MODELLING MIXED LANGUAGES: SOME REMARKS ON THE CASE OF OLD HELSINKI SLANG¹

Merlijn de Smit
Stockholm University

Abstract

Old Helsinki Slang (OHS) a linguistic variety spoken in the working-class quarters of Helsinki from approx. 1900 to 1945, is marked by the usage of a virtually wholly Swedish vocabulary in a Finnish morphosyntactic framework. It has recently been subject of two interestingly contrasting treatments by Petri Kallio and Vesa Jarva. Kallio argues that the morphosyntactic base of OHS gives cause to analyzing it as unambiguously Finnic, and therefore Uralic, from a genetic perspective, whereas Jarva, drawing attention to the possible origins of OHS in frequent code-switching, believes it deserves consideration as a mixed language alongside such cases as Ma'a and Media Lengua.

The contrasting approaches of the two authors involve contrasting presuppositions which deserve to be spelled out: should the genetic origin of a language be based on the pedigree of its structure (with mixed structures pointing to a mixed genetic origin) or on the sociolinguistic history of its speakers? Taking the latter course, I argue that the most valid model for the emergence of genetically mixed languages is the code-switching one proposed by Peter Auer. Measuring OHS against Auer's model, however, it is a marginal case for a mixed language, particularly as Auer's and similar models imply some composition in structural domains, which seems wholly absent in the case of OHS. Thus OHS is not a genetically mixed language, even if it may have developed as one, had early OHS taken a different course as it eventually did.

1. Introduction

The question of mixed languages—languages with more than one ancestor in the genetic sense—has risen anew lately in general historical linguistics (Thomason & Kaufman 1988; Mufwene 2001:19) and in Uralic linguistics as well (Raukko & Östman 1994, 1995; Künnap 1998:21-39; Esa Itkonen 1998). The question itself is anything but new, as the rise and fall of Marr's ideas in the Soviet Union testifies. In recent discussions concerning possible mixed

¹ My thanks go to Petri Kallio and Vesa Jarva for their comments on a draft version of this paper.

languages, there appears to be little consensus on either models for their emergence (Thomason 2003:27) or on their structural features (Matras 2000:79). Not all suggested models necessarily involve mixing in the genetic sense (Thomason 1997a:450) and the existence of genetically mixed languages itself is still contested (Greenberg 1999). Part of the reason for this lack of consensus is, I believe, that the question of genetic mixing directly confronts us with the need to model language as a historical entity to a degree not required by the traditional models of language divergence.

The theoretical issues concerning genetic mixing are exemplified by two recent papers offering contrasting approaches to Old Helsinki Slang (henceforth OHS), namely that of Kallio (2007) and Jarva (2008). Kallio, basing himself on an analysis of OHS core grammar as unambiguously Uralic, does not believe OHS to be a mixed language, whereas Jarva, drawing attention to the role of code-switching in the emergence of OHS, is more sympathetic to an analysis of OHS as genetically mixed. The purpose of this paper is to tease out some of the theoretical issues raised by these two contrasting approaches. Below, I will first argue what an adequate model of genetic mixing should be. I will then discuss some problematic consequences of one of these requirements (namely, that an approach to genetic mixing should be consistently realistic with regards to language as a historical entity). One of these consequences is that the genetic model itself may have a more or less restricted applicability in describing the evolution of languages. After this, I will discuss the models of genetic mixing which I consider to be the most promising: namely, the code-switching models proposed by Auer (1999) and Myers-Scotton (2003). Measured against these models, however, OHS is not a mixed language. Thus, while I am in agreement with Jarva's (2008) approach with its attention to the sociohistorical circumstances surrounding the genesis of OHS and the role of code-switching, my conclusion is the same as Kallio's (2007).

2. Mixed languages: looking for a model

The question of genetic mixing itself is posed by a number of languages which, aside from pidgins and creoles, seem to confound historical reconstruction. So-called intertwined languages (Bakker & Muysken 1994) show whole subsystems (such as the lexicon on the one hand and bound morphology on the other) with various genetic origins: Michif, Ma'a, Media Lengua and Angloromani are cases noted in the literature. Michif, for example, combines Cree verb morphology and French nominal morphology; Media Lengua has a largely Quechua grammar and Spanish vocabulary. Ma'a, spoken in Tanzania, has a non-Bantu core vocabulary (probably relatable to Cushitic) with a largely Bantu morphology (Mous 1994, Thomason 1997b). The challenge posed by mixed languages is that determining whether they are indeed genetically mixed, or whether they should rather be described as having a singular ancestry obscured by massive borrowing, directly involves our notions of what a language is in a genetic sense. Is the history of a language the history of its usage, or the history of its structure—reflected in individual instances of usage, or individual idiolects? Either answer may lead to different points of view on the historical emergence of languages such as Ma'a.

Below, I will discuss two approaches to genetic linguistics which I believe to be mistaken. The first one, which attempts to specify a 'minimal unit' of linguistic change (the idiolect, or the utterance) and regard genetic lineages, mixing, etc. as emergent from variation of these minimal units, runs the risk of making the distinction between genetic mixing and language contact impossible to state. The second one, departing from structure and defining genetic relationships as dependent on correspondences in structure, risks implicitly or explicitly making a category mistake in substituting abstract, a-historical structures for language as a historical entity. Instead, I suggest that any model of genetic mixing must allow for both the 'ideal' aspects of language (language as a social system of norms stretching across various generations of speakers) and the notion that the historical aspect of language is rooted in the usage of language by a specific speech-community. These lead to some specific demands that one may make of any model of genetic language mixing.

2.1. The persistent aspects of language

Language is one of those objects of our human world which, like persons and cultural practices, change and yet stay the same. This is what grants those objects their historicity: abstract objects such as numbers neither ‘change’ nor ‘persist’, rather, temporal attributes do not even apply to them. And the smaller building-blocks of the physical world are subject to substitution, a popping in and out of existence, rather than change. With languages, this property of being ‘smeared-out in time’ is even more problematic than with biological organisms, in that we cannot speak of discrete ancestors and descendants: ‘Finnish’ as a language itself involves variation in both time and place, and the distinction between various ancestral stages of the language is notoriously hard to draw. A child acquiring a language will acquire the language spoken by the speech-community he is born in by virtue of the social, normative nature of language itself. Intergenerational differences may occur, and in some cases, notably related to language death, these differences may be radical (the differences in syntax, morphology and phonology between traditional Dyirbal and Young People’s Dyirbal being one example (McMahon 1994:305-307), and that between the various generational varieties of another Australian language, Tiwi, being another (McConvell 2008:198)). Yet communicative capabilities across generations, guaranteed by a shared set of linguistic norms, will generally be preserved. Thus children of bilingual parents will either learn one of their parent’s languages, or both, but not a mixture of both. The challenge for models with incorporate genetic mixing is thus to allow for a mixing process gradual enough, or a period of coexistence between the two languages long enough, to not imperil communication across generations.

This may seem to be begging the question, as one of the models argued for in the emergence of mixed languages, namely that of intertwining (Bakker 2003:139-140), argues for precisely a sudden emergence: children are argued to combine the grammatical system of one of their parents’ languages (typically the mother’s) with the lexicon of the other in the context of the emergence of a new, mixed ethnic group. I would argue nonetheless that in such a scenario, the continuity of language as a system of norms stretching across the generations would be severed and the mixed language would arise non-genetically: it would be ‘mixed’ only in the make-up of its structure, but not in terms of genetic origins. Without restricting the notion of genetic descent to diachronically continuous languages, we end up, as I will argue below, with a model in which contact-induced change and genetic descent can no longer be distinguished.

The closest approach to a working model involving genetic mixing in Uralic linguistics has been provided by Raukko & Östman (1994, 1995). Their model leads, however, to any contact-induced innovation to constitute genetic language mixing, focusing, as it does, on the details of linguistic usage as the cost of language as a persistent entity (Laakso 1995:70). The same goes in lesser degree for some other recent models which allow for language mixing: both Croft (2000, 2006) and Mufwene (2001) are seduced by the superficial similarities between linguistic and biological evolution, and seek to specify a ‘minimal unit’ of linguistic change analogous to genes: just as the evolution of species may be seen as emergent on the variation in the gene pool, the development of languages from ancestral forms may be seen as emergent on the variation of such ‘minimal units’. For Mufwene (2001:2), the idiolect is the minimal unit; for Croft (2000:26) it is the utterance. Seeing linguistic change as historical variation between unchanging idiolects is not new, of course: it was a mainstay of the attempts to have one’s generativist pie (an unchanging grammatical competence situated in the individual speaker’s brain) and historically eat it. The problem with specifying the idiolect as some kind of minimal unit is the same with the viewpoint that regarded structural change as occurring mainly in the context of language acquisition: it ignores that an individual innovation becomes a *change* only when it spreads through a speech-community, and as Aitchison (2001:209) points out, young children do not form the social networks needed for innovations to be propagated as change.

Croft's more fine-grained notion of the utterance as minimal unit suffers from a similar defect: the equivalents of genes are linguistic variants, "linguemes", which are replicated in utterances, sometimes in a slightly altered way, which may then propagate across a speech-community. A language, then, is defined as a historically bounded population of utterances (Croft 2006:106). This view tends to reduce our notion of language to usage, or to knowledge of language, and to leave aside the ideal, structural domain immanent in the conventional, social side of language—a reductionism which occurs more widely in the cognitivist/constructionalist paradigm (for example, Bybee 2006:730). Yet, abstract structure is necessary to account for the production and interpretation of utterances, not to speak of such notions as reanalysis and analogy in language change. As many applications of evolutionary models in the human sphere, Croft's is explicitly anti-teleological (Croft 2000:66-71, 2006:119) but this is a disadvantage, not an advantage of these models. Teleology, as the role of abstract objects (not future concrete objects) in guiding the development and ordering of concrete objects is indispensable in making sense of such mechanisms of change as reanalysis and analogy (Andersen 2006:76-77).

In terms of their accounting of genetic mixing, both Mufwene's and Croft's proposals have the same problems as Raukko & Östman's. For example, Croft (2006:113-114) draws an analogy between hybridization in plant genetics and interference in language. This potentially extends 'genetic mixing' to all possible contact-induced changes, leading to a hopelessly complicated model. The same may be said for Mufwene's (2001:126, 143-144) model. In both cases, the focus on a 'minimal unit' of change has been too relentless, and in itself justified criticism directed at regarding linguistic structures, rather than speakers as the agents of change has extended into a disregard for the ideal side of language as a whole: of language as a 'persistent entity'. But without this notion of language, genetic transmission can not be comprehensibly stated. Models such as Croft's lead to a conception of language as an aggregate of its components (idiolects, or utterances) in the same way that a species is seen as a population of individuals. However, there is no role for species or "essence" as such in the mutation and propagation of genes, whereas with language change this role is indispensable: reanalysis, extension and analogy all refer to rational operations on structure as a whole, as does markedness as an organizing principle in language structure and linguistic interpretation (Andersen 2006:77 ; Shapiro 1983:18, 74-75).

Significantly, in his treatment of mixed languages, Croft (2003:49), argues for a more-or-less *ad hoc* distinction between normal genetic transmission and genetic mixing:

[...] every language has a small proportion of vocabulary from parents other than its normally recognized ancestor language. This is technically a mixed language in our terms; the mixture arose by the processes described above. However, the term 'mixed language' is commonly construed very narrowly to languages in which significant proportions of linguemes come from more than one parent language. This definition, while not precise, narrows the field of "really" mixed languages dramatically.

The same problem of demarcation also appears, however, in treatments of mixed languages which focus on synchronic structure, rather than on speech events analogous to genes. Thus Matras (2000:79) argues that "*most languages are to some extent at least "mixed", in the sense that they have components that can be traced back to more than one source language as a result of a situation of contact in the language's earlier history*", so that "*when languages are referred to in the literature explicitly as "mixed" [...] it is presumably to highlight that they go beyond the commonly attested patterns of mixture.*" Greenberg (1999:351-352) correctly criticized this notion of mixture as synchronic, not historical, in nature. If there is no conceptually necessary (not just an empirical) link between a language's history and a language's synchronic structural make-up (and I will argue below there is none), one cannot define genetic mixing on the basis of the presence of various structures in the synchronic state of any language. Models of linguistic evolution which (in analogy to biological evolution) seek to specify a minimal unit of change but deny the role of the

normative, ideal aspects of language fall in the same trap as models which tend to hypostasize synchronic structure at the cost of linguistic history. We cannot account for genetic linguistics without placing into center stage those ideal, conventional aspects of language which seem to transcend history.

2.2. Realism and the limitations of the genetic model

Above, I argued that a model of genetic transmission and mixing must be based on historical process rather than synchronic structure. This is closely related to what Lass (1993) dubs realism in historical linguistics, namely, the assumption that our reconstructions and the lines of transmission we draw between reconstructed stages of a language refer to languages spoken by real, concrete speech-communities at a specific time and place in the past, and refer to a real transmission of language through various generations of a speech-community down time. This does not mean that our reconstructions, and our knowledge of past linguistic relationships, is not partial, map-like and quite possibly wrong in some places. Rather, the notion that they very well may be crucially depends on a realist assumption.

In other words, genetic relationships are not directly based on correspondences between synchronic states of languages. Rather, those correspondences point to common origins which in turn testify of genetic relationships—which exist whether we know of them or not. This notion is implicit in the practice of historical linguistics: historical linguists thus avoid reconstructing semantic anachronisms; try to keep reconstructed structures within the bound of what is typologically possible and generally attested in languages of today; and are conscious of the way the comparative method itself may produce artefacts probably not part of the real, historical languages which are the object of reconstruction.

There is a consequence to realism: it means that the importance of any particular subsystem of linguistic structure (such as bound morphology) in reconstruction is contingent, not conceptually necessary. Assuming our earliest human ancestors spoke a language, Finnish is genetically related to a language spoken in Africa more than 100,000 years ago—but, likely, total turnovers of lexical and morphological material occurred multiple times during the historical time-line between that language and modern Finnish. This affects relatability, our abilities to uncover genetic relationships—but not the genetic relationships themselves. Subdomains of language which we know to be historically more or less stable, such as bound morphology and basic lexicon, may help us uncover genetic relationships, but they do not define those genetic relationships themselves. No subdomain of language is stable in any absolute sense—bound morphological items may fuse to their stems, lose their specific grammatical meanings, disappear, or even be borrowed (Thomason & Kaufman 1988:98; Pakendorf 2009).

Hierarchies of borrowability do not hold in an absolute sense either. Aikhenvald (2003) sketches a contact situation in the Vaupés region of the Amazon, involving Tariana, a moribund Arawak language, undergoing contact-induced changes from nearby East Tucanoan language. These changes manifest themselves as restructuring in morphosyntax, however, lexical borrowing appears to be virtually absent due to strong resistance against borrowing words on the part of the speakers of Tariana. Cases in which lexical borrowing is strongly resisted, but contact-induced morphosyntax change is not, can easily be found and suggest that salience, and speakers' attitudes towards borrowing salient items such as words and morphological markers, play an important role in language-specific borrowability hierarchies (Anttila 1989:169; Thomason 2001:82-83). According to Thomason (1997b:478), the retention of Cushitic vocabulary in Ma'a against overwhelming Bantu influence on morphosyntax may be explained by cultural resistance to borrowing as well. Ma'a exists symbiotically as an in-group language with Mbugu, an unquestionably Bantu language which is used for communication with outsiders (Mous 1994:196). The sociolinguistic function of

Ma'a is precisely to be incomprehensible to outsiders, in other words, to be not Bantu (Mous 1994:199; Thomason 1997b:476-478). In other words, speakers and their attitudes, rather than structure, determines in the end what may be borrowed and what not.

This suggests, as emphasized by Croft (2000:2, 17) and Brinton & Traugott (2005:8) that we must strongly resist hypostasizing structure. Structures are abstract, they have no history; history is constituted by the immanence of various structures in concrete speech events. Genetic relationships are not contingent upon correspondences between structures such as bound morphological paradigms, but contingent on the historical networks of speech events in which those structures are represented. In other words, the history of a language is the history of a speech-community, or, more accurately, a 'community of speech events'.

This view may seem unexceptional, but it has wide-ranging consequences. It suggests, first of all, that genetic continuity is determined by the continuity of a distinct 'community of speech events', implying a distinct community of speakers, more so than the continuity of any particular linguistic structure. Consider the case of Ma'a: Bakker & Muysken (1994:49-50) argue that structural concerns should override sociohistorical ones in analyzing the genetic make-up of Ma'a:

[...] the speakers of Ma'a are ethnically and historically Cushitic. Speaking from the perspective of their origin, they took over the Bantu grammatical system. Linguistically, however, there are arguments to claim that the Cushitic lexicon was adopted into a Bantu grammatical framework.

However, if we avoid looking at structural similarities as a direct indication of genetic relationships, as I think we should, the precise genetic analysis of Ma'a depends on the social and historical circumstances of the shift to Bantu. If, as Greenberg (1999:352) and Thomason (1997b:481-482) suggest, the importation of Bantu grammatical structures was a gradual, centuries-long process, there would be strong reasons to regard Ma'a as genetically Cushitic, despite the virtually wholly Bantu grammar. This gradual Bantuization may well be seen as a process of language shift—but importantly, it did not reach its end-point, and the Cushitic core vocabulary of Ma'a remains as the one feature distinguishing Ma'a from the Bantu Mbugu language it is in close contact with. But, it is not hard to find cases in which, conversely, massive lexical borrowing has obscured genetic relationships. One such case is Bai, a Sino-Tibetan language whose precise position in that language family is a matter of controversy: Yeon-Ju & Sagart (2008:379) argue that Bai is in fact Tibeto-Burman but that centuries of lexical borrowing from Chinese have led to a situation in which 47 lexical items on the 100-word Swadesh list are borrowings from Chinese, against 12 Tibeto-Burman items. So, both vocabulary and structure may be, contingently, markers of the genetic relationships in question.

If languages that look relatively mixed, such as Ma'a or Bai, may have an unmixed genetic origin, something like the converse may also be possible in cases where a group of speakers shift to another language, importing a large number of structural features from their previous language, yet remain relatively isolated from the speech-community of the language they are shifting to. Ringe, Warnow & Taylor (2002:64-65) draw attention to cases like these, and the to the problems they pose for genetic classification. Examples they mention include the case of a northeast Midlands dialect of Middle English which is the result of imperfect learning by Scandinavian speakers and subsequent isolation, as well as Ma'a and Asia Minor Greek. Normally, historical linguists would fit linguistic varieties emerging through language shift to the standard family tree model: a daughter language would emerge from an ancestral form through shift-induced substratum influence. However, if the community of shifting speakers remains relatively isolated from the community of speakers of the language they are shifting to—and normally, of course, they would not—this becomes problematic. As mentioned above, the continuities underlying genetic lineages in languages are continuities of speech superimposed on continuous speech-communities. In the case of shift-induced linguistic varieties, there are no such continuities between the target language

shifted to and the linguistic variety eventually emerging, if the emerging speech community remains isolated; neither are there any between the community's previous language and the language shifted to, as there has been a language shift. The case would thus fit neither as an instance of linguistic divergence nor as an instance of genetic mixing.

Thus, the realist view implies cases of linguistic genesis which cannot be genetically modeled, and perhaps more widely so than the cases of non-genetic linguistic origin pinpointed by Thomason & Kaufman (1988:200-201). This should not be surprising, as languages, as distinct from biological organisms, are not discrete individuals. If linguistic genetic relationships are based on a continuity of linguistic usage down time, with a gradual change in the linguistic structures immanent in usage, ancestor-descendant relationships can be specified to the extent that a breakdown in communication across generations did *not* take place, in other words, to the extent that there are no discrete ancestors and descendants. Any distinction between mother and daughter languages is then based on sociopolitical concerns (North Finnish and Meänkieli), structural concerns (the Germanic consonant shift) or indeed the structure of the tree model itself (Finnish and Estonian are daughter languages of Late Proto-Finnic precisely because it is the most recent linguistic stage ancestral to both). What is lacking in linguistics is a level below the level of species and forms varying or changing through time, analogous to the level of individual genotypes or genes, to which the higher level can be reduced. The application of genetic models to languages (and perhaps to other cultural practices or semiotic systems liable to historical change) is thus always more or less metaphorical. They apply well to specific language families such as Uralic, Indo-European and Polynesian; possibly less well to languages in Australia and New Guinea if Dixon's (1997) thesis on the non-applicability of the family tree model to those languages is correct; and possibly less well to creoles and a number of purported cases of language mixing if Thomason & Kaufman's (1988) analysis of their origins as nongenetic is correct—all of this quite aside from the question whether genetic mixing can be adequately modeled.

To sum up: even if structural subsystems such as bound morphology may, in normal cases, be a reliable guide to uncovering genetic relationships, these genetic relationships are ultimately rooted in the history of speech-communities and cannot be deduced directly from correspondences in structure. The usual structural correspondences may be deceptive in some cases: the possible Cushitic origin of Ma'a is indicated by its vocabulary, not its bound morphology—but there are many cases where it is the other way around. Also, there may well be processes of language genesis which can neither be covered by metaphors of genetic mixing or divergence. This underscores the need to model language mixing as a historical process, rooted in the behaviour of a mixing speech-community, rather than as a matter of linguistic structure.

2.3. Requirements for a model

On the basis of the above, I believe the following can be stated as requirements for any theoretical model incorporating genetic language mixing:

- 1) It must be gradual, to do justice to the smeared-out-in-time aspect of language as a social, normative system which is transmitted as a whole; in other words, it must account for a drawn-out mixing process in which both languages are involved as wholes (as incorporating just parts of another language into the framework of another constitutes borrowing rather than mixing).
- 2) It must base itself on a non-reductionist conception of language as a historical entity. Even if the presence of linguistic constructions and "chunks of discourse" (Bybee & Hopper 2001:8) in linguistic memory may have an important role to play in linguistic usage, language as a historical entity may not be reduced to linguistic usage or knowledge of language: we cannot

understand language change without reference to the abstract and normative side of language.

- 3) It must base itself on a realist conception of reconstruction and genetic relationships, as realism is presupposed in the practice of historical linguists, if not always in theoretical pronouncements.

I believe one proposed model potentially fulfills these requirements—namely Auer's (1999), which I will discuss below. I will then apply his model to the problems posed by Old Helsinki Slang for genetic linguistics.

3. Code-switching and genetic mixing

A number of researchers have emphasized the role of code-switching in the genesis of 'mixed' languages such as Media Lengua, Michif, Ma'a and Mednyj Aleut (Myers-Scotton 1993:216, 221, 2003:94-95; Thomason 1997a:464-465; Auer 1999; McConvell & Meakins 2005; McConvell 2008). The relevance of code-switching for a model of the genesis of mixed languages has been doubted as well (Bakker 2003; Backus 2003). Arguments mentioned by Bakker and Backus include that code-switching involves quantitatively less lexical insertion than exhibited by mixed languages, that code-switching tends not to involve basic vocabulary (Bakker 2003:129) and that code-switching in intense contact situations tends to involve alternational rather than insertional code-switching, which is harder to conceptualize as a prior stage to language mixing (Backus 2003:237). In contrast, McConvell & Meakins (2005) claim that a development of a mixed language from code-switching has been observed in the case of Gurindji Kriol, a language in northern Australia sharing features of Kriol, an English-based creole, and Gurindji. According to McConvell & Meakins (2005:15), Gurindji Kriol was formed in the 1960s and 70s as children were exposed to a variety of Gurindji involving pervasive code-switching. Nowadays, a mixed Gurindji Kriol is transmitted to next generations, unlike Gurindji or Kriol themselves. Additionally, the Evenki dialect of Sebjan-Küöl, which is in the process of borrowing whole verbal paradigms from Yakut in a situation marked by intensive code-switching, may be an example of this ongoing process (Pakendorf 2009:101).

The mixed languages mentioned in the literature do not have all the same structural features—Media Lengua has a lexicon-grammar split, Mednyj Aleut and Michif appear to have a split between the nominal and verbal areas of lexicon and grammar, and Ma'a a split between core-vocabulary and non-core vocabulary and grammar. It seems unlikely that all are the result of similar contact mechanisms and unreasonable to demand that a single mechanism explain them all. My main interest here is whether a proposed mechanism for language mixing through code-switching would indeed constitute genetic language mixing. Myers-Scotton's (2003:89-90) model proceeds from structural concerns: crucial in the genesis of a mixed language is a turnover in the Matrix Language which does not reach its end-point (at which it would constitute a finished process of language shift) but gets arrested somewhere on the way. Auer's (1999) model departs from the functions code-switching has in a speech community. Auer (1999:310) compares pragmatically functional code-switching (CS), unmarked code-switching or language mixing (ML) and fused lects to the grammaticization cline proposed by Givón which regards discourse and structure as opposite sides of a continuum. Code-switching is often pragmatically motivated, with the degree of difference in languages involved having no direct relevance to this motivation: code-switching may as easily occur between closely related dialects or sociolects (Alvarez-Cáccamo 1999:37-38; Auer 1999:313). However, in certain circumstances, code-switching may become an unmarked feature of the variety of speech in question, often emblematic for a new or mixed identity (Auer 1999:314-315, 320-321).

Language varieties which are distinguished by unmarked and pervasive code-switching are not hard to find. Two examples from South Africa are Cape Town Afrikaans, the variety of Afrikaans

spoken in the working-class, coloured districts of Cape Town for which code-switching with English as an embedded language is emblematic (McCormick 2002), and Tsotsitaal, a language variety spoken by urban blacks in the townships of Johannesburg which likewise has Afrikaans as a Matrix and English as an Embedded language (Slabbert & Myers-Scotton 1996; Makhudu 2002). With Cape Town Afrikaans, code-switching indeed seems to correspond to the mixed ethnic heritage of its speech community (McCormick 2002:222). With Tsotsitaal, the need for a “secret” language emblematic for group membership may have been a factor in its emergence in addition to the need for a vehicle for interethnic communication (Slabbert & Myers-Scotton 1996:321). A Uralic example—that of unmarked code-switching or in Auer’s (1999) term language mixing between Karelian and Russian—has been presented in detail by Sarhimaa (1999:195-199). The final stage in Auer’s model is the “fossilization” of unmarked code-switching in which the use of a specific code for a specific constituent becomes conventional: the result is a fused lect (Auer 1999:321).

I believe that the process modelled by Auer does indeed constitute mixing in a genetic sense; it may be the only process to do so. The main criterion mentioned above, that a mechanism for the emergence of mixed languages allows for a mixing gradual enough or a co-existence between two languages long enough not as to constitute a ‘break’ in the existence of language down time is fulfilled by the model in its crucial second phase: in the period of unmarked code-switching, the two languages as conventional systems of signs are still discrete, even if they are ‘mixing’ on the discourse level. However, a new speech community to whom this mixing on the discourse level is emblematic is already emerging, and both languages are, as a whole, available as registers to the speakers of that community. Only after this new speech community emerges does a new language, a new conventional system of signs (having some features from one, some from the other ancestor) emerge as well. In the second phase, we may still speak of separate languages in the structural sense, but on the discourse level, the two languages do no longer form two distinct ‘networks of speech events’ in as far as the emerging speech community is concerned (they still do, of course, in the sense that both are spoken more widely in non code-switching varieties). In the third stage, mixing on the discourse level is projected as mixing on the structural level. The process does not conform to the normal, monogenetic processes of language transmission and contact-induced change, neither does it, in any obvious sense, involve a break in transmission radical and punctual enough to speak of non-genetic origins. Instead, we really have an ‘intertwining’ of linguistic lineages that conforms to the three criteria I mentioned above: the process is gradual; it is compatible with a holistic view on language as social and conventional in nature, rather than on the isolation of single speech events; and it is realist in that it is based on primarily the social history of its speakers and of their usage of language, rather than the genetic origins of any particular subsystem.

Above, I argued that no linguistic subsystems may serve as markers of genetic identity as a matter of conceptual necessity. Bound morphology may be very often a reliable guide to genetic relationships; but it is not always so. Moreover, it seems to be possible for various historical processes to lead to the same result, as the controversy surrounding the genetic identity of Ma’a seems to attest. This does not mean that there is no connection at all between historical processes and linguistic results. Thus mixing as a consequence of unmarked code-switching will lead, obviously, to a linguistic variety combining features of both the original Matrix and the Embedded language. Even if code-switching may occur between very closely related or very similar linguistic varieties, some perceptible difference seems to be required for the functional needs of code-switching itself to be fulfilled, or for a unmarked code-switching variety to function as emblematic of a mixed ethnic heritage. Is it possible to state unique linguistic results of processes such as that sketched by Auer, which distinguish it from ‘mixed’ languages which emerged in different ways?

Both Myers-Scotton and Auer argue that it is: Myers-Scotton (2003:91) argues that genuine mixed languages should show a composition in structure beyond the lexicon: thus Media Lengua (which has a largely Spanish lexicon and a Quechua grammar) is rejected as a mixed or ‘split’ language, in contrast to such cases as Mednyj Aleut. Auer (1999:327) likewise argues for some measure of mixed structure to be necessary in order to speak of a fused lect. McConvell (2008:187-188) stress that hybrid languages in Australia such as Gurindji Kriol do show such a split in structure; in the case of Gurindji Kriol, the tense-aspect-mood system as well as most basic verbs are Kriol in origin, whereas case-marking, nominal derivational morphology and the like are Gurindji in origin (McConvell & Meakins 2005:10; McConvell 2008:198). Using the historical make-up of a specific subsystem of a language as a defining criterion of genetic mixing is, as I mentioned above, highly problematic. However, to speak of ‘genetic mixing’ requires the two languages involved to be, potentially, available as a whole (thus going beyond the lexicon), to the code-switching community in the ‘unmarked code-switching phase’. A degree in split structure provides proof of this process. Relexification—the substitution of the lexicon—does not in itself imply genetic mixing: if Ma’a may substitute most of its bound morphology but may still be, on good historical arguments, considered a Cushitic language, then the same obviously goes for Media Lengua and its lexicon.

4. Contrasting approaches to Old Helsinki Slang

Old Helsinki Slang and its possible nature as a mixed language has been the subject of two contrasting recent treatments: that of Kallio (2007) who argues that OHS is unambiguously Uralic (Kallio 2007:178) and Jarva (2008) who argues that OHS may well be described as a mixed language (Jarva 2008:77-78). Both writers, however, agree on the basic facts about OHS. These are, briefly, the following:

Old Helsinki Slang arose in the bilingual working-class quarters of Helsinki at the end of the 19th century and was in use, among males and in youth gangs, until the Second World War. At the time of its inception, Helsinki, originally a majority Swedish-speaking town, had undergone a large influx of Finnish speakers in tandem with industrialization, and by 1900 about half the city spoke Finnish. OHS became a vehicle for intergroup communication for speakers of both languages, and emblematic for a new, urban identity (Kallio 2007:178; Jarva 2008:56-57), and developed in close contact and interaction with Swedish slang varieties (Cantell *et al.* 1989:11-12, Forsskåhl 2006:60-62). Postwar colloquial varieties of Finnish in Helsinki have been influenced by OHS to an extent but have a much more clearly Finnish lexicon (Paunonen 2000:16).

And the most striking feature of OHS is indeed its borrowed vocabulary, most of it originating from non-standard varieties of Swedish, a much smaller amount from Russian. Paunonen (2000:16) and Kallio (2007:178) claim that approximately 75% of OHS vocabulary is of Swedish origin, whereas Jarva (2008:66) expresses scepticism with regards to this estimate, pointing to difficulties in defining OHS vocabulary as a discrete system, distinct from the lexical resources of Finnish and Swedish with which OHS co-existed. However, Jarva (2008:68) agrees with Kallio (2007:183-187) that the vast majority of the Swadesh 150-word list can be stated in OHS using Swedish-origin lexical items. Most Swedish lexicon are content words, however, OHS sports some adverbs of Swedish origin as well and as Kallio (2007:187) points out even some adpositions: these are however, in accordance with Finnish morphosyntax, used as postpositions and take Finnish case markers. Originating from Swedish as well are a number of derivational endings such as *-is*, which is used for integrating Swedish lexicon into the framework of OHS (Paunonen 2000:25, 2006:339). Most of these derivational suffixes are unique to OHS though a number have survived into present-day colloquial Finnish. Pronouns, conjunctions and numerals, however, are Finnish (Jarva 2008:67).

OHS phonology is convergent, containing phonemes (voiced stops, affricates) as well as phonotactic features (most prominently, initial consonant clusters) alien to Standard and dialectal

varieties of Finnish. These phonotactic features seem to pragmatically serve as a marker of identity (Kallio 2007:188, Jarva 2008:71). Vowel harmony is absent in the Swedish but not in the Finnish lexicon of OHS (Kallio 2007:189; Jarva 2008:71). OHS bound morphology, in contrast, is Finnish; there are no Swedish grammatical markers used in OHS and the full structure of Finnish nominal and verbal morphology is retained (Kallio 2007:178-179; Jarva 2008:72). In some respects, OHS morphosyntax differs from that of Standard and dialectal Finnish: an innovation unique to OHS is the emergence of a new verbal stem type which has levelled the morphophonemical alternations that mark its closest equivalent in Standard Finnish (Paunonen 2000:22-23, Jarva 2008:73). Also, there are irregularities in object case-marking with the unmarked and the marked *-n* object both showing counterposed tendencies for generalization (Paunonen 2000:25); this feature is well-known for varieties of Finnish in close contact with Swedish, such as Meänkieli in the Torne valley region of Sweden (Kangassalo, Nemvalts & Wande 2003). None of the morphosyntactic features of OHS seem to imperil its classification as Finnish: there is no mixed structure analogous to that of Michif, Mednyj Aleut or Gurindji Kriol. OHS phonology is obviously mixed, but this is hardly unique among the Finnic languages: the phonologies of Russian and Karelian, for example, appear to be largely convergent (Sarhimaa 1999:178). The feature of OHS giving rise to the discussion of OHS as a mixed language is precisely its lexicon, and in that regard, OHS is similar to Media Lengua, but not Mednyj Aleut or Michif.

In the integration of its lexicon into a Finnish morphosyntactic frame, OHS shows some typical features of what Halliday (1976) calls an anti-language: a linguistic variety used at the margins of society to embody an alternative social reality. A typical anti-linguistic feature of OHS is the overlexicalization of certain semantic domains (Halliday 1976:571, see Makhudu 2002:402-403 for Tsotsitaal), thus Paunonen (2000:32-33) mentions boys and girls, money, the police, fighting, alcohol, sexuality and adjectives such as 'stupid' and 'nice' as central lexical domains in OHS, with various expressions for 'girl' or 'woman' numbering in the hundreds. Another feature of anti-languages according to Halliday (1976:576-579) are the usage of lexical substitution processes such as metathesis, metaphor, compounding, etc. to distinguish the anti-linguistic lexicon from the lexicon of the main, dominant language. In OHS, the usage of derivational suffixes (Paunonen 2000:25-28) and most strikingly the use of initial consonant clusters (Paunonen 2000:19-20) serve to set the specific OHS slang lexicon apart from that of Finnish. OHS initial consonant clusters were often not part of the original borrowed lexical item, thus Jarva (2008:60) mentions that the English word *cowboy* may appear in OHS (with the derivational suffix *-ari*) as *koobari*, *scoobari*, *goobari*, *skoubari*. The OHS word *rigi* 'suit' has been noted to occur as *brigi*, *drigi*, *krigi*, *sprigi*, *skrigi*, *strigi* etc. (Paunonen 2006:337). Variation in the usage of derivational suffixes as well serves to create an enormous amount of lexical synonymy in OHS, however, lexical synonymy and variation in the usage of initial consonant clusters seems have to reached its greatest extent after the Second World War, when OHS had given way to more finnicized varieties of slang (Paunonen 2000:21, 2006:340-341). The use of metaphorical expressions in hyperlexicalized semantic domains such as 'girl' or 'woman' likewise is a feature of postwar Helsinki slang much more than of OHS (Paunonen 2006:360). Anti-language features in Halliday's sense seem thus to be particularly a feature of later varieties of Helsinki slang.

Jarva and Kallio both have different opinions as to the mixedness of OHS. Kallio (2007:178) argues OHS is unambiguously Uralic on the basis of its Finnish grammatical structure:

As most scholars throughout linguistic history have considered grammatical structure the surest proof of genetic relatedness, there should be no doubt that Stadin Slangi is genetically Uralic, or more precisely, Finnic.

The Swedish lexicon of OHS is regarded as the result of massive lexical borrowing (Kallio 2007:179). I would disagree with this line of reasoning, as structure is neither a necessary indicator

of genetic relatedness, nor is it empirically always so—as mentioned above, Ma’a would possibly a counterexample if Greenberg’s (1999) and Thomason’s (1997b) description of its historical emergence is to be accepted (note, however, that Kallio (2007:190) regards Ma’a as Bantu-based). Though Kallio does distinguish OHS and Media Lengua, which show a lexicon-grammar split, from cases such as Michif and Mednyj Aleut with their split structure (Kallio 2007:179), he relates with scepticism to the existence of genetically mixed languages in general:

[...] it is the very essence of linguistic borrowing that a target language always maintains its genetic identity, no matter how many lexical or even grammatical items it borrows. Namely, when a target language fails to remain itself, the process is no longer called borrowing but language shift. (Kallio 2007:182-183).

Jarva, on the other hand, argues that OHS “*might well be described as an intertwined mixed language.*” (Jarva 2008:79), on the basis of the similarities between OHS and languages such as Media Lengua and Ma’a. Jarva’s argument rests on his conception of the sociohistorical circumstances surrounding the emergence of OHS, arguing it likely arose through frequent, unmarked code-switching between Finnish and Swedish leading to relexification—Finnish speakers substituting the basic vocabulary of their speech variety with Swedish lexical items, and regrammaticization—Swedish speakers acquiring a Finnish grammatical framework for the lexicon of early OHS (Jarva 2008:78). The latter is considered to have possibly lead to some of the morphosyntactic peculiarities of OHS (Jarva 2008:78). What is not explicit in Jarva’s treatment of the question is whether OHS is to be considered *genetically* mixed rather than mixed in terms of synchronic make-up (Swedish lexicon, Finnish grammar). It is absolutely true, as Jarva argues, that OHS bears such strong similarities to Media Lengua and Ma’a that, as a problem case for genetic linguistics, it deserves consideration alongside them (Jarva 2008:79), however, in the literature concerning such languages, various notions of what it means to be ‘mixed’ (some clearly historical, others synchronic/structural in nature) seem to compete. The languages commonly grouped under the moniker ‘mixed’ or ‘intertwined’ may well be the results of very different historical processes, and not all of them necessarily mixed in the genetic sense.

Jarva’s focus on the sociohistorical circumstances of the emergence of OHS as a key to resolving its mystery is, however, entirely correct. Here, Jarva’s opinion differs from that of Kallio and Paunonen. Paunonen (2000:16) argues that OHS originated as a kind of pidgin used for communication between Finnish and Swedish speakers. Kallio (2007:180) likewise draws attention to the function of early OHS as a vehicle for intercultural communication and its pidgin-like features: among these, he names a limited vocabulary of some 2000 words and a simplified grammar. Later on, this simplified language expanded and gradually merged with Finnish (Kallio 2007:181). Jarva (2008:76) criticizes this point of view, noting that it is impossible to estimate the size of OHS vocabulary because OHS vocabulary was to a large extent open to introducing new (especially Swedish) lexical items around an integrated core (with the specific derivational suffixes and phonological features marking the lexicon as precisely OHS) (Forsskåhl 2006:63; Jarva 2008:66), and noting that we have no proof of a radically simplified OHS grammar. Though Kallio (2007:180) plausibly argues that we have very little documentation of OHS grammar as it is, most of the material we have being vocabulary and most text samples being too late to serve as proof of the structure of early OHS, the fact remains that OHS grammar does seem to have retained its full complexity. The specific OHS verbal conjugation does level a number of the morphophonemic alternations specific to other varieties of Finnish, and irregularities in the usage of object cases have been noted—but these are both well within the range of variation in the Finnic languages in general.

5. Is OHS a mixed language?

At first sight, however, Kallio (2007) is quite correct in pinpointing to the Finnish structure of OHS as proof it is genetically Uralic, rather than mixed. In code-switching models of the emergence of mixed language such as those of Auer (1999) and Myers-Scotton (2003) some amount of composite structure is necessary for a language to be seen as genetically mixed, in other words, as evidence that during a prior ‘mixing’ stage of unmarked code-switching, both languages as a whole were potentially available as registers to use. This evidence is absent in the case of OHS, which is closely similar to *Media Lengua* rather than to Michif or Mednyj Aleut with their characteristic composite structures. In the unmarked Russian-Karelian code-switching varieties that Sarhimaa (1999:241-246) describes, speakers switch effortlessly between Russian and Karelian constituents and also between Russian and Karelian as a Matrix language, and with code-switching varieties of Cape Town Afrikaans, the same has been observed (McCormick 2002:225-226). It would be easy to conceptualize such a code-switching variety to, at some point, conventionalize the use of Russian for a particular type of constituent and Karelian for another, and the result would be something very similar to Mednyj Aleut or Michif. However, if OHS does have its roots in Finnish-Swedish code-switching, evidence for prior variation in the use of Matrix languages or of a Matrix language turnover (Myers-Scotton 1993:214) is lacking in the case of OHS, and thereby, evidence that OHS is the conventionalized result of a prior ‘mixing stage’ through unmarked codeswitching and that OHS is thus a genetically mixed language is lacking as well.

However, the ‘mixed’ part is only part of the question relevant here, the other part is constituted by to what extent OHS can be considered a language. OHS, after all, existed symbiotically with various varieties of Finnish and Swedish. Notably, OHS did not have monolingual native speakers. Though Kallio (2007:180-181) points out that OHS may well have been acquired in a nativelike fashion by young people on the streets of Helsinki, no speakers have exclusively acquired OHS as a native language. Of course, regarding the existence of monolingual speakers as a criterion for languagehood would quite radically cull the inventory of the world’s languages. Kallio (2007:151) considers OHS to be a distinct dialect of Finnish, possibly a language; Paunonen (2000:36) regards it as a distinct language. The question at issue here is to what extent OHS was conventionalized. It might be worthwhile to consider a thought experiment proposed by Thomason (1997b:476-477) in response to Mous’ argument (Mous 1994:196, 199) that the symbiotic relationship between Ma’a and the clearly Bantu Outer Mbugu must be regarded as a crucial factor in genetically classifying Ma’a. Thomason envisions the Ma’a severing their sociolinguistic links with the surrounding Bantu speakers and using exclusively Ma’a in the future. In this hypothetical case, the same problems with classifying a language with Cushitic vocabulary and Bantu morphology would remain: the sociolinguistic ties between Ma’a and Outer Mbugu themselves would have no bearing on the question. Could we envision a group of speakers of OHS moving to, say, a distant tropical island and only speaking OHS among themselves?

In terms of grammatical structure, nothing about OHS would seem to make this scenario impossible. The question centers on its vocabulary: to what extent was the vocabulary of OHS a more or less closed, conventionalized system, and to what extent was it open to adopting code-switches and new lexical items from Swedish? More precisely, if OHS does not conform to the final stage in Auer’s (1999) model, does it perhaps correspond to an intermediate stage: would it be possible to regard OHS as an unmarked code-switching variety of Finnish? At first sight, the Swedish lexicon of OHS appear to be borrowings rather than code-switchings: they are well integrated into the morphological system of Finnish and can even be seen to be doubly integrated: first into the matrix of phonological features and derivational suffixes that pragmatically distinguish OHS, second into the grammatical structure of Finnish. Morphological integration as a criterion of borrowing is, however, highly problematic: Poplack, Wheeler & Westwood (1989) do not regard

morphologically integrated single lexical items as code-switchings in their research on Finnish-English bilingualism but have to resort to a third category, that of nonce borrowings, instead (Poplack, Wheeler & Westwood 1989:396-397). This conception was criticized by Myers-Scotton (1993:23) who points out single-word switches are common with agglutinative Matrix languages (Myers-Scotton 1993:31) and that code-switching and borrowing rather form a continuum with frequency as the main distinguishing factor between them (Myers-Scotton 1993:176). Sarhimaa (1999:194) is likewise critical towards the notion of nonce borrowings. Rather than morphosyntactic and phonological integration, conventionality seems to be the main criterion distinguishing borrowing from code-switching. With code-switching, the embedded language as a whole is potentially open to usage. Jarva (2008:66) argues that this was precisely the case with OHS and its Swedish lexicon. Another indicator for the openness, or lack of conventionality, of the OHS lexicon would be the tremendous variation in the phonotactics of the Swedish lexical items, though the variation in initial consonant clusters and the like seems to have reached its heights with later, postwar varieties of Helsinki Slang.

As a further indication of the possible role of code-switching in the emergence of OHS, Jarva (2008:77-78) mentions the presence of longer code-switched phrases in some Helsinki slang material, and draws an interesting parallel between OHS and *kyökkisuomi* or 'Kitchen Finnish', an umbrella term for older Finnish-Swedish code-switching varieties used, for example, in communication between Swedish landowners and Finnish-speaking kitchen staff (Jarva 2008:59). A lot of anecdotal evidence for earlier code-switching varieties between Finnish and Swedish exist, such as the statement by an inhabitant of Sjundea in Nyland in 1847 that "*almost all the Finns that one can find here, in the municipality of Sjundea, speak Swedish too. A consequence of this is that the Finnish dialect which you can hear around here is mixed with Swedish for almost one fourth of it*" (Wallén 1932:102) or the words of the 19th century grammarian Renvall (quoted in Cannelin 1926:79):

In some cities and in many places in the countryside as well language is such mumbo-jumbo, that it is not Finnish, neither Swedish, but an anarchic mixture of both which clearly proves how ugly a language can become, if it is unthinkingly composed of different substances, either two different languages or two different dialects. Try, for example, the company of Finnish women in Turku! What a barbaric pig's latin for a language, for example *onk syster ollu visitin pääl, tul helseman minun päällen, se on farlit seilata tormiss yli haavin* and so on. (Translation of both quotes is my own).

The examples posited by Renvall both involve integrated lexical items (*seilata* 'to sail'; *tormiss* 'storm' with an inessive ending and the reduction of the initial consonant cluster characteristic for Finnish loanword integration; as well as *haavin* 'sea-GEN') and the unintegrated *farlit* which retains Swedish phonology and apparently the suffix *-t* (Sw *farlig-t*). The existence of earlier code-switching varieties is beyond doubt, and possibly parallels very interesting; though there is obviously a dearth of material. But, though Jarva (2008:59) mentions possible lexical inflow from Kitchen Finnish varieties to OHS and also possible contact in the area of morphosyntax, it is doubtful whether relationships between Kitchen Finnish and OHS go much beyond those of interesting parallels. Notably, OHS gained a function as identity-marker that earlier code-switching varieties of Finnish likely never had.

A strong argument against regarding early OHS as an unmarked code-switching variety is that the embedding of Swedish lexical forms itself becomes conventional to the extent that the vast majority of content words are substituted with Swedish-origin words. There is a stark difference in that regard between OHS and unmarked code-switching varieties such as Russian-Karelian mentioned above and Tsotsitaal, an Afrikaans-English code-switching variety which is unmarked in the sense that embedding English lexicon and phrases is a marking feature of the variety itself, may share some similarities with OHS in its emergence: possibly, Tsotsitaal emerged as a vehicle of

interethnic communication which later developed more and more to a slang with “anti-language” features (Slabbert & Myers-Scotton 1996:322). However, in Tsotsitaal, the Afrikaans Matrix language is dominant in lexicon as well, with the frequency of English items falling short of the near-wholesale vocabulary substitution encountered in OHS, and a significant number of longer code-switches phrases occurring (Slabbert & Myers-Scotton 1996:331-334). For OHS, in contrast, the proper name of the process seems to be relexification: the near-wholesale substitution of content vocabulary, which Muyskens (1981:62-63) sees as lying at the basis for Media Lengua. On the other hand, it may be possible to regard relexification, as Myers-Scotton (1993:216-217) does, as a very intense and frequent type of code-switching.

It is, in any event, hard to imagine the emergence of a variety such as OHS without very frequent and intense code-switching: borrowings do not enter a language fully phonologically and morphologically integrated like Metis springing from Zeus’ head. But we lack the footprint, in the shape of composition in structure, that would enable us to unambiguously see OHS as a descendant or as an example of an unmarked code-switching variety. Unfortunately, the variety of OHS most relevant to its genetic identity would be the earliest one, and we have little material on its features except its vocabulary (Kallio 2007:180). Thus, *pace* Jarva (2008:79), I am unsure of the relevance of OHS to the question of the emergence of mixed languages, at least in accordance with Auer’s model.

5. Consequences: language genesis in a black box

As mentioned, Media Lengua and OHS would be rejected as genetically mixed languages based on the models provided by Auer (1999) and Myers-Scotton (2003). But should they? There is no evidence of a preceding unmarked code-switching stage in the structural domains of either language, but does absence of such evidence preclude a mixed genetic origin? A moment’s reflection shows that it does: even if OHS in its early stages arose from an unmarked code-switching or ‘language-mixing’ variety similar to those found in Russian-Karelian and English-Afrikaans contact situations, OHS as it is strongly suggests a variety with Finnish, and only Finnish, as its Matrix language lay at its roots. Unmarked code-switching varieties such as Kaaps Afrikaans are not (yet) languages in and of themselves, but rather intermingling usage patterns of two languages in use by a specific speech community, and if it is conceptually possible that such a variety may, through conventionalization of a specific language for specific kinds of constituents, give rise to a genetically mixed language. But it is equally conceptually possible that one of the languages in use takes over as an exclusive matrix language. Thus the notion that early OHS may have been a ‘language-mixing’ variety does not commit one to the incoherent notion that a mixed language has given birth to one that is unambiguously Uralic (as the more recent Helsinki slang, if not OHS, clearly is). OHS is therefore not a mixed language in the genetic sense.

Thus, I agree in essence with Kallio’s (2007) position that OHS is to be seen as Uralic on the basis of its clearly Finnish grammatical structure, even if I consider the crucial role of code-switching in the emergence of OHS as argued by Jarva (2008:66) to be highly likely. My reasoning differs from Kallio’s in that I believe it is quite possible to state what a genetically mixed language would look like and how it would emerge—we just lack the evidence to call OHS one. Also, I consider the importance of structural relationships in assigning genetic identity to be very relative: though a lack of composition in structure would seem to imply monogenesis, it does not tell us, in case of languages sporting a grammar-lexicon split, which ancestor the language would need to be assigned to. Regardless of the actual history of Ma’a, the scenario sketched by Thomason (1997b) and Greenberg (1999) of a very gradual Bantuization of an originally Cushitic language seems quite conceivable, and therefore, it is conceivable that there are cases in which basic vocabulary, rather than structure, is the more reliable guide to genetic relationships. What may be plausible in the case of Ma’a—a language spoken by people eager to preserve and assert a threatened ethnic identity in

contrast to surrounding speakers of Bantu languages—is not, of course, in the case of OHS, a language originally developed as a vehicle for interethnic communication, then slowly becoming a badge for a new, emergent urban identity. In both cases, vocabulary is the most salient component of language, and open to conscious efforts of retention, substitution or manipulation: but it is not predetermined what speakers will do with that salient component. It is really the social history of the speakers that holds the key to the genetic relationships of their languages.

References

- Aikhenvald, Alexandra. 2003. Mechanisms of change in areal diffusion: new morphology and language contact. *Journal of Linguistics* 39:1-29.
- Aitchison, Jean. 2001. *Language change: progress or decay?* Cambridge: Cambridge University Press.
- Alvarez-Cáccamo, Celso. 1999. From ‘switching code’ to ‘code-switching’. Towards a reconceptualization of communicative codes. In: Peter Auer (ed.), 29-48. *Code-switching in conversation. Language, interaction and identity*. London: Routledge.
- Andersen, Henning. 2006. Synchrony, diachrony and evolution. In: Ole Nedergaard Thomsen (ed.), 59-90. *Competing models of linguistic change: evolution and beyond*. Current Issues in Linguistic Theory 279. Amsterdam: John Benjamins.
- Anttila, Raimo. 1989. *Historical and comparative linguistics*. Current Issues in Linguistic Theory 6. Amsterdam: John Benjamins.
- Auer, Peter. 1999. From code switching via language mixing to fused lects: toward a dynamic typology of bilingual speech. *International Journal of Bilingualism* 3-4:309-332.
- Backus, Ad. 2003. Can a mixed language be conventionalized alternational codeswitching?. In: Yaron Matras & Peter Bakker (eds.), 237-270. *The Mixed Language Debate. Theoretical and Empirical Advances*. Trends in linguistics Studies and monographs 145. Berlin: Mouton de Gruyter.
- Bakker, Peter & Muysken, Pieter. 1994. Mixed languages and language intertwining. In: Jacques Arends, Pieter Muysken & Norval Smith (eds.), 41-52. *Pidgins and creoles. An introduction*. Creole Language Library 115. Amsterdam: John Benjamins.
- Bakker, Peter. 2003. Mixed languages as autonomous systems. In: Yaron Matras & Peter Bakker (eds.), 107-150. *The Mixed Language Debate. Theoretical and Empirical Advances*. Trends in linguistics Studies and monographs 145. Berlin: Mouton de Gruyter.
- Brinton, Laurel J. & Traugott, Elizabeth Closs. 2005. *Lexicalization and language change*. Cambridge: Cambridge University Press.
- Bybee, Joan & Hopper, Paul. 2001. Introduction to frequency and the emergence of linguistic structure. In: Joan L. Bybee & Paul Hopper (eds.), 1-24. *Frequency and the emergence of linguistic structure*. Typological studies in language 45. Amsterdam: John Benjamins.
- Bybee, Joan. 2006. From usage to grammar: the mind’s response to repetition. *Language* 82:711-733.
- Cannellin, Knut. 1926. Ruotsinkielen vaikutus nykyiseen yleiskieleemme. *Virittäjä* 30:76-80.
- Cantell, Ilse, Lehti-Eklund, Hanna, Nyholm, Leif & Salokivi, Jouni (eds.). 1989. *Helsingforsslang nu. En skolenkät 1988*. Meddelanden från institutionen för nordiska språk och nordisk

- litteratur vid Helsingfors universitet B:12. Helsingfors: Institutionen för nordiska språk och nordisk litteratur.
- Croft, William. 2000. *Explaining language change. An evolutionary approach*. Harlow: Longman linguistics library.
- Croft, William. 2003. Mixed languages and acts of identity: An evolutionary approach. In: Yaron Matras & Peter Bakker (eds.), 41-72. *The Mixed Language Debate. Theoretical and Empirical Advances*. Trends in linguistics Studies and monographs 145. Berlin: Mouton de Gruyter.
- Croft, William. 2006. The relevance of an evolutionary model to historical linguistics. In: Ole Nedergaard Thomsen (ed.), 91-132. *Competing models of linguistic change: evolution and beyond*. Current Issues in Linguistic Theory 279. Amsterdam: John Benjamins.
- Dixon, Robert M.W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.
- Forsskåhl, Mona. 2006. Helsinki slang around 1900: A new slang variety is born. *Revue d'Études Françaises* 11:53-67.
- Halliday, Michael A.K. 1976. Anti-languages. *American Anthropologist* 78-3:570-584.
- Greenberg, Joseph. 1999. Are there mixed languages?. In: Joseph Greenberg & William Croft (eds.), 349-357. *Genetic Linguistics: Essays on Theory and Method*. Oxford: Oxford University Press.
- Itkonen, Esa. 1998. Sukupuu ja kontakti. *Virittäjä* 102:90-103.
- Jarva, Vesa. 2008. Old Helsinki Slang and language mixing. *Journal of Language Contact – Varia* 1:52-80.
- Kallio, Petri. 2007. How Uralic is Stadin Slangi?. In: Rogier Blokland & Cornelius Hasselblatt (eds.), 176-191. *Language and identity in the Finno-Ugric world*. Studia Fenno-Ugrica Groningana 4. Maastricht: Shaker publishing.
- Kangassalo, Raija, Nemvalts, Peep & Wande, Erling. 2003. Ruotsin itämerensuomalaisten kielten syntaksi: kieliopilliset sijat meänkielessä, ruotsinsuomessa ja ruotsinvirossa. In: Pirkko Muikku-Werner & Hannu Remes (eds.), 156-166. *Viro ja suomi: kohdekielet kontrastissa*. Lähivertailuja 13. Joensuu: Joensuun yliopisto.
- Künnap, Ago. 1998. *Breakthrough in present-day Uralistics*. Tartu: Tartu university.
- Laakso, Johanna. 1995. A spade is always a spade - comment on "The 'pragmareal' challenge to genetic language tree models". In: Seppo Suhonen (ed.), 70-75. *Itämerensuomalainen kulttuurialue*. Castrenianum Toimitteita 49. Helsinki: Suomalais-Ugrilainen Seura.
- Lass, Roger. 1993. How real(ist) are reconstructions?. In: Charles Jones (ed.), 156-189. *Historical linguistics: problems and perspectives*. London: Longman Linguistics Library.
- Makhudu, K. Dennis Papi. 2002. An introduction to Flaaitaal (or Tsotsitaal). In: Rajend Mesthrie (ed.), 398-406. *Language in South Africa*. Cambridge: Cambridge University Press.
- Matras, Yaron. 2000. Mixed languages: a functional-communicative approach. *Bilingualism: Language and Cognition* 3 -2:79-99.
- McConvell, Patrick & Meakins, Felicity. 2005. Gurindji Kriol: A mixed language emerges from code-switching. *Australian journal of linguistics* 25-1:9-30.

- McConvell, Patrick. 2008. Mixed languages as outcomes of code-switching: recent examples from Australia and their implications. In: Robert Nicolai & Bernard Comrie (eds.), 187-212. *Language Contact and the Dynamics of Language: Theory and Implications*. Journal of Language Contact, Thema 2. <http://www.jlc-journal.org/>.
- McCormick, Kay. 2002. Code-switching, mixing and convergence in Cape Town. In: Rajend Mesthrie (ed.), 216-234. *Language in South Africa*. Cambridge: Cambridge University Press.
- McMahon, April. 1994. *Understanding language change*. Cambridge: Cambridge University Press.
- Mous, Maarten. 1994. Ma'a or Mbugu. In: Peter Bakker & Maarten Mous (eds.), 176-200. *Mixed languages. 15 case studies in language intertwining*. Studies in language and language use 13. Amsterdam: Institute for Functional Research into Language and Language Use.
- Mufwene, Salikoko. 2001. *Ecology of Language Evolution*. Cambridge: Cambridge University Press.
- Muysken, Pieter. 1981. Halfway between Quechua and Spanish: the case for relexification. In: Arnold Highfield & Albert Valdman (eds.), 52-78. *Historicity and variation in creole studies*. Ann Arbor: Karoma publishers.
- Myers-Scotton, Carol. 1993. *Duelling Languages. Grammatical Structure in Codeswitching*. Oxford: Oxford University Press.
- Myers-Scotton, Carol. 2003. What lies beneath: Split (mixed) languages as contact phenomena. In: Yaron Matras & Peter Bakker (eds.), 73-106. *The Mixed Language Debate. Theoretical and Empirical Advances*. Trends in linguistics Studies and monographs 145. Berlin: Mouton de Gruyter.
- Pakendorf, Brigitte. 2009. Intensive contact and the copying of paradigms: An Even dialect in contact with Sakha (Yakut). *Journal of Language Contact* – Varia 2:85-110. <http://www.jlc-journal.org/>.
- Paunonen, Heikki. 2000. *Tsennaaks Stadii, bonjaaks slangii. Stadin slangin suursanakirja*. Helsinki: WSOY (Werner Söderström Corporation).
- Paunonen, Heikki. 2006. Synonymia Helsingin slangissa. *Virittäjä* 110:336-364.
- Poplack, Shana, Wheeler, Susan & Westwood, Anneli. 1989. Distinguishing language contact phenomena: evidence from Finnish-English bilingualism. *World Englishes* 8-3:389-406.
- Raukko, Jarno & Östman, Jan-Ola. 1994. *Pragmaattinen näkökulma Itämeren kielialueeseen*. Publications of the department of general linguistics n. 25. Helsinki: University of Helsinki.
- Raukko, Jarno & Östman, Jan-Ola. 1995. The 'pragmareal' challenge to genetic language tree models. In: Suhonen, S. (ed.), 31-69. *Itämerensuomalainen kulttuurialue*. Castrenianiumin Toimitteita 49. Helsinki: Suomalais-Ugrilainen Seura.
- Ringe, Don, Warnow, Tandy & Taylor, Ann. 2002. Indo-European and computational cladistics. *Transactions of the philological society* 100-1:59-129.
- Sarhimaa, Anneli. 1999. *Syntactic transfer, contact-induced change and the evolution of bilingual mixed codes. Focus on Karelian-Russian alternation*. Studia Fennica Linguistica 9. Helsinki: Suomalaisen Kirjallisuuden Seura 1999.
- Shapiro, Michael. 1983. *The Sense of Grammar. Language as Semeiotic*. Bloomington: Indiana University Press.

- Slabbert, Sarah & Myers-Scotton, Carol. 1996. The structure of Tsotsitaal and Iscamtho: code switching and in-group identity in South African townships. *Linguistics* 34:317-342.
- Thomason, Sarah Grey & Kaufman, Terrence. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
- Thomason, Sarah Grey. 1997a. Mednyj Aleut. In: Sarah G. Thomason (ed.), 449-468. *Contact languages. A wider perspective*. Creole Language Library 17. Amsterdam: John Benjamins.
- Thomason, Sarah Grey. 1997b. Ma'a (Mbugu). In: Sarah G. Thomason (ed.), 469-487. *Contact languages. A wider perspective*. Creole Language Library 17. Amsterdam: John Benjamins.
- Thomason, Sarah Grey. 2001. *Language contact: an introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah Grey. 2003. Social factors and linguistic processes in the emergence of stable mixed languages. In: Yaron Matras & Peter Bakker (eds.), 21-39. *The Mixed Language Debate. Theoretical and Empirical Advances*. Trends in linguistics Studies and monographs 145. Berlin: Mouton de Gruyter.
- Wallén, Holger. 1932. *Språkgränsen och minoriteterna i Finland svenskbygder omkr. 1600-1865*. Åbo: Åbo Akademi.
- Yeon-Ju, Lee & Sagart, Laurent. 2008. No limits to borrowing: the case of Bai and Chinese. *Diachronica* 25-3:357-385.