Language contact, language endangerment, and the role of the 'salvation linguist'

YARON MATRAS

_____

_____

EL Publishing

Language contact, language endangerment, and the role of the 'salvation linguist'

Yaron Matras

1. Preface

In the first part of this paper I address concerns in respect of certain images and notions that surround the current agenda of the study of endangered languages. In particular, I question the usefulness of, and point out some risks attached to, the self-proclaimed moral high ground of the 'salvation linguists', those who seem to carry the banner of language conservationism. I then go on to discuss some examples for the potential research yields in the field of endangered languages, basing my observations on two case studies, the documentation of Domari (an Indo-Aryan language spoken by traditionally peripatetic communities in the Middle East), and the documentation of so-called Angloromani (the mixed speech varieties used by Romanies in England and Wales). Special attention will be given to the role of language contact, arguably a principal common denominator of endangered languages. I conclude by returning to the issue of activism and the role of the linguist studying an endangered language.

2. Language conservationism and the moral high ground

It is not unusual for research and (to a lesser extent, perhaps) teaching agendas to be susceptible to trends and fashions. In linguistics, trends tend to comprise both theoretical approaches to language (see, for example, the rise – and fall? – of Optimality Theory over the past decade), as well as topical areas, such as the study of pidgins and creoles, or speculations on language evolution. Language endangerment has conquered a central position on linguists' agendas over the past few years. Not only are there conferences, societies, and newsletters devoted to the subject, there are now also several programmes that provide grants specifically for the documentation and study of endangered languages, and there are training programmes for postgraduate students who wish to specialise in this area.

Non-clinical linguistics has offered non-academic audiences inspiration in the past, in the study of slang and dialects, in teaching foreign languages, and on issues of language standardisation, language policy, and linguistic identity. The primary concern around such questions has been the practical support that research can give in solving problems of communication, or in increasing awareness of identities and so potentially encouraging participation and self-fulfilment. Labov's (1970) work on Black Vernacular English is a classic example. But language endangerment is arguably the first-ever topic in linguistics to attract and demand attention on moral grounds. There are, of course, academic

arguments in support of the documentation of endangered languages. The most obvious of those is the fact that a full and comprehensive documentation of today's languages, living and moribund, offers a sample of linguistic diversity on a scale which, as a result of the ongoing massive decline of languages, we shall never be able to witness again. Linguists therefore have a strictly academic interest in securing this sample of data as a matter of urgency. Other questions relate to the sociolinguistic aspects of endangerment, the causes of language abandonment, for instance, or the ways to support language maintenance, as well as to structure-related aspects, the role of variation in situations of lax normative control, the role of multilingualism in cross-generation communication, or the structural impact of language contact.

But alongside these and other issues of concern, endangered language activists tend to emphasise a kind of moral responsibility to document languages before they disappear, or even to offer support to prevent the disappearance of languages. Maffi's (1996) position paper on endangered languages, for example, states that:

> An ever growing body of literature on endangered languages, vanishing cultures, and biodiversity loss has been accumulating in recent years, attesting to the perceived gravity and urgency of such issues. Underlying this concern is a common interest in the future of humanity and the earth's ecosystems.

In adopting such an attitude toward research on endangered languages, some linguists have chosen to imitate the role of activists, similar to the role assumed by those engaged in conservationist, environmentalist, and similar political initiatives (cf. Grinevald 2003). An attitude is gaining ground among linguists that the description of endangered languages is not just of interest to the research community, but a service to the speaker communities and so to human society in general; it is not just a topic of research, but a mission (cf. Nettle and Romaine (2000:153)), who even coin a term to refer to those who do not share the mission – "advocates of the benign neglect position"). Unlike the case of linguists who assume advisory roles in education, media, or the drafting of language policy, many linguists specialising in endangered languages view themselves as pursuing an urgent mission, one that is of top priority because of the pace at which languages are disappearing. Crystal's (2000:166) closing words illustrate this:

> The alternative is to act, using as many means as possible to confront the situation and influence the outcome. We know that intervention can be successful. Revitalisation schemes can work. But time is running out. It is already too late for many languages, but we hold the future of many others in our hands. The linguists in the front line, who are actually doing the fieldwork, therefore need as much support as we can mobilise.

Such wordings as Crystal's "linguists in the front line" adds to the dramatisation, as does the depiction of numbers of languages that disappear every year[1].

It is naturally gratifying to be able to say, as a linguist, that one has rendered a service to a community, and perhaps managed to 'save' an important aspect of the community's culture, though I am uncertain whether any linguist, however active and engaged, is actually in a position to say this. But the claim to a moral high ground in the preoccupation with endangered languages carries with it a number of risks. First, it risks dividing the community of linguists into those who regard themselves as linguists who 'do good', and those who are simply and plainly concerned with pursuing academic agendas in their research. A number of grant schemes already link research grant eligibility to the provision of a letter of reference from what is referred to as '*the* speech community', itself a kind of neo-colonialist image, as if the speech community of an endangered language is by necessity a monolithic interest group. The procedure implies that a class of eligible linguists can be identified who are officially licensed to carry out activities in the name of communities, 'salvation linguists'.

In the process, there is the risk of projecting attitudes and expectations on to the respective community. In many communities around the world, language is not an issue of particular concern, and the disappearance of a language is an accepted fact of life. The present author, for instance, grew up in a community in which the elderly generation spoke Yiddish, and even believed, as a child, that one learns Yiddish when one becomes a grandparent. The fact that younger generations did not speak Yiddish was simply a fact of life, and was neither contested nor mourned. An outsider approaching the community at the time with the suggestion that a tragedy or other dramatic event was under way merely because Yiddish was spoken only by the old people, and not passed on to the young, would have encountered a bewildered and bemused reaction.

That it is not to say that it is wrong to suggest that the process of language shift might be reversed, and that measures to maintain languages might be adopted. But the presumption that every community wishes its language to be salvaged by a linguist, and hence is waiting to entrust a linguist with the task, is a rather pretentious self-ascription on the part of the salvation linguists themselves, and has little to do with empowerment of speakers to act in their own interest, the alleged goal of language conservationism (cf. the discussion in Nettle and Romaine (2000:150ff)). It is noteworthy that part of the imagery surrounding endangered language communities is their exotic depiction as groups that are not only disenfranchised and disempowered, but also archaic and remote from urban civilisation, both physically and in the stage of their technological and institutional development. Endangered European languages, like Frisian, Yiddish, or Low German, have become almost unwelcome distractions in the new trend of research events devoted to endangered languages. It is true that much more is known about their structure and history

---

[1] See UNESCO's website on endangered languages: http://www.unesco.org/courier/2000_04/uk/doss03.htm.

compared to languages of the Amazon, British Columbia, or Irian Jaya; though on the other hand the reasons behind their endangerment are arguably even more complex and certainly not unworthy of in-depth sociolinguistic investigation. But the preferred image of linguistic fieldwork as involving discomfort and even danger to the linguist reveals a need to celebrate the researcher as a hero.

We take a further risk if we choose to focus our training efforts among a new generation of linguists on language conservationism and activism, at the expense of preparing them for the contradictions that one encounters in the field, and indeed for the very realistic possibility of a conflict of interests between the researcher and parts of the speech community whose language is the subject of the research. Not every speech community is interested in language conservationism, or in celebrating a linguist who salvages its culture. Among the millions of people who are speakers of endangered languages, some are keen to exploit outsiders' interest in their culture in order to make a personal profit, and others will jealously guard and control access to their language; others still may be entirely indifferent. While ethical standards should form part of any fieldwork training, and we should not encourage researchers to work against the wishes of the community as a whole or of individuals within it, it seems irresponsible to prepare trainee researchers solely for the possibility that they will be celebrated and supported by the speakers whom they study. I will return to this issue below, and argue that a responsible training programme must advocate a more modest role, while preparing students for the fact that *the* expectation of *the* community *per se* does not exist, and that a range of attitudes may be encountered, for which a range of different response strategies is needed (see also Dobrin, this volume).

3. The languages

The remainder of this paper deals with examples from work on two endangered languages. The first is Domari (also referred to by its speakers as *domi*). Domari and related varieties (known as *qurbati*, *domani*, as well as by several other names) are Indo-Aryan languages spoken in the Middle East by populations who are normally referred to in Arabic as *nawar*, in a free English translation, 'Gypsies'. The variety with which we are concerned here is spoken in Jerusalem, and is more or less identical to dialects of the language spoken in Gaza and in Jordan. There is little documentation on Domari or related varieties from other countries in the Middle East, and no reliable figure about the numbers of speakers in other countries are available. Most linguistic information is limited to wordlists or short samples of sentences, much of which was collected and published in the nineteenth and early twentieth centuries. Informal reports suggest that Domari and related varieties are spoken in eastern Anatolia, northern Syria, the Damascus area (for the latter see Meyer 1994), Lebanon, and Jordan. My own brief encounters with Dom from Jordan and from Gaza suggest that, there too, as in Jerusalem, Domari is only spoken by the older generation. Jerusalem Domari was first documented rather extensively – through a collection of over

100 translated short stories, a dictionary of some 1,000 words, and a 35-page grammatical sketch – by Macalister (1914). Between 1996-2000 I visited the Jerusalem Dom community on a regular basis, and collected both conversational data and elicited data.[2] Some data and evaluation has been published (Matras 1999, 2000, 2003a, 2005), and a grammar is in preparation.

The Palestinian Dom were a tent-dwelling, peripatetic community specialising in metalwork and entertainment (musicians and dancers). The Jerusalem community, comprising altogether three extended families, emphasise their specialisation in metalwork, though dancing, music, and begging were also traditional sources of income until the early 1970s. The community went through a major transformation during the early period of British rule in Palestine, in the 1920s, when municipal services were set up and Dom men were recruited to work for the sanitation department as waste collectors and caretakers in public lavatories. The move into wage labour was followed by a move from tents into rented permanent accommodation in a narrow street in the north-eastern corner of the Old City of Jerusalem, adjacent to the City Wall, where most of the community members still reside today. The first generation of wage labourers continued to work for the Jordanian, then (after 1967) for the Israeli municipal authorities, and retired after the annexation of East Jerusalem by Israel and the introduction of the Israeli pension system. The generation born in the 1960s was the first to attend school (in Arabic), and now typically seeks employment in construction or industry in and around (Israeli) West Jerusalem.

The Jerusalem community is one of two communities of Palestinian Dom. The other is in Gaza, and consists of refugees who left Jaffa during the 1948 war, and their descendants. In both communities, only the oldest generation uses the Domari language. In Jerusalem, the shift occurred in the late 1960s, when the first generation of children who attended compulsory schooling were spoken to in Arabic, and Arabic quickly took over as the language of cross-generation communication. Individuals born after 1960 generally understand Domari, but are unable to repeat sentences word by word, and are unable to form grammatical sentences, beyond repetition of isolated phrases. The youngest fluent speaker was born in 1955, and is an exception in the community, having been orphaned at an early age and raised by her grandparents, without attending school. Most fluent speakers were born before 1945. No precise figures are available on the total number of the Dom population of Jerusalem, and it is estimated at somewhere between 600-900 (including a number of households who have moved out of the Old City and into suburbs on the outskirts of Jerusalem). No more than ten percent of those are fluent Domari speakers. The numbers in Gaza are unknown, but from casual encounters with Dom families from Gaza it appears that the proportion of speakers is similar (the language being preserved only by those born before ca. 1945).

Domari preserves an archaic Middle Indo-Aryan present-tense conjugation, and a past-tense conjugation system that resembles that of the northwest Indian 'frontier' languages, such as Kashmiri (derived from attaching oblique personal pronouns to the past participle). So-called Layer II case markers, which are clitics in most modern Indo-Aryan languages, have developed in Domari (as in Romani) into fully agglutinated case suffixes. Although these general morphological traits are shared with Romani, some early phonological features, as well as the overall composition of grammatical vocabulary indicate that the two languages may have undergone similar developments before leaving India. They were not simply dialects of the same ancestor language, but derive from different Indian idioms. The similarity in the origins and socio-economic profile of the two populations, as indeed the similarity between the self-appellations *dom* and *řom*, can be derived from an origin of both groups in the caste of the *ŋom* commercial nomads of the Indian subcontinent.

There is thus an ancient connection between Domari, and the second language under consideration here, which we shall call 'Angloromani'. The term is not, strictly speaking, a self-appellation, but rather a technical term coined by linguists (cf. Hancock 1970, 1984). Speakers tend to refer to the language as *romanes* or *romani jib* 'Romani language'; this is the reason that linguists have chosen an alternative designation – in order to distinguish between the (inflected) Romani language and its various dialects, and the particular kind of selective use of Romani-derived lexicon within an English utterance and discourse framework, which is what 'Angloromani' essentially is. Although the circumstances of the transition are not entirely clear, and interpretations are controversial, it is generally assumed that Angloromani largely replaced (inflected) Romani as the community language of Romanies in England and Wales toward the end of the nineteenth century. Monograph-length descriptions of inflected Romani document the use of the language in England (Smart and Crofton 1875) and in Wales (Sampson 1926), by what was probably the last generation of fluent speakers.

The form of the language that still survives today, but is reported by speakers to be in decline, consists of the optional insertion of Romani-derived words and phrases into an ethnolectal English discourse ('ethnolectal' English being the distinct variety of English that is spoken by English Romani Travellers). These words and phrases constitute a repertoire of perhaps several hundred core items, with considerable variation among speakers and families. The survival of Romani-derived vocabulary after language shift is attested in other Romani communities in Europe, most notably in Scandinavia and the Iberian peninsula, but also in other regions, and has been coined 'Para-Romani' in the Romani linguistic literature (Cortiade 1991, Bakker and Cortiade 1991, Matras 1998b). For speakers, unaware of an inflected Romani until the recent immigration of Romani-speaking immigrants from eastern Europe, this style of inserting Romani items into everyday discourse is what sets apart 'Romani' speech from 'ordinary' speech. Regardless of the essentially stylistic or register-like character of Angloromani, it is perceived by speakers as

a language in its own right, one that is reserved to particular contexts and settings, and which flags group identity and, in the modern sense, ethnicity.

There are no figures or even estimates of the number of users of Angloromani. The number of English Travellers might be estimated at between 40,000-60,000[3]. Due to the nature of Angloromani as a Romani-oriented style or register, manifested primarily through the use of particular lexicon, it is difficult to even construct strict criteria for 'speakers' or 'users'. In the perception of community members, certain older people are known as individuals who are familiar with a greater extent of vocabulary, and who use or used this vocabulary more frequently in the past. Middle-aged users tend to describe their own knowledge as being in decline, and to report much more frequent exposure to Romani in the past, as children witnessing their parents' linguistic behaviour. This overall impression is confirmed, to some extent at least, by word elicitation, which suggests that some (usually older) individuals are more likely to have knowledge of a more extensive vocabulary than others. Apart from the self-evaluations and so far rather fragmented elicitation comparisons, however, it is difficult to provide an objective picture of an ongoing decline; indeed, political activists appear to over-report use of Romani lexicon. It is, however, beyond doubt that school attendance has increased over the past generation, as has work outside the strict family network. Both these factors lead to a decrease in the amount of time spent within the family and so in the proportion of communicative interaction contexts in which Angloromani is used even potentially. Moreover, there is apparently also an increase in intermarriage between Travellers of Romani background, and those of Scottish and Irish background (who have no Romani roots), carrying with it, again potentially, a decline in the function of Angloromani as a family or in-group code.

It is in any event worth noting once again that Angloromani itself succeeds an inflected form of Romani, which is documented and which was spoken by Romanies in England and Wales until the end of the nineteenth century.[4] In this sense, Angloromani has survived inflected British Romani, which was endangered a century ago, and has since died out. At present, Angloromani itself appears to be in decline, though due to its nature as an optional style characterised primarily by lexicon, the boundary between using Angloromani, and not using it, is extremely fuzzy.

---

[3] Emphasising the absence of any actual data, the Commission for Racial Equality cites an estimate of between 200,000-300,000 British and Irish Travellers (http://www.cre.gov.uk/gdpract/g_and_t_facts.html), which seems fantastically high, even though it includes Irish Travellers (who are not users of Romani).

[4] While this interpretation seems indisputable, it has in fact been challenged by anthropologist Judith Okely (1983), who suggested instead that the source of Indic vocabulary was not an inflected language, once spoken by a community with Indian roots, but rather that Traveller-Gypsies picked up lexical items of Indic origin along the trade roots. The most obvious argument against such an interpretation is the fact that inflected Romani is indeed documented in Britain, and that British Para-Romani clearly descends from those inflected British Romani varieties (taking into account their distinct structural traits).

4. Endangered languages as unique sources of data

Is there a common denominator to endangered languages, one that would justify lumping them together as a focal area of investigation? Apart from their position on an agenda of urgent documentation, and the documentation methodology which such an agenda entails, it is not obvious that endangered languages should have any features in common, certainly not in the conventional sense of 'structural' features. What they share is, rather, the fact that they are potential sources of data that cannot easily be found elsewhere. There are two ways to approach this. First, the sample of endangered languages may contain greater structural diversity than the sample of 'larger' languages that enjoy a more secure position, hence the potential of endangered languages to offer insights into structural features that cannot be found in a less extensive language sample. Second, there is something about the situation of endangerment itself which shapes languages in a particular way – socially, and perhaps even structurally. I shall explore several issues related to both these assumptions.

4.1 Structural oddities

A well-cited argument in support of extensive documentation of endangered languages is the extent of structural diversity among today's languages, which is likely to be lost within a generation or so. Consequently, we will no longer be able to observe certain structural features of language unless they are documented now.

Domari offers such a unique, hitherto undocumented feature in its consistent use of suppletive comparative/superlative form of adjectives. Domari adjectives agree with the noun in gender and number. They are traditionally placed in front of the noun, thus *tilla zara* 'the big boy', *tillī čōnī* 'the big girl', though a predicative construction is gradually taking over all attributive constructions, presumably due to the fact that, like Arabic, it presents the adjective in post-nominal position (cf. *zara tillēk* 'the boy, being big = the big boy', *čōnī tillik* 'the girl, being big = the big girl'). Domari does not have any inherited, productive morphological procedure to transform positive forms of the adjective into comparative or superlative forms. Instead, the respective comparative/superlative form of the corresponding Arabic adjective is used: *aha zara/ ihi čōnī akbár-i* 'this boy / this girl is bigger/biggest'. As a result, each and every inherited adjective, such as *tilla* 'big', has a suppletive comparative/superlative form, such as *ákbar* (the exception are, of course, adjectives whose positive form is also borrowed from Arabic).

We can hypothesise that this process derives from a cognitive motivation to borrow the procedure used to construct comparison[5] (cf. Romani dialects, which also tend to

---

[5] 'Cognitive', since comparison, much like focus particles, connectors, privative expressions, etc. can be considered as high on the relevance scale of information-processing and so susceptible to cognitive pressure to reduce the inventory of forms in the bilingual repertoire to just one single set, with that of the 'outsider' language taking over for obvious reasons of prestige and functionality (since a reduction of the two sets in

borrow the comparative and superlative markers from their respective contact languages, Turkish *daha*, Romanian *mai*, Slavic *po*, and so on, but attach them to the inherited positive form of the Romani adjective). In Arabic, however, the comparative is not easily analysable morphologically, as it is based on a vocalic alternation that is internal to the stem, and so not segmentable (cf. Arabic *kbīr* 'big', *ákbar* 'bigger'). As a result, the complete Arabic word-form is borrowed. The case of Domari adjectives is an interesting illustration not just of the fate of adjectives, but of the possible discourse-functional motivation behind suppletion in the first place. (The ordinal 'first', for instance, is also borrowed from Arabic; in numerous other languages, it is suppletive. Suppletion, like borrowing, then, appears to be a strategy to prioritise the processing of information in discourse (cf. Elšík and Matras (2005)).

Angloromani too offers a rather exceptional structural profile, not just in a particular sub-system, but in its basic structural composition, taking word-roots and some grammatical vocabulary from Romani, while relying on English for grammatical inflection, clause and discourse structure, and most of the phonology. This formula for what has been termed 'Mixed Languages' (cf. Bakker and Mous 1994, Bakker 1997, Matras and Bakker 2003) is documented in very few varieties around the globe, all of them are in the process of decline and endangerment. There are essentially three major questions surrounding the phenomenon of Mixed Languages that demand particular attention, are the focus of debate and even controversy in the study of Mixed Languages, and for which one must rely empirically on the small sample of documented varieties. The first concerns the structural profile of what is to be termed a 'Mixed Language'. Bakker and Mous's (1994) formula of a 'lexicon-grammar split' is known to be a simplification. In fact, researchers have tended to lump together varieties rather intuitively as Mixed Languages, rather than follow any strict definition. A somewhat more precise characterisation of the common denominator of those languages that have, so far, been treated as Mixed is a separation between the source of finite verb inflection, and that of a considerable part of the core vocabulary (cf. Matras 2003b). Consider that under more conventional conditions of structural borrowing it is extremely rare to find either wholesale borrowing of finite verb inflection, or wholesale borrowing of core lexicon. This trait, then, the etymological separation of finite verb inflection and core lexicon, is only found in a very small number of languages, which arise under very particular sociolinguistic conditions, and it is therefore this trait that makes them 'more mixed than conventionally mixed' and so worthy of the term 'Mixed Languages'(see also Woodbury, this volume).

The second question concerns the circumstances of emergence of Mixed Languages. This is perhaps the most controversial issue in the discussion, views ranging from Mixed Languages as the outcome of gradual structural borrowing, through Mixed

---

favour of the one from the 'insider' language would not be functional when communicating outside the group, but it is functional for communication within the minority group, all members of which are bilingual); cf. Matras (1998a).

Languages as a fossilisation of codeswitching, through Mixed Languages as abrupt creations shaped by a pre-defined cognitive blueprint (separating lexicon from grammar), and on to deliberate creations originating in word-play among bilinguals. In the absence, in most cases, of direct historical documentation of the early stages of Mixed Languages, the only empirical clues can be found in the actual structural composition of the sample languages, as well as in the comparison with synchronic language behaviour in bilingual communities. Finally, the third question is related to the mere 'languageness' of the variety, the extent to which it is both conventionalised and so predictable as a self-contained system, and its communicative functions within the community. Here, we find a broad continuum, with Mixed Languages serving as principal all-purpose family languages (native languages), as languages of choice at the level of the entire conversation or discourse, and through to mere utterances inserted occasionally in certain key positions within the (otherwise non-mixed) discourse.

As one of very few varieties that constitute our universal sample of Mixed Languages, Angloromani is therefore in a position to shed some light not just on issues of individual structures, but on the theoretical conception of what constitutes a Mixed Language. In respect of its structural composition, Angloromani nicely fits into the split between finite verb inflection, and core lexicon. A modest amount of derivational morphology is carried over from Romani, as are some function words, including deictics, further strengthening doubts as to the usefulness of the more wholesale notion of a lexicon-grammar split. As far as the history of its emergence is concerned, Angloromani arguably supplies confirmation of various hypotheses. The fact that it succeeds and replaces inflected Romani suggests a connection between the emergence of the mixed code and the decline of the inflected in-group language. Angloromani is therefore the product of language shift, co-existing with an attempt to preserve an in-group code with limited functions (and hence with limited structural resources). Nonetheless, some of the morphological erosion that characterises Angloromani appears to have actually preceded language shift (e.g. loss of gender distinctions and erosion of synthetic case), as a result of English influence, while on the other hand some stereotypical constructions (occasional use of concord markers with some verbs) appear to have survived the shift. The most plausible emergence scenario is therefore one that relies on various notions put forward in the literature so far: Gradual attrition leads to language shift away from Romani. A mixed form grew, first parallel to inflected Romani (cf. Bakker 2002), for use in particular contexts, by bilinguals consciously and habitually inserting Romani lexicon into English discourse. It then became widespread as the principal in-group code to compensate for the loss of Romani, reinforced by the selective replication of fossilised impressions of inflected Romani, and finally through lexical expansion through word-play and lexical creativity. Angloromani is thus an example of a language acquiring 'life after death'.

As for the 'languageness' issue, there are still many questions to be addressed and especially empirical evidence to be collected. So far, almost all documentation of Angloromani consists of purposefully composed text material, i.e. texts that are the product

of users' deliberate attempt to demonstrate knowledge of Romani lexicon, and of researchers' attempt to document mixture. Very few recordings of naturally occurring discourse in Angloromani exist.[6] Both observations and speakers' own reporting suggest that the insertion of Romani elements occurs primarily at the utterance level, in key utterances in the discourse, often involving some kind of pragmatic marking of the relevant utterance. At the same time, however, speakers also report both continuous insertion of Romani material in the family context, and deliberate insertion of Romani in the presence of outsiders in order to convey warnings or intimate evaluative statements. A rough idea of the range of employment of Romani can be gained from the following conversation excerpt, in which a user of Angloromani from north-west England reconstructs usage patterns (translation of Angloromani segments appears in square brackets and italics):

(1)

I'll put this in a scenario, it's like me chingerin' [*fightin'*] him, sorry, me telling him off, right, and me saying to you mandi's pooked doba [*I've said this*], or mandi's pooked lesti [*I've told him*], I told him, I've chinged lesti [*told him off*], but chinga [*scold*] means to scold, I've chinged [*scolded*] him. Things like that; if I've been somewhere and there's a man with purple lips, me dad says: dika the mush's moi, nafli zee [*look the man's face, sick heart*], look at the man's lips, he's got a bad heart, things like that. Yeh, akai [*here*], say we're working and a man comes, watch akai [*here/there*], but not directly, we'd say: mush akai [*man here*], you'd say: the konligera's avin', gava ya kukri [*the congregation's comin', hide your self/alone*], hide yourself the school board's here!

Superficially, then, it appears that Anglromani consists primarily of Romani words inserted into English utterances. However, at closer scrutiny it is obvious that different grammatical rules apply to the English component as well. Thus, the auxiliary have appears in the same form (English 3sg) regardless of person: cf. mandi's for 'I have'. Infinitive and imperatives have an inflectional vowel attached to the verb stem: chinga 'to scold', dika 'watch!', gava 'hide!'. Contextual presupposition is strong and is relied on at the expense of anaphoric, existential, possessive and other overt grammatical constructions: nafli zee [*bad heart*] for 'he's got a bad heart'. Dialectal phonology is prevalent: avin' 'coming', me 'my', ya 'your', etc. Angloromani is therefore best described as a register with a continuum of discourse-related functions, triggered by the pragmatics of interaction in response to extremely subtle contextual factors.

---

6 The Romani project at the University of Manchester has begun to collect narratives and conversational data in which Angloromani is used.

4.2 Structural variation

Moribund languages have been said to be less normative, and hence more variable (cf. Dorian 1981, 2001). On the other hand, they are often spoken in small, close-knit communities, where the inventory of extra-linguistic factors that differentiate speakers is limited. There is therefore a special challenge in trying to describe the nature and the correlates of internal structural variation in endangered languages. Domari for instance exhibits broadly two types of internal variation. The first may be defined as primarily phonetic/phonological free variation at the level of the individual speaker's use of form, and is reflected by alternation primarily of vowel quality, more rarely also of phonetic features of come consonants (for instance presence or absence of pharyngealisation, a feature imported from Arabic, in the pre-Arabic or Indic component). For the word 'small', for instance, the following variants are all attested: [qaštoṭa, qištoṭa, qištuṭa, quštuṭa, qæštoṭa, kaštoṭa, kaštoṭa]. A second type of varation is primarily, though not exclusively morphological. It concerns the tendency toward consistent use of a cluster of features by individual speakers, rather than random alternation between forms, as in the phonological variation pattern. There appear to be two such clusters of features, depicted in Table 1 as Clusters A and B:

Table 1: Cluster variation in Domari

function	Cluster A	Cluster B
Comitative case	*-san*	*maʕ .. -ki*
Ablative case	*-kī*	*min … -kī*
Dative case	*-ta*	*-ka*
3PL present tense concord on verb	*-ndi*	*-di*
Phonology	/č/	/š/

Speakers tend to be consistent in their choice of variant for the respective functions; in other words, speakers will use either *-ndi* or *-di* consistently as the marker of the 3PL, and so on. A choice of one variant from a particular cluster will usually also imply a choice in favour of the other variant in the same cluster; in other words, speakers who use *-ka* as a dative case ending will also tend to use *-di* as 3PL concord marker, while speakers who retain the affricate /č/ are also likely to retain the older comitative ending *-san*. These correlations are not absolute, however, and some exceptions do occur. On the whole, we can speak in a modest form of quasi-registers that exist within the community. Sorting out

the distribution of these registers is extremely problematic. Variants of Cluster A are on the whole more conservative: The comitative ending *-san* is inherited from Middle Indo-Aryan, while the alternative structure involves borrowing of the Arabic preposition *maʕ* with the default prepositional case (originally Ablative) *-ki*. The use of *-ki* as an Ablative case in its own right correlates with the more modest infiltration of Arabic prepositions and so with a less widespread use of *-ki* in the function of a general prepositional case, while the alternative Ablative is expressed by the Arabic preposition *min* in conjunction with *-ki* as a prepositional case. The Dative ending *-ta* is similarly a Middle Indo-Aryan retention (cf. Romani locative *-te*), while the form *-ka* appears to be a contamination of this original *-ta* with the prepositional and ablative *-ki*. The 3PL ending *-ndi* is a direct continuation of the Middle Indo-Aryan conjugation ending, while *-di* is a shortened, eroded form. And finally, the affricate is a conservative sound, which is not found in the contact language Arabic. The corresponding sibilant is the innovation; it not only has an Arabic counterpart, but the process of affricate-to-sibilant reduction in Domari resembles the one attested in Jerusalem Arabic for the voiced pair /dž/ > /ž/ (Arabic *dždīd* 'new', Jerusalem Arabic *ždīd*). Cluster B forms are thus on the whole innovations, some of them triggered directly or indirectly by contact with Arabic.

One might therefore expect the distribution of the clusters to follow age delimitation within the community. This is partly indeed the case, at least to some extent. The oldest living speakers tend to use Cluster A forms, while the youngest living speakers (though only some 25-30 years younger) tend to use Cluster B features. The distinction is blurred however for some speakers, born around the late 1920s or early 1930s, as one might expect from a transitional generation. However, Macalister (1914) already notes a tendency toward a shift from individual features which appear in our Cluster A, to those in our Cluster B. The introduction of Cluster B features therefore is not a property of the generation born in the 1920s-1930s. It seems more likely that we are dealing with traces of what used to be family dialects, at a time at which families often led independent lives that may have included a different schedule of mobility and so a different pattern of contacts with other Dom communities (and in all likelihood with other Arab communities as well). At present, family coherence at the linguistic level is almost impossible to identify due to the presence of just a few speakers in each family, and the cross-generational language divide between Domari and Arabic. Intermarriage, loss of nomadic tradition, and settlement in neighbouring building complexes that tend to house members of extended families have created a new pattern of loyalties, where individuals may flag loyalty to other individuals in the group by taking on some features of their speech. The variation pattern among the very small group of speakers has thus become a pattern of personal sets of preferences.

Angloromani too shows considerable variation patterns, most appear to be the property of family constellations. The types of variants are in line with the formation history of the language. Thus, some variation is phonological, and might reflect different degrees of accommodation to English phonology (cf. *vast* vs. *wast* 'arm'). Some variants

testify to distinct developments during the process of grammatical attrition in Romani, Thus we find in nominal forms the historical feminine in some varieties (as in *kukri* 'self'), and the historical masculine in others (*kukro*), or the original Romani feminine derivation (*gras* 'stallion', *grasni* 'mare'), alongside a newly-constructed feminine derivation (*gras* 'stallion', *gris* 'mare'). Other variants reflecting various effects of contact appear in grammatical compositions, as in *rati* 'night, tonight', preserved from inflected Romani, alongside *rati* for 'night' and *torati* for 'tonight', modelled on English. The bulk of variation reflects the degree of individual creativity in lexical composition. While some speakers retain *kans* for 'ears' (Romani *kan* 'ear', PL *kan-a*), others have the metaphorical creation *shoningras* (Romani *šun-* 'to hear', GEN.PL derivation *-engr-*), in all likelihood modelled on Cant *hearing chetes*. In some families, *linda* is retained for 'mirror' (cf. Kelderash Romani *glinda*, Latvian Romani *gendalos*, North Russian Romani *gindalo*, East Slovak Romani *gendalos*, German Romani *glenderi*, all from Balkan-Romance *oglindă*, cf. Balkan-Slavic *ogledalo*), while in others we find the metaphorical creation *dikamengri* (from *dikh-* 'to see'). Less common objects of reference are even more likely to display shared forms (cf. individual composition *duvla's pani* lit. 'God's water', for 'rainbow'). Variation in Angloromani thus reflects not just the family-anchored use of the language as an intimate in-group code, but also the family-internal creativity that it is at the heart of its retention as a symbol of clan and group identity.

4.3 The impact of language contact

Although languages can become endangered without language contact, through dwindling group numbers in extremely isolated populations, or through genocide of isolated populations, for example, language shift is the most common process of language death. This of course presupposes language contact. In fact, researchers are sometimes inclined to see a continuum of structural borrowing leading ultimately to the replacement of most of the grammar and lexicon and so to language shift (cf. Thomason and Kaufman 1988).

Arabic structural borrowing into Domari is massive. The two languages share most syntactic patterns, intontation, and a fair amount of phonology, and for a great number of grammatical categories Domari relies almost entirely on Arabic loan forms. These include prepositions (2), negation markers (3), aspectual auxiliaries including their Arabic inflection (4-5), modal auxiliaries including their Arabic inflection (6-7), adverbial subordinators and the structure of adverbial clauses (8-10), coordinating conjunctions, indefinites, focus particles, and discourse markers (11), relative and resumptive pronouns and their inflection (12), and more:

(2)

  min   *kury-a-ki*

  from   house-OBL-ABL

'from the house'

(3)

  pandžī   **mišš̌**   *bizzot-ēk*

  3SG   NEG   poor-PRED.M

'He is not poor'

(4)

  kunt   *aw-ami*

  was.1SG   come-1SG

'I used to come'

(5)

  baqēt   *kamk-ami*

  continued.1SG   work-1SG

'I continued to work'

(6)

  ṣārat   *mangišk-ari*

  began.3SG.F   beg-3SG

'She began to beg'

(7)

bidd-ī   *dža-m*   *kury-a-ta*

want-1SG   go-1SG.SUBJ   house-OBL-DAT

'I want to go home'

(8)

qabel   mā   *dža-m*   *xałḷaṣk-ed-om*   *kam-as*

before   COMP   go-1SG.SUBJ   finish-PAST-1SG   work-OBL

'Before I left I finished my work'

(9)

na   kil-d-om   *bara*   **liʔann-hā**   *wars-ari*

NEG   go.out-PAST-1SG   out   because-3SG.F   rain-3SG

'I did not go out because it is raining'

(10)

iza   *wars-ari*   *n-aw-am-eʔ*

if   rain-3SG   NEG-come-1SG-NEG

'If it rains, I shall not come'

(11)

ū   daʔiman/ yaʕnī/ kun   *ama*   *kury-a-m-ēk*   **wala**

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

kil-šami   **wala**   *aw-ami.*

exit-1SG   and.not   come-1SG

'And I was always/ I mean/ at home, not going out nor coming'

(12)

mana	*illi*	to-r-im	*iyyā-h*
bread	REL	gave-2SG-1SG	RES-3SG

'the bread you gave me'

In fact, it is possible to identify just a core set of grammatical relations on which Arabic has had little or no impact. It consists of deictic and anaphoric personal pronouns (with the exception of the resumptive pronoun) and bound pronouns (with the exception of those that accompany complementisers), synthetic case markers, bound person concord marker on the verb (with the exception of Arabic marker on Arabic-derived modal verbs), bound tense-aspect and modality markers (with the exception of inflected Arabic modal verbs and auxiliaries), and one converbal constructions.

This is clearly, by any measure, a case of heavy or massive structural borrowing. But is language contact responsible for language attrition and ongoing language death? As far as we can tell, borrowing on this scale has been present in the language for the past two hundred years, as it is attested already in early 19th-century sources. There is no reason to assume any significant decline in the number of speakers since that time, until the 1960s. With the exception of parts of the lexicon and lexico-grammatical domain, where words have been lost, for instance, Macalister (1914) lists Indic numerals for the full set, while present-day speakers tend to know only lower numerals, it appears that language contact has been stable in Domari, and has not led either to massive language attrition, or to language loss. The causes for language loss are external to the linguistic system itself.

In Angloromani, itself a product to some extent of attrition and language loss, language contact is in fact responsible for the creation of the idiom as a mixed system that relies on components from both the old language, Romani, and the surrounding language, English. The particular patterns in which material from the source languages is combined belong to the question of structural composition and etymological compartmentalisation in Mixed Languages, which has already been addressed above. It is sufficient to remark here that Angloromani (and other Mixed Languages) provide a useful indicator for the actual structural definition of 'shift': In Angloromani, finite verb inflection comes from English (with the exception of very few fossilised forms, cf. *shom shilo* 'I-am cold', but *lesti shilo* 'he [is] cold'), even if it does not correspond to standard English usage (cf. *mandi's pooked* 'I've said'). The absence of productive finite-verb inflection means that the predication is not initiated in Romani, but in English. The utterance thus becomes an English utterance, in which Romani structures are hosted. Irrespective of the continuation of structures from (inflected) Romani, the shift in the construction of the predication justifies the definition of Angloromani in historical perspective as a shift away from Romani and into English, with accompanying selective replication of structural material from Romani.

4.4 Sociolinguistic dimensions

By definition, endangered languages share a unique sociolinguistic condition, that of endangerment. This condition can be broken down to a series of factors, both internal and external. The challenge in the study of endangered languages is to examine the extent to which similarities and differences exist between types or clusters of factors. Without attempting any comprehensive survey, we might list as external factors:

(1) the power of the language - the ability of speakers to constitute social units that exercise powers over their own affairs and possibly also the affairs of their neighbours, and

(2) links between the language and socio-economic opportunities; and as internal factors:

(3) the role of language in cementing and manifesting group-identity,

(4) the prevailing attitudes to language (resulting partly from that, and from the other factors).

Limiting ourselves here to those cases in which language death is triggered by language shift (gradual shift to another language, rather than by sudden disappearance of the population of speakers), and assuming that language shift involves a change in the balance of domain-particular use of languages in the multilingual group's repertoire, we are seeking to identify the causes of language abandonment and so of endangerment in the social changes that have brought about a shift in the balance of domain-particular language use.

    As indicated above, the Dom have been bilingual for many centuries, in fact, as far back as their history in the Middle East can be reconstructed. They have always been in a social position that was subordinated to and dependent on their Arab neighbours, and there was never a link between group membership and power over the group's destiny, save in the sense that nomadism allowed the group in the past to avoid danger and dominance by others to some extent. The chain of events that led to language abandonment is related strictly to the change in the group's socio-economic profile, and so the loss of its identity as an economic unit. It begins with the shift in the occupation of the men, from metalwork, to paid employment, accompanied by the move from tents to permanent accommodation. The generation that was born after the disappearance of the nomadic, family-oriented, tent-based service economy, in the 1940s, was also the generation that abandoned the language. The clear cut-off point, however, was in the late 1960s, with the introduction of compulsory primary school education. At this point, the first generation of Dom children, both girls and boys, attended school, alongside their Arab neighbours. By this time the community was simply another clan of poor urban workers. It was this generation of children, growing up in the 1960s, who were no longer spoken to in Domari. Language loss must be regarded not so much as a conscious attempt to integrate into the Arab

surroundings, however, but rather as the loss of the symbol of a self-contained socio-economic community living apart from mainstream society and not participating in any of its activities apart from providing it with specialised services. This, the mobile service economy, was the framework in which Dom traditions and values, including language, were functional and sustainable. With the decline of the mobile service economy, group-identity is reduced to merely an origin, and primarily a stigma, and language is abandoned as it no longer flags any operational network of skills, survival, or social cohesion. The remaining speakers are the sole survivors of a socialisation in an earlier period, that of the tent-dwelling, mobile service-economy.

The case of Romani in Britain is quite different. Much like the Middle Eastern Dom, the British Roma constituted a small, mobile and isolated community, specialising in specific trades, between the 16th and 19th century. There is however indirect evidence of linguistic assimilation or at least of contacts between Roma and other, non-Romani travelling groups as early as the 17th century (cf. Bakker 2002, who discusses a Para-Romani wordlist). Certainly within the past century, English and more recently also Welsh Romanies have intermarried widely with Travellers of indigenous (non-Romani) origin. It is noteworthy that these indigenous groups had in-groups codes of their own: English Cant, Scottish Traveller Cant, Irish Traveller Cant or Gammon (also known as Shelta). Unlike inflected Romani, these are not all-purpose everyday languages, but, much like Angloromani, primarily lexical-stylistic registers employed in selective utterances in the discourse. They are, however, the linguistic expression of group-identity in these contexts. British Romanies appear to have undergone a process of accommodation and even partial assimilation to the indigenous travelling communities. Although demarcation lines between English Romanies and Irish and Scottish Travellers remain clear to members of the groups, despite intermarriage, in England and Wales there remains to distinction between Romani and non-Romani Travellers, possibly hinting at an amalgamation of distinct populations into one group, with the retention of clan divisions within it.

With the increasing immersion with Travellers of non-Romani origin, inflected Romani as an everyday family or community language began to lose its functionality. However, since the overall character of the population as a distinct, mobile, family-based economy and with it its values and traditions remained intact, possibly integrating values of the various groups (as is the case in present-day cases of intermarriage of Irish Travellers and English Romanichals or Welsh Roms or Kååle), there remained a functional slot for an in-group code, consisting of primarily lexicon and some stylistic features. Romani was re-functionalised to fill this particular slot. Grammar and with it the ability to form predications (and so utterances) distinctly in Romani was abandoned, but lexicon and creative processes of lexical formation were retained, selectively, in order to maintain a linguistic expression of in-group identity, and as a means of communicative demarcation of

insiders and outsiders, much like the function of Cant in the other travelling populations.[7] Historically, then, Romani has been reduced functionally from an everyday family language to just an expressive, pragmatic marker of group identity; this has carried with it a structural reduction of grammatical inflection and selective retention and replication primarily of lexicon. Contrasting with the case of Domari, group-identity has not been lost, and the maintenance of group-identity is accompanied by the selective maintenance of an in-group linguistic code, a kind of linguistic life after death. The Dom, by contrast, have not developed a 'Para-Domari' as an in-group code, presumably due to the almost complete collapse of social organisation, economic structures and positive awareness of group identity.

5. Community activism and the role of the linguist

I return now to the issue addressed at the beginning of this paper, and that is the role of the linguist investigating endangered languages. Let us begin with the role of the linguist working among the Dom of Jerusalem. Here, we find a situation in which ongoing language death is accepted by the speakers and the community. In conversations and interviews, speakers show no passion toward the preservation or even documentation of their language, and in daily life they show no effort or initiative to teach it to the younger generations. There are also no community institutions that could provide such an organised opportunity, or even inspire individuals to act or even to reflect upon the ongoing disappearance of the language.

Activism within the community consists of two roles for individuals. The first is the traditional *Mukhtar* or community leader, an office that is in part hereditary (the present Mukhtar is the son of the previous one), in part elected (his right to succeed his father was confirmed by an assembly of the community elders), and in part appointed (he is officially recognised by the Israeli authorities, and was before that officially recognised by the Jordanian authorities, through a letter of appointment). In the absence of any distinct traditions or community-internal organisations or ceremonies, the role of the Mukhtar is reduced to that of occasional arbitrator in cases of community-internal feuds, or, more often even, a go-between between the authorities and members of the community. Born in 1933, the present Mukhtar is a fluent speaker of Domari, and was one of the author's principal language consultants. After co-operating with the author during visits to the community over a period of more than three years, he began imitating the structure of language elicitation sessions by writing down wordlists, phrases and improvised inflection paradigms in Domari in Arabic script, asking the author (as well as, apparently, other visitors to the community, including journalists) whether his notes might be published abroad and whether they could earn him some money.

---

[7] In fact, in casual conversation Irish and Scottish Travellers occasionally refer to their own respective forms of Cant as "our Romani", re-interpreting the term 'Romani' as a generic terms indicating 'the in-group lexicon of a travelling population'.

The Mukhtar's political challenger in the community is a young woman, born in the early 1970s, who is not a speaker of Domari. She is, however, fluent in English, and has since 1995 established links with non-governmental organisations abroad and in Israel, and frequently hosts visitors to the community, including political activists, Christian missionaries, and journalists. Around 1998 she founded a Dom association, with the help of a left-wing Israeli political party, and established a web page which has since become one of the principal sources of information on the community for the outside world, though its content is limited primarily to declarations of goals (such as social reform and provisions for education within the community). The association did not, at least between the years 1998-2002, seem to have any real support from the community, save the activist's own siblings, who were listed in the organisation's literature as holding various offices within the society's board of directors.

Like all other members of the community, this young activist and her family are Muslims. However, much of the support for the activities of her association comes from a small organisation operated from Cyprus by Baptists Christians from North America, under the name 'Dom Research Centre'. The Dom Research Centre devotes its website (www.domresearchcenter.com) and publications to textual and visual documentation of the Dom communities – both in Jerusalem and elsewhere in the Middle East. There is no trace of any missionary content in its official publications, and its members deny any religious orientation of their organisation. The present author has, however, heard repeated confirmation from activists who are closely associated with the Dom Research Centre that it is operated by the Wycliffe Foundation and the Co-operative Baptists Fellowship.[8]

Thus, the linguist approaching the Dom community in the mid- and late 1990s in order to document the Domari language, is confronted with a situation of a very small group of speakers, largely elderly and illiterate, who are not concerned about the ongoing disappearance of their language, and who appear to have no strong collective desire either to maintain or document the language or to preserve or flag any aspect of their group-internal traditions or values. Rather, their primary concerns are with day-to-day issues of survival, health, and personal developments in their immediate neighbourhood. The Mukhtar shows a faint interest in the documentation work, motivated partly by curiosity, partly by the honour of having an opportunity, though the interviews, to display his role as a prominent community member, and partly by a remote dream of a possible financial gain should his own 'publication' about the language receive an audience abroad. This interest in cooperating is often, however, put under the condition of exclusivity, in an attempt to

---

[8] In personal communication, Donald Kenrick (1999), Thomas Acton (2004), as well as Keith Holmes of the Co-operative Baptist Fellowship (2004). It is not unusual for missionary sects to conduct 'covert' religious activities, posing as research enterprises. There is no mention of any missionary background on the web pages or printed information of the Summer Institute of Linguistics (SIL). Everett (2003: 144-145), however, who admits to having been a member of SIL, writes: "it is nevertheless true that the majority of missionary linguists are motivated mainly by the desire to 'convert' indigenous peoples to a fundamentalist form [of] Christianity".

prevent the linguist from speaking to any other members of the community. Then there is the young activist, who is prepared, similarly on condition of exclusivity, to mediate between the linguist and speakers of the language. Not being a speaker, she is not, however, in a position to help directly. Moreover, keen to secure financial support for her activities from the missionaries, and so to secure her position as a leader of a viable community initiatives, she prefers to minimise her involvement with any alternative projects. She even rejects the linguist's offer to assist in the documentation of Domari-language stories and to help produce instruction materials in Domari, claiming that such activities will be carried out in collaboration with the Dom Research Centre.

Who, then, represents the community toward the linguist? Is it the Mukhtar, the 'officially acknowledged' leader who, however, does not seem to have any active role in the community, and certainly does not engage in any community-wide activity? Is it the young activist, who is not a speaker, but who seeks to control all contacts between members of the community and outsiders, and whose considerations in respect of academic co-operation are essentially political, based on a need to secure a financial basis for her own activities? And who are the potential co-operation partners for a community-based project? Are those the speakers, who lack any motivation or inspiration to document their language (although they are willing to co-operate with the researcher in interviews)? Or the younger members of the community, who are concerned mainly with day-to-day survival? Under the present circumstances, there is no interest on the part of speakers in language-oriented activities, no political benefit to the activist to engage in collaboration with the linguist, and no financial gain to the traditional leader in intensifying the collaboration beyond the interviews. There is, in short, little that the linguist can do beyond the strictly academic task of documenting and analysing the language through interviews with the remaining elderly speakers.

Curiously, a different kind of situation present itself to the same linguist among users of Angloromani in England. Increased interest in the Romani language is motivated by at least three different factors. The first is an intellectual curiosity, sparked in part perhaps through knowledge about ongoing scholarly activities through the interface between scholarly research and political activism, going back to the late 1960s, as well as the presence of book stalls, some of them displaying popular and scholarly research works on Gypsy culture and language, at fairs and other events frequented by Travellers. This pathway is particularly hard to assess; nonetheless, we have evidence that casual exposure to research literature on Gypsies has led individual Romanies to contact linguists and ask for sources of information on their language – contact which then marked a common interest in 'discovering' the language, and gave rise to occasional co-operation in its documentation. A more powerful motivation on the part of English and Welsh Gypsies to 'explore' the Romani language results from encounters between British Romanies and Romani immigrants to the UK from central and eastern Europe. Such encounters normally begin as economic enterprises: immigrants are interested in finding jobs, but are hesitant to expose themselves to exploitation (especially when their legal status in the country is

ambiguous, e.g. if they lack employment permits). Working for other Gypsies seems a safe solution. Local Romani families in turn often need additional workers for seasonal enterprises. Such business partnerships give rise to a practical need to communicate, in which the British Romanies try to capitalise on their knowledge of Romani-derived lexicon. They soon discover that this is a useful tool in communication, but that it is limited, and that they are unable to actually converse with the immigrants. Realising that the British form of Romani is tightly embedded into English, they regard the European dialects of Romani as prestigious, due to their function as full-fledged languages. Curiosity and practical communication needs are thus coupled with the prestige of learning the more 'original' or 'proper' Romani language.

Finally, large parts of the British Romani community are involved in an international Romani pentecostal church. Leaders of the organisation have close ties with Romani communities in other European countries as well as North and South America, and are keen to 'adopt' communities abroad, as well as to integrate Romani immigrants to Britain into the local churches. The church provides a closed and protected space in which Romani cultural identity can be re-invented, still run by Romanies, still subjected to Romani community organisation structures and value, but at the same time autonomous to act overtly, free of the limitations that are often imposed on public movement and appearances of Gypsies.

In these contexts, a kind of individual academic curiosity, practical communication needs, and the search for new and international symbols of Romani unity and identity have prompted British Romanies to explore their own language and its relation to the inflected Romani dialects used in other countries. Some have taught themselves to speak inflected Romani, through close associations with immigrants or colleagues abroad or through periods of work in Romani communities abroad; the author knows of at least twenty such speakers, and the actual number is likely to be much higher. Some have contacted me to seek advice on how to learn inflected Romani. These contacts have resulted in the production of a series of tapes, of which hundreds of copies have been distributed by one of the Romani missionary organisations, providing instruction in inflected Romani grammar. They have also resulted in the launching of a documentation project of Angloromani, with the aim of producing educational materials to promote awareness of Romani culture and language among Romani and non-Romani pupils in the Lancashire and Greater Manchester areas.

Having been, in a sense, re-cycled as a discourse-level register whose primary function is pragmatic identity-flagging, and so having retained a reduced function after language death, Romani is now being revitalised in numerous communities in Britain as the language of interaction with foreign Romanies and a symbol of international bonds with a wider Romani community. Time will tell whether this revitalisation process will become more widespread and lead to a genuinely popular acquisition process of Romani in Britain, and, if it does, whether inflected Romani will replace Anglromani as an identity symbol, or

whether it will, rather, occupy a place alongside Angloromani as a language reserved for transactions with immigrant and foreign Roma, for 'international' Romani contexts, so to speak. Linguists are, in any event, currently able to witness a unique and exceptional process, namely the beginning of a process leading to possible language revitalisation.

6. Conclusion

Linguists have a right and a duty to be interested in endangered languages. Even the two modest examples discussed in this paper show structural features and even 'system types' that are rare among the languages of the world and cannot be observed strictly within the sample of viable, 'larger' and more stable languages. Moreover, endangered languages constitute a unique sample in their own right in respect of types of individual variation, and in respect of the factors that support or impede language maintenance. They often also constitute valuable case-studies in language contact, illustrating, among other things, that 'extreme' contact can be prolonged, that it need not be the cause of language death, and that it can even, in fact, be the trigger for language birth. All this is sufficient to justify an academic interest in the study of endangered languages. And if the sample of endangered languages is dwindling, then the documentation of endangered languages is a matter of academic urgency. It is correct, in this respect, to make the documentation of endangered languages a matter of high priority.

But in doing so, we as linguists must be open about our motivations to study endangered languages. Alongside our academic duty, we may or may not also assume active roles in the communities. But to link academic documentation with activism on behalf of the community is not a productive approach to the tasks and duties of linguists. Not every community is able to or interested in 'saving' or even documenting its language. Indeed, not every community is able to act as a collective on such issues. If community activism and consensus within the community were made pre-conditions for the documentation of Domari, then we would not have insights into the specific process of ongoing language death in this community, nor into some of the unique contact-related structures of this language. It would eventually disappear without leaving a trace in our sample of documented languages – merely due to the expectation that linguists should adopt the role of heroes to the community of speakers, a role that speakers of Domari are not willing to entertain (even if the linguist were willing to act it out). On the other hand, even the less pretentious linguist might occasionally be rewarded with the privilege of advising speakers on how to re-discover their linguistic heritage, and even witness the rare initiation of language revitalisation efforts at the spontaneous level of the (largely non-literate) community. In training young researchers to engage in the documentation of endangered languages, we must therefore avoid any pressure on them to become community 'heroes', and instead prepare them for the potential pitfalls, but also for the possible rewards that are associated with the urgent task of securing a diverse linguistic sample corpus for the sake of future generations of students of language.

7. References

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.

Bakker, Peter 2002. *An early vocabulary of British Romani 1616: A linguistic analysis. Romani Studies series 5*, Volume 12:2, 75-101.

Bakker, Peter, and Marcel Cortiade (eds.) 1991. *In the margin of Romani. Gypsy languages in contact.* Amsterdam: Institute for General Linguistics.

Bakker, Peter and Maarten Mous (eds.) 1994. *Mixed languages: 15 case studies in language intertwining*. Amsterdam: IFOTT.

Cortiade, Marcel 1991. Romani versus Para-Romani. In: Bakker, Peter, and Cortiade, Marcel. eds. *In the margin of Romani. Gypsy languages in contact*. Amsterdam: Institute for General Linguistics 1-15.

Crystal, David 2000. *Language death.* Cambridge: Cambridge University Press.

Dorian, Nancy 1981. *Language death: the life cycle of a Scottish Gaelic dialect.* Philadelphia: University of Pennsylvania Press.

Dorian, Nancy. 2001. *Surprises in Sutherland: linguistic variability amidst social uniformity.* In: Paul Newman, and Martha Ratliff (eds.) Linguistic fieldwork, 133-151. Cambridge: Cambridge University Press.

Elšík, Viktor and Yaron Matras 2005. *Markedness and language change: The Romani sample*. Berlin: Mouton de Gruyter.

Everett, Daniel L 2003. Documenting languages: A view from the Brazilian Amazon. In: Peter K. Austin (ed.) *Language documentation and description,* Volume 1, 140-158. London: Hans Rausing Endangered Languages Project.

Grinevald, Colette 2003. Speakers and documentation of endangered languages. In: Peter K. Austin (ed.) *Language documentation and description*, Volume 1, 52-72. London: Hans Rausing Endangered Languages Project.

Hancock, Ian F 1970. Is Anglo-Romanes a creole? *Journal of the Gypsy Lore Society*, third series, 49: 41-44.

Hancock, Ian F 1984. Romani and Angloromani. In: Peter Trudgill (ed.) *Languages in the British Isles*, 367-383.

Labov, William 1970. *The Study of Nonstandard English.* Champaign: NCTE.

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.

Maffi, Luisa 1996. Endangered Languages, endangered knowledge, endangered environments. Position paper for the Berkeley Conference. http://ucjeps.berkeley.edu/Endangered_Lang_Conf/Endangered_Lang.html.

Matras, Yaron. (ed.) 1998. *The Romani element in non-standard speech*. Wiesbaden: Harrassowitz.

Matras, Yaron 1999. The state of present-day Domari in Jerusalem. *Mediterranean Language Review* 11, 1-58

Matras, Yaron 2000. Two Domari legends about the origin of the Doms. *Romani Studies*, 5th series, 10, 53-79.

Matras, Yaron. 2003a. Defining typical features of minor languages (with special reference to Domari. In: Sherzer, Joel and Stolz, Thomas (eds.). *Minor Languages. Approaches, definitions, controversies.* Bochum: Brockmeyer. 1-14.

Matras, Yaron 2003b. *Mixed languages: re-examining the structural prototype*. In: Yaron Matras, and Peter Bakker. (eds.) The mixed language debate, 151-176. Berlin: Mouton de Gruyter.

Matras, Yaron 2005. The full extent of fusion: A test case for connectivity and language contact. In: Walter Bisang, Thomas Bierschenk, Detlev Kreikenbom und Ursula Verhoeven. eds. *Kulturelle und sprachliche Kontakte: Prozesse des Wandels in historischen Spannungsfeldern Nordostafrikas/Westasiens*. Akten zum 2. Symposium des SFB 295, 241-255. Würzburg: Ergon Verlag.

Matras, Yaron and Peter Bakker. (eds.) 2003a. *The mixed language debate. Theoretical and empirical advances.* Berlin: Mouton de Gruyter.

Matras, Yaron and Bakker, Peter 2003b. The study of Mixed Languages. In: Yaron Matras and Peter Bakker (eds.). *The mixed language debate. Theoretical and empirical advances.* 1-20. Berlin: Mouton de Gruyter.

Meyer, Frank 1994. Dōm und Turmān in Stadt und Land Damaskus. Erlangen: Fränkische Geographische Gesellschaft.

Nettle, Daniel and Suzanne Romaine 2000. Vanishing voices. The extinction of the world's languages. Oxford: Oxford University Press.

Okely, Judith 1983. The Traveller-Gypsies. Cambridge: Cambridge University Press.

Sampson, John 1926. The dialect of the Gypsies of Wales, being the older form of British Romani preserved in the speech of the clan of Abram Wood. Oxford: Clarendon Press.

Smart, Bath C. and Henry Thomas Crofton 1875. The dialect of the English Gypsies. London: Asher, and Co.

Thomason, Sarah G. and Terence Kaufmann 1988. Language contact, creolisation and genetic linguistics. Berkeley: University of Calfornia Press.