da*, Greek *na*) (cf. MATRAS 2002, Ch. 8).

The purpose of this contribution is to examine the extent of Fusion is language with so-called 'heavy' borrowing. The term 'heavy' – used by THOMASON & KAUFMAN (1988) – is to a considerable degree impressionistic, and there is no agreed (and even few proposed) objective methodology of measuring and quantifying the extent of borrowing (aside from the representation of basic vocabulary as measured by the Swadesh lists). 'Heavy' can be taken to imply that a language has borrowed not just a large number of lexical items, but also a large number of grammatical structures, and from a comparatively large and diverse range of grammatical categories. Moreover, 'heavy' might be taken to imply that borrowing has taken place in grammatical categories that are normally, that is in cases of less 'heavy' contact, not prone to borrowing. In this latter sense, any evaluative judgement about the extent of borrowing takes for granted some general appreciation, at some level, of a cline or hierarchy of borrowing. If a language shows 'heavy' borrowing and is therefore predicted to show borrowing also at the level of lower-ranking categories (on the borrowing cline), then we might expect this language to exhibit even more consistent borrowing at the higher-ranking levels of the cline. If we follow the model suggested above for Fusion, we would therefore expect particularly extensive borrowing at the level of clause combining and connectivity, as well as the levels associated with connectivity functions, directly or metaphorically (see discussion above).

2. Domari as a language with heavy borrowing

Domari is the neo-Indic (New Indo-Aryan) language of the Near Eastern *Dom*, who are referred to by the majority populations of the region by a variety of names, including *Kauli* (Iran), *Qurbati* (Syria and Iraq), and *Nawar* (Jordan and Palestine). The Dom were traditionally peripatetic societies, engaged in itinerant service-providing trades such as metalwork and entertainment. They are part of the more general phenomenon of Indian service-providing castes who have

sought livelihood opportunities outside of India but have preserved their social structure, to some extent customs and beliefs, and language. The best known and best documented of those groups are the Rom (or Romanies, or Gypsies) of Europe, and indeed an issue of continuous debate in Romani linguistics has been the relationship between the Romani language, and Domari (see Hancock 1988, 1995; Matras 1999, 2002. Domari first became known to the scholarly community in the form of wordlists published in the nineteenth century (POTT 1846). A first grammatical sketch, based on the dialect of Jerusalem, was published by MACALISTER (1914). No empirical work documenting Domari has followed however until my recent descriptive and documentary publications on the language (MATRAS 1999), also based on the Jerusalem dialect. The Jerusalem Dom community is small, and the language ceased to be transmitted to the younger generation in the 1960s. The younger generations have consequently switched to Arabic, and Domari is restricted to everyday, oral conversation in the context of the family, and used actively by no more than a few dozen people in the community, most of them born before 1940. Domari is therefore an endangered language.

Domari is also a language in contact. The Dom have always been bi- or multilingual as a result of the structure of their economy, which makes interaction and transactions with the majority society an absolute necessity. There are no formal codes on the use of the Domari language, and no community-internal institutions to safeguard it. In the past, Domari served a function as the symbol of community identity and so it was retained and passed on as long as the community economy remained intact. The abandonment of the language is connected to the gradual transition from specialised work to paid and regular labour, and from accommodation in tents at the outskirts of towns to rented accommodation in the midst of non-Dom population. Lack of formalisation (as a written language or ritual language), bilingualism, limited domains of use and limited functional scope have thus allowed Domari to absorb foreign influences, in particular from Arabic. Yet Domari – as still spoken by the elders of the Jerusalem community – remains an inflected, clearly-defined, separate language.

It is not to be confused with the occasional insertion of foreign (sometimes, indeed, Indic-derived) vocabulary into the speech of peripatetics in the region to form a so-called 'secret language' or special lexicon. Despite the massive amount of Arabic borrowings in lexicon and grammar, Domari maintains its core vocabulary and inflectional patterns of Indic origin. It is, in fact, one of the most conservative modern Indo-Aryan languages, retaining for instance to a considerable extent the Middle Indo-Aryan present-tense conjugation and the Middle Indo-Aryan subjunctive and synthetic passive.

The Indic features of Domari include a core lexicon (around 50% of the Swadesh-list is an indicator), and most inflectional grammar. The latter includes the system of person concord (subject and object agreement) on verbs in all tenses, and person markers on nouns and location expressions (possessors), the case endings (a complex and layered system expressing altogether up to six non-nominative nominal cases), bound markers on the lexical verb (expressing passive, causative, subjunctive, and tenses-aspect categories), plural and derivational markers on the noun, gender and number agreement markers on adjectives, participles, converbs and their distribution patterns, all personal pronouns and demonstratives, the enclitic copula and modal verb 'to be able to', a bound indefinite marker, lower numerals (below 5, as well as 10 and 100), most interrogatives ('what', 'who', 'where', 'how', 'when'), converb markers, the negator of lexical verbs, and phonemic distinctions that have no counterpart in Arabic (among them the opposition *p:b*). It is clear then that Domari is not a 'Mixed Language' in the sense argued for by BAKKER (1997; cf. also BAKKER & MOUS 1994), MUYSKEN (1997), THOMASON (1995, 1997), or others: It does not show a split between the source languages of the grammar and the lexicon, and it does not defy genetic classification, for most of its structures can be traced back without considerable difficulty to an Indic ancestor.

What then amounts to the 'heavy' or 'massive' Arabic borrowing in Domari? Here is a bird's-eye view of Arabic-derived structures: Much of the lexicon, in all likelihood even the majority of the lexical types (though not necessarily tokens) used in any Domari conversation, comes from Arabic; this includes 50% of the

Swadesh list entries of assumed 'core vocabulary'. Arabic items include numerals above 5 (excluding 10 and 100). Arabic nouns are incorporated with their plural counterparts (although Indic plural endings may be added on top of those), and they are only partly adopted phonologically, and tend to preserve their original phonemes. Moreover, Arabic phonemes, such as pharyngeals, are adopted and diffused into the pre-Arabic lexical component of Domari too. Arabic provides a series of modal verbs and auxiliaries, including 'want', 'must', 'begin', 'stop', 'continue' (and even 'remain') and the aspectual auxiliary for the habitual-frequentative; all these carry Arabic-derived person and tense-aspect inflection and Arabic negators. The entire inventory of unbound prepositions is Arabic-derived, as are the comparative and superlative forms of adjectives (including the lexical form: thus Indic-derived *tilla* 'big', Arabic-derived *akbar* 'bigger'). All conjunctions, co-ordinating and subordinating, are Arabic, as are focus particles, discourse markers, most indefinite expressions, and most non-deictic adverbs, including phasal adverbs ('still', 'no more'). Arabic-derived are also the complementiser introducing complement clauses, along with its agreement inflection with the subject of the complement clause, the relativiser, and the resumptive pronoun in relative clauses, along with its Arabic agreement inflection with the head noun. The two languages share many syntactic-typological features: Word order in the verb phrase is virtually identical, with the exception of the presence in Domari of an enclitic copula (which however is often complemented by the Arabic copula functioning as an aspectual auxiliary; see below). Domari has inherited adjective-noun order, while Arabic has noun-adjective order; but Domari is restructuring its attributive word order, and speakers use attributives with predicational markers in postnominal position (instead of *qašt\ot\e zare* 'small children', or *gulda xatm-ak* 'a pretty ring', the preferred structures are *zarēnī qašt\ot\ēnī* 'chidren, small ones', *xatm-ak guld-ēk* 'a ring, being pretty', cf. Arabic *awlād z̧Vār, xatm h\ilu*).

3. The case for Fusion in Domari

From the above survey, however brief, of borrowed categories in Domari it is already apparent that the language shares its entire connectivity structure with Arabic. The following examples provide an illustration of this state of affairs (all examples from tape-recorded narratives in Domari, recorded by the author between 1996-2000 in Jerusalem; Arabic-derived items are italicised):

(1)   a.   *lamma kunt*   ama qašṭōṭik,   na   nērdedim   *madāris*anka.

                  when was.1SG   I small.PRED.F.SG   NEG sent.3PL.1SG   schools.PL.DAT

         b.   *ū baqēt*   kuryama   *zayy xaddām*ēk

                  and stayed.1SG   house.LOC   like servant.PRED.F.SG

         c.   *ū da'iman yaʕnī kunt*   ama kuryamēk   *wala*

                  and always   that.is was.1SG   I   house.LOC.PRED.F.SG   and.not

                  kilšami *wala*   awami,   *wala*   *waddi*karmi   *maḥall*ak.

                  exit..1SG and.not   come.1SG   and.not   bring.3PL.1SG   place.INDEF

         d.   *bass kānat*   dāyos   ḥayatēki ghāy wāšīm.

                  but was.3SG.F   mother.POSS   H.ABL   good   with.1SG

         a.   When I was small, they didn't send me to [any] school.
         b.   And [so] I stayed at home like a servant
         c.   And I was always I mean at home, not going out nor coming, nor
                 do they take me anywhere.
         d.   But Hayyat's mother was good to me.

All clause combining elements are Arabic-derived: the subordinating conjunction *lamma* 'when', the co-ordinating conjunctions *ū* 'and' and *bass* 'but', the filler/tag *yaʕnī* 'I mean', and the negative focus marker *wala* 'neither…nor'. Also Arabic-derived are the past-tense copula *kunt* 'I was' and *kānat* 'she was', which carry Arabic tense and person inflection; the verb *baqēt* 'I stayed', with its Arabic tense and person inflection; the comparative particle *zayy* 'like', the

adverb *da'iman* 'always', and the indefinite marker *maḥall-ak* 'anywhere' (lit. 'a-place'). Further Arabic loans are the words for 'school', 'servant', and 'to bring'.

The next examples further illustrate the non-separation of Domari and Arabic for clause-combining structures:

(2)   *qabel mā* džam *xałłaṣ*kedom kamas

before COMP go.1SG.SUBJ finished.1SG work.ACC

'Before I left I finished my work'

(3)   *iza* warsari, nawame'

if rain.3SG NEG.come.1SG.NEG

'If it rains, I shall not come'

(4)   *law* ērom xužoti *kān* laherdomsa

if came.1SG yesterday was saw.1SG.3SG.PAST

'If I had come yesterday, I would have seen him'

From Arabic we have the conjunctions *qabel mā* 'before', *iza* 'if' (realis), *law* 'if' (irrealis), and the counterpart auxiliary *kān* 'was' which marks out the Arabic counterfactual. There are no non-Arabic variants for any of these connectors, and the only conclusion is that Domari has adopted the complete set of Arabic connectors, which have replaced any older system that the language may have had prior to contact with Arabic. Domari clause-combining structures have thus undergone Fusion with Arabic.

A frequently cited view is that the indeclinable nature and sentence-peripheral position of connectors make them easy to integrate and therefore attractive targets for flagging bilingual competence (see already Poplack 1980). Some Arabic connectors however do indeed take inflection, and those keep their Arabic inflection when used in Domari. Whether or not the absence of inflection reinforces borrowing is debatable, but these examples show that inflection is at least not a hindrance for borrowing when the overall trend is toward Fusion of connectivity devices:

(5)   ama sindom   *innak*   atu mnēn hrori

       I   heard.1SG   that.2SG you   here   are.2SG

       'I heard that you are here'

(6)   na kildom   bara *li'annhā*   warsari

       NEG went.out.1SG   out because.3SG.F rain.3SG

       'I did not go out because it was raining'

(7)a.   mana *illi*   torim   *iyyāh*

       bread REL   gave.2SG.1SG   RES.3SG

       'the bread you gave me'

   b.   ple   *illi* torim   *iyyāhum*

       money REL gave.2SG.1SG   RES.3PL

       'the money(pl) you gave me'

In (5), the Arabic complementiser introducing complement clauses (*inn-*) takes the Arabic 2.SG.M inflection marker *-ak*, agreeing with the subject of the complement clause, which is expressed by the Domari inherited pronoun *atu* 'you.SG'. In (6), the Arabic causal conjunction *li'ann-* similarly agrees with the subject of the subordinated, adverbial clause expressing causality. In the particular example, the subject is impersonal, and agreement is with the F.SG, using the Arabic F.SG marker *-ha*. The choice of agreement target is itself modelled on Arabic, where the implicit subject of weather conditions is *ad-dunya* 'the world' – a feminine noun.

   In (7), finally, we find Arabic resumptive pronouns. Domari personal endings are in principle capable of marking back-reference to head nouns in object roles in relative clauses; consider *kažža illi laherd-om-is* 'man REL saw-1SG-3SG' 'the man whom I saw [him]'. However, there is no way to mark two direct objects, as required by the verb 'to give' (based on the Arabic model for this verb; an alternative, also found in Domari, is to assign to the recipient the Benefactive rather than Accusative/Direct Object case: *illi tor-is amake* 'REL gave-3SG to.me' 'which you gave [it] to me'). The first accusative, expressing the recipient, is

marked by the personal object affix on the verb, while the Arabic-derived resumptive pronoun agrees with the head noun, marking it too for the Accusative/Direct Object case. The productivity of Arabic inflectional patterns is illustrated by the switch to plural agreement in (7)b.. Now, in Arabic, *iyyāh* can express a variety of Direct Objects, and is not confined to the function of a resumptive pronoun (*aʕṭīk iyyāh* 'I shall give *it* to you'). These other usages of *iyyāh* are not found in Domari, however. Clearly, what is borrowed here is not a pronoun (with its inflection agreement patterns), but strictly a *resumptive* pronoun as part of the frame of connectivity in a complex clause. Looking at the same facts again in another perspective, the presence of inflection and agreement patterns does not hinder the incorporation of connectivity devices into Domari. Rather, these are carried over as part of the wholesale adoption of connectivity structures. Inflection, we might say, is replicated as a kind of 'accommodation layer' to allow the wholesale, smooth and efficient replication of the Arabic system of connectivity devices.

We now turn to a level beyond connectivity. Recall that not just conjunctions but also focus particles, phasal adverbs, and comparatives/ superlatives are imported from Arabic. We could treat these, functionally, as elements that convey an attempt to match the speaker's evaluation of the propositional content to that of the hearer: thus something is *bigger* than expected, an attribute applies *even* under stated conditions and hence somewhat contrary to expectations, and event is *still* ongoing and hence we presuppose knowledge about it having begun earlier, and so on and so forth. In other words, there is a degree of explicit reference to hearer-sided expectation in respect of the relevance of the proposition. Unlike connectors and discourse-operators, though, there is no evidence for actual monitoring of hearer-participation or of intervention with hearer-sided processing in the form of a call for revision of expectations from the course of the turn (as with contrastive connectors, for instance, or with fillers and tags). We might therefore rank this pool of expressions lower than the connectivity devices discussed above for monitoring-and-directing operations, and so for what I shall term – in an ad hoc way, for the

time being – the 'co-operation effort' extended by the speaker to take into
consideration the hearer's communicative state-of-mind and readiness to accept
the speaker's assertions and follow the speaker's line of argument. It makes sense
when we find these particular categories also in second position on the borrowing
hierarchy, though nonetheless closely following connectivity structures.

       Along with these devices that are second-ranking for co-operation effort
we find auxiliaries. Auxiliaries are a grammaticised and so regular structure that
modifies predicates as expressions of events and actions. They are evaluative in
that they modify statements about predicates. Modality auxiliaries directly relate
to expectations, attitudes, and external forces influencing the main event:

(8)   *xallīhum* skunnhōšad   barariyamma

       let.3PL   live.SUBJ.3PL   outside.LOC

       'Let them live outdoors.'

(9)   *baʕdēn kānat*   šara   amake *biddhā*   qumnar

       then   was.3SG.F   say.3SG.PAST me.BEN wants.3SG.F eat.SUBJ.3SG

       'Then she used to say to me that she wants to eat'

(10)   *ū*   *ṣār*   *mna*ʕkari dāyim *yaʕni* na *xarriʃ*hōšar

       and began.3SG.M prevent.3SG mother.1SG I.mean NEG talk.SUBJ.3SG

       wāšīm, *ū* ma laharim

       with.1SG and NEG see.SUBJ.3SG.1SG

       'and he started to prevent my mother I mean she shouldn't talk to me and
       shouldn't see me'

   In (8) the modal verb *xallīhum* 'allow them!' is Arabic-derived, and takes
Arabic inflection (imperative inflection + 3.PL object). In (9) the modal verb
biddhā 'she wants' is Arabic, and takes again the Arabic 3.SG.F person inflection
(here: nominal inflection on the impersonal modal expression). This also
introduces a new typological distinction into Domari, namely gender distinction,

which is otherwise lacking in Domari 3.SG anaphora. In (9) we also see the aspectual auxiliary *kānat* 'she used to …', again with tense, person and gender inflection from Arabic; I shall return to this auxiliary below. In (10) finally we find the verb *ṣār* 'he began', once again with its Arabic tense, person and gender inflection. We can conclude from these representative examples that, much like connectivity, the domain of modal auxiliaries is replicated wholesale in Domari based on its Arabic source. There are however exceptions, and Fusion here has not reached its full extent: Thus the verb *sak-* 'to be able to' is Indic (inherited). I interpret this as a hint on the direction of the cline. The expression of speaker's ability signals independence of external forces, and matches metaphorically those positions on the pragmatic-interactional cline that express less need to convince the hearer, less dependency from a speaker's point of view on the hearer's co-operation, and so a reduced rather than enhanced co-operation effort. The picture that is emerging is therefore one of a conceptual categorisation onto which structural categories are mapped hierarchically. This categorisation comprises parallel clines which could be labelled 'close vs. remote', 'independent vs. dependent', requiring 'less vs. more' co-operation on the hearer's part, and 'less vs. more' intervention with hearer-sided processing (monitoring-and-directing). I shall return to this issue in the concluding remarks.

Another auxiliary that is Arabic-derived is *kān* 'to be'. Existentials belong to the elements for which borrowing is rarely attested. In Domari, the imperfect tense, which can be used to express past-progressive or past-habitual, is formed by altering the final vowel on the present indicative: *ka(ra)di* 'they are doing', *ka(ra)da* 'they were doing'. The same device derives the pluperfect/counterfactual from the simple past: *kardom* 'I did', *kardoma* 'I had done'. Although this device is productive in Domari, it is usually reinforced by the Arabic auxiliary *kān* carrying Arabic (past) tense and concord inflection:

(11)   *kānu*   *lamma* qayiškada   *kunt*   wēštama   wāšīsan
      was.3PL   when   cook.3PL.PAST was.1.SG   sit.1SG.PAST   with.3PL
      'When they were cooking I used to sit with them.'

In Palestinian Arabic, *kān* derives past progressive, habitual, and counterfactual meanings from the present participle, present tense, and past tense respectively. There is no synthetic construction. The semantics of these derivations are clearly aspectual, thus having to do with the internal structure of the event, rather than with attitudes and external forces acting upon it (modality). How then does the borrowing of aspect fit into the cognitive-based cline hypothesised above (and throughout this paper)?

A detailed analysis of the Domari TAM-system is beyond the scope of this paper, but the essentials are these: The primary opposition, and one that is expressed (as elsewhere in New Indo-Aryan), through verb stem inflection, is that of perfectivity vs. absence of perfectivity (*karami* 'I do', *kardom* 'I did'). Tense is external to the stem as well as to the expression of concord – again as elsewhere in New Indo-Aryan (cf. MASICA 1991) – and marks relevance to the present context of speech or discourse (*karama* 'I do [+remote]' = 'I used to do, was doing'; *kardomi* 'I have done', *kardoma* 'I had done'). Strictly speaking, then, what Arabic auxiliaries do is apply a situative perspective to the degree of relevance of an action to the present context, in other words they help evaluate the event from the shared perspective of speaking time. But a crucial key to understanding the role of Arabic *kān* in Domari is its role as a TAM-carrier for the copula or existential verb.

The Domari copula itself is not borrowed from Arabic. Domari and Arabic differ typologically in their formation of existential clauses. Domari retains an enclitic copula (deriving from the perfect tense of the verb *bhu- > ho-* 'to become'), which can be regarded as a reflection of earlier (Middle Indo-Aryan and pre-contact Indic) verb-final order in the language:

(12)   ama mišta-hromi

       I   ill-COP.1SG.PRES

       'I am ill'

Arabic on the other hand lacks a present-tense copula altogether, and shows a nominal predication instead, with the negator *mišš* acting as a nominal negator:

(13)   a.   'ana ʕayyān   (Arabic)

                I   ill.SG.M

              'I am ill'

       b.   'ana mišš ʕayyān

                I   NEG ill.SG.M

              'I am not ill'

Domari copies this construction in the negative, thus equating its enclitic copula – notably, once again, the only instance of a clause-final verb in the language – with the Arabic nominal sentence:

(14)   ama *mišš* mišta-hromi

        I   NEG ill-COP.1SG.PRES

        'I am not ill'

Once this equation of the Domari enclitic copula with the Arabic nominal existential construction appears, similar modification is applied not just for negation, but also for TAM. For the past and future tenses, Arabic existential constructions draw on the inflected copula *kān* and, for negation, on the verbal negator *mā* (15), and this construction is replicated in Domari as well (16):

(15)   'ana mā kunt   ʕayyān   (Arabic)

        I   NEG was.1SG ill.SG.M

      'I was not ill'

(16)   ama *mā kunt*   mišta-hroma

        I   NEG was.1SG ill-COP.1SG.PAST

      'I was not ill'

Let us then review the process: Domari undergoes syntactic convergence with Arabic in virtually all aspects of clause organisation (the only exception being the choice of indicative rather than subjunctive in Domari for the complement verb of the modal auxiliary 'to begin'; cf. MATRAS 1999). The only syntactic-typological discrepancy between the languages is the retention in Domari of an enclitic copula, and thus the existence of a copula-verbal predication in the present tense. This discrepancy is reconciled by re-interpreting the Domari copula predication as matching the Arabic non-verbal (nominal) present-tense existential construction. An indication for this is the use, in Domari, of the Arabic nominal negator *mišš* with the enclitic present-tense copula. The matching extends to the past (and future) tense, where Arabic has a verbal copula *kān*, which however is also used with lexical verbs as an auxiliary. The presence of *kān* without Arabic person inflection in the Domari corpus published by MACALISTER (1914) gives some historical evidence in support of this development path: *kān* seems to have made its way from an impersonal marker to a genuine auxiliary. The important thing is that this development is an expression of accommodation to Arabic at two distinct levels: It is part of the process of syntactic convergence, affecting also existential predications; and it is part of the gradual wholesale replication or Fusion with Arabic in the domain of auxiliary representation and so evaluation of propositions (specifically: predications).

4. Conclusion: Fusion and a model of contact

We have seen that Domari is a language with 'heavy borrowing', and further that entire categories have been replicated, wholesale, from Arabic. One of these categories is the set of prepositions, which I have not discussed here, but which indicate that Fusion reaches far beyond the levels hypothesised to be affected first and foremostly by contact-induced change: the system of connectivity, and the system of expressing attitudes toward propositions and predications. But although Domari has borrowed extensively, adopting entire categories, borrowing has not
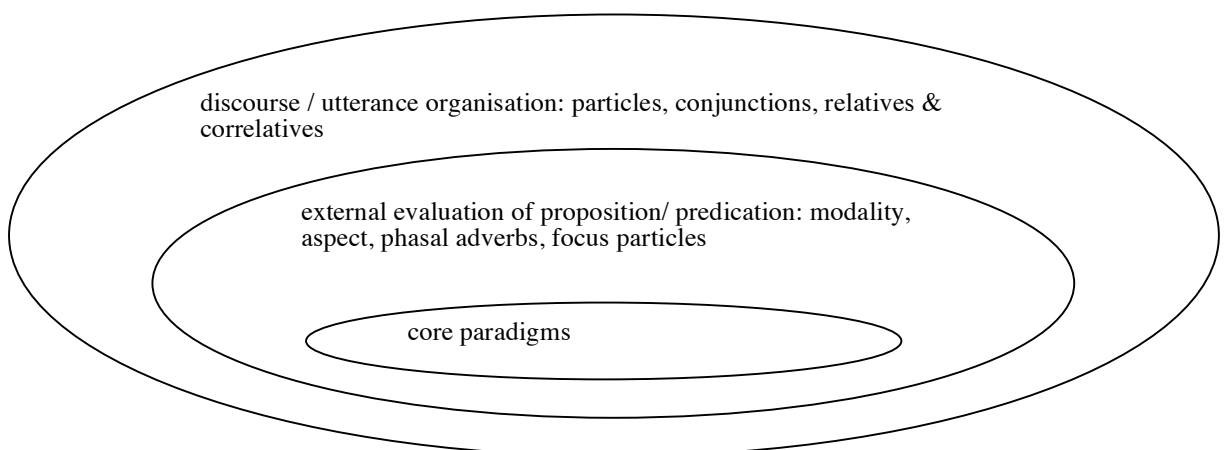
infiltrated certain structural domains: The morphology of lexical verbs, nominal inflection, and the systems of deictic and anaphoric reference remain intact. There seem therefore to be core paradigms that are resistant to borrowing. Nonetheless, Fusion at the level of connectivity and expression of attitudes may override some of these constraints on direct borrowing into paradigms, and hence we have Arabic-inflected resumptive pronouns or complementisers in the connectivity domain, or Arabic-inflected auxiliaries in the 'attitude' domain. The result is a language that maintains its structural autonomy, yet undergoes Fusion – non-separation of form and structure – with its contact language for a distinct set of categories.

My argument in this paper (as in some previous work which I referred to above) is that, identifying the categories that undergo Fusion, and perhaps even some clues as to the sequence in which they undergo Fusion as separate categories, as well as the sequence in which individual member forms of these categories fuse – all this may provide insights into the functionality of contact-induced change. This is not to say that any language will always change through contact in order to be functional; that of course cannot be maintained. Rather, what I mean by 'functionality' is that, if languages change through contact, this change will not be entirely accidental. So far, the non-arbitrariness of contact-induced change has been dealt with in the literature primarily at the level of formal structures (e.g. bound vs. unbound), or at the sociolinguistic level (more intense and prolonged cultural contact will lead to heavier borrowing). I suggest that non-arbitrariness of contact-induced change must be examined at the level of functions – with which I mean the functions of categories in processing language, in categorising linguistic structures in relation to categories of perception of the real world (cognitive categorisation), and in regulating communicative interaction.

Above I tentatively identified several inter-related clines of borrowing (see also MATRAS 2002, Ch 8.): 'close–remote', 'independent–dependent', requiring 'less' vs. 'more' explicit co-operation with the hearer, and 'less' vs. 'more' intervention on the part of the speaker with hearer-sided processing (also

referred to as monitoring-and-directing operations). Naturally, the more we try to identify a common denominator for these clines, the more abstract the common theme is likely to be. My conclusion on this particular issue of an overall motivation for contact-induced change will therefore remain an abstract hypothesis, and some might say speculation. But putting the pieces together, what seems to me to merit consideration as an overall motivating force might be defined as a principle of 'co-operation effort' in communication: the 'broader', 'upper-level', 'larger' units or 'remote' concepts are those that require stronger, more controlled, more intense, and more explicit effort on the part of the speaker to successfully manage the interaction. This in turn brings us to the, in evolutionary terms, 'primitive' functions of language – to negotiate situations with interlocutors. The more intense the speaker's effort to maintain control of the communitcative situation and ensure cooperation between himself and the hearer, the more instinctive or automatic his linguistic gestures are. And the more automaticised the gestures, the more likely they are to form an indiscriminate system in which choice among different sets is abolished in favour of efficiency of reaction. Thus, the roots of contact-induced change might be said to reflect the roots of the language capacity itself.

Figure 1: 'Outer' and 'inner' categories on the cline of contact-susceptibility

discourse / utterance organisation: particles, conjunctions, relatives & correlatives

external evaluation of proposition/ predication: modality, aspect, phasal adverbs, focus particles

core paradigms

References

BAKKER, PETER & MOUS, MAARTEN. eds. 1994. *Mixed languages. 15 case studies in language intertwining*. Amsterdam: IFOTT.

BAKKER, PETER. 1997. *A language of our own. The genesis of Michif — the mixed Cree-French language of the Canadian Métis*. New York: Oxford University Press.

DE ROOIJ, VINCENT A. 1996. *Cohesion through contrast: Discourse structure in Shaba Swahili/French conversations*. Amsterdam: IFOTT.

HANCOCK, IAN F. 1988. "The development of Romani linguistics", in: JAZAYERY, MOHAMMAD ALI, & WINTER, WERNER. eds. *Languages and cultures. Essays in honour of Edgar Polomé*, 183-223. Berlin: Mouton de Gruyter.

HANCOCK, IAN F. 1995. "On the migration and affiliation of the D\ōmba: Iranian words in Rom, Lom, and Dom Gypsy", in: MATRAS, YARON. ed. *Romani in contact*. 25-51. Amsterdam: Benjamins.

MACALISTER, R. A. S. 1914. *The language of the Nawar of Zutt, the nomad smiths of Palestine*. (Gypsy Lore Society Monographs 3). London: Edinburgh University Press.

MASCHLER, YAEL. 1994. Metalanguaging and discourse markers in bilingual conversation, in: *Language in Society* 23, 325-366.

MASICA, COLIN. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.

MATRAS, YARON. 1996. Prozedurale Fusion: Grammatische Interferenzschichten im Romanes, in: *Sprachtypologie und Universalienforschung* 49: 60-78.

MATRAS, YARON. 1998. Utterance modifiers and universals of grammatical borrowing, in: *Linguistics* 36: 281-331.

MATRAS, YARON. 1999. The state of present-day Domari in Jerusalem, in: *Mediterranean Language Review* 11, 1-58.

MATRAS, YARON. 2000. Fusion and cognitive basis for bilingual discourse markers, in: *International Journal of Bilingualism* 4: 505-528.

MATRAS, YARON. 2002. *Romani: a linguistic introduction*. Cambridge: University Press.

MUYSKEN PIETER. 1997. "Media Lengua", in: THOMASON, SARAH G. ed. *Contact languages*. 365-426. Amsterdam: John Benjamins.

POPLACK, SHANA. 1980. Sometimes I'll start a sentence in Spanish y termino en español: toward a typology of code-switching, in: *Linguistics* 18, 581-618.

POTT. AUGUST F. 1846. Über die Sprache der Zigeuner in Syrien, in: *Zeitschrift für die Wissenschaft der Sprache* 1: 175-186.

ROSS, MALCOLM. 2001. "Contact-induced change in Oceanic languages in north-west Melanesia", in: DIXON, R.M.W. & AIKHENVALD, ALEXANDRA Y. eds. *Areal diffusion and genetic inheritance: problems in comparative linguistics*, 134-166. Oxford: Oxford University Press.

SALMONS, JOE. 1990. Bilingual discourse marking: code switching, borrowing, and convergence in some German-American dialects. *Linguistics* 28: 453-480.

STOLZ, CHRISTEL, and STOLZ, THOMAS. 1996. Funktionswortentlehnung in Mesoamerika. Spanisch-amerindischer Sprachkontakt Hispanoindiana II. *Sprachtypologie und Universalienforschung* 49: 86-123.

THOMASON, S. G., and KAUFMAN, T. 1988. *Language contact, creolization, and genetic linguistics*. Berkely: University of California Press.

THOMASON, SARAH GREY. 1995. Language mixture: ordinary processes, extraordinary results. In: Silva-Corvalán, Carmen. ed. Spanish in four continents. Studies in language contact and bilingualism. Washington D.C.: Georgetown University Press. 15-33.

THOMASON, SARAH GREY. 1997d. Mednyj Aleut. In: Thomason, Sarah Grey, ed. Contact languages. Amsterdam: Benjamins. 449-468.