

## Do languages die of 'structuritis'? On the role of code-copying in language endangerment

Lars Johanson

When languages cease to exist, the decisive factors are social in nature. The processes may be accompanied by what is called extensive 'borrowing' and 'simplification'. Structural phenomena of these kinds are, however, sometimes viewed as the very reasons for decay and death. Languages displaying these 'corruptions' are thought to demolish themselves and to destroy their own identity. But the phenomena in question are neither unique to 'language death', nor are they unequivocal signs foreboding it. The structural development itself does not necessarily contribute to the endangerment. There is no empirical evidence for language shifts through gradual structural development. It would indeed be impossible for foreign structures to enter into a linguistic code and infect it with 'structuritis'. Ideas of this kind are the logical consequence of viewing contact-induced influence as 'borrowing' or 'import' of elements from one code to another. According to the Code-Copying Framework, however, elements inserted into a given contact-affected code are not taken to be identical with the originals, but just copies of them. Due to the necessary accommodation of the copies, even typologically highly different elements may be incorporated. Languages do not cease to exist for structural reasons, but because they are no longer acquired by new generations. Even considerable degrees of structural change do not prevent them from being passed on. Languages do not die of 'structuritis'.\*

### *Gradual decline and shift*

Why do languages 'die'? The social circumstances under which they normally cease to exist are rather well known. In asymmetric multilingual settings, a socially dominated linguistic code A may come under the strong pressure of a socially dominant code B, which influences the linguistic behavior of the A speakers. The dominated code becomes restricted to certain domains. Since its speakers need the dominant code in other domains, they become increasingly bilingual, using a primary code A<sub>1</sub>, and a secondary code B<sub>2</sub>. The latter may eventually become their main code. (The terms *code* and *code-copying* will be discussed below.)

A process of shifting away from the recessive code may begin. If the social pressure leads to a negative attitude towards it, its speakers may avoid transmitting it to their children, who acquire it incom-

pletely and, at best, grow up as semi-speakers. The use of  $A_1$  is further restricted, until the rest of its speech community finally abandon it and shift to the dominant code, which thereby becomes their primary code  $B_1$ . For a description of a scenario of this kind, see, e.g., Sasse (1992); compare Campbell & Muntzel's classification of types of language obsolescence (1989).

At the sociolinguistic level, we can thus state that the use of socially strong codes in asymmetric settings tend to threaten weaker codes, reducing their functional domains. The result may be marginalization, decreasing maintenance efforts, declining linguistic competence and finally code shift. The replacement of one code in favor of a co-areal contact code under overwhelming social pressure is a radical qualitative leap. (The question whether the shift preferably goes from more 'complex' to less 'complex' languages cannot be dealt with here.)

### *Linguistic corollaries*

The linguistic corollaries of the social processes mentioned above may encompass several phenomena.

They may include extensive code-copying. The dominated code A may copy increasingly more features of the dominant code B. It may acquire new elements in lexicon, phonology, morphology, syntax, pragmatic organization, etc., mostly substituting them for native elements.

On the other hand, a strongly dominated code may undergo simplification in a gradual process of structural attrition. It may not only lose native lexicon, but ultimately also native structures in phonology, morphology, syntax, etc. The significant structural attrition often observed in endangered languages seems to be due to the collapse of marked categories and massive copying. For characteristic symptoms of attrition see, for example, Dorian 1981 and 1989. The structural simplification may also be a result of foreign influence in the sense of code-copying.

Extensive structural code-copying may thus be said to be typical of 'moribund' or 'dying' languages. The same is true of structural simplification, whether contact-induced or not. This does not mean, however, that the phenomena in question are the *reasons* for the gradual death. Numerous high-copying codes are fully viable, and even clearly recessive codes may creatively develop new features through copying. The structural development itself does not necessarily contribute to the endangerment as if it were some kind of perilous 'structuritis'.

*Structural decay and death*

While this may seem evident, there is a long tradition regarding language death as a result of structural decay. The comparative linguists of the nineteenth century conceived of language as an organism with a predestined life cycle, consisting of birth, development to a Golden Age, and finally decay and death. The peak of the cycle was represented by the classical languages Sanskrit, Greek and Latin, whereas the contemporary languages represented the phase of decay, on the threshold to death.

According to a popular opinion that is still quite widespread, linguistic change, in particular under foreign influence, is a negative phenomenon. Purists want to preserve their languages from external processes that might make them change and lose their supposed 'purity', maintaining that these processes lead to corruptions and, in the long term, to decay and death.

There is thus an infelicitous tradition conceiving of code shift as a result of gradual development brought on by progressively stronger foreign structural influence. Even in today's linguistics, code-copying is often referred to as 'interference', a term that has negative connotations, implying deviations from monolingual norms, and suggesting phenomena that cause impaired communication. Terms such as 'interference' and 'negative transfer' are justified in the context of foreign-language learning, as long as they refer to errors introduced as a result of contact with another language and reflecting the level of competence achieved. However, deviations from norms due to code-copying cannot generally be regarded as failure to learn a full native code.

Furthermore, I argue that the very use of metaphors such as 'borrowing', 'import' and 'transfer' often lead linguists to false conclusions concerning the nature of contact-induced change, since it suggests some kind of implantation of foreign elements into the body of a Basic Code. In the last part of the present paper, I will summarize some principles of the Code-Copying Framework as an alternative conceptual basis.

*Shift through extensive copying*

Even in relatively modern linguistic work we are confronted with the idea that extensive contact-induced influence would, as it were, lead to some kind of 'structuritis' with lethal effects. Increasing structural influence on a code would lead to its abandonment.

Many contact linguists maintain, in more or less clear ways, the idea of a successive transition from an old primary code  $A_1$  to new a primary code  $B_1$  through intermediary stages characterized by increasing copying. Speakers of a dominated code with growing proficiency in a dominant code are thought to take over larger and larger parts of the dominant code, until they end up speaking the dominant code instead of the dominated one.

Einar Haugen quite correctly states that speakers of dominated languages may proceed to shift via different, more or less extended bilingual phases. The process is thought to create various forms of speech positioned between the typically monolingual forms of speech, intermediary stages that 'represent points on a continuum from one language to another' (Haugen 1972: 303).

It is often assumed that extensive 'borrowing' may lead to 'language suicide'. Having correctly noted that, when a language dies, it is because another language has gradually ousted the old one, Jean Aitchison claims that speakers of the old language may continue speaking it, 'but will gradually import forms and constructions from the socially dominant language, until the old one is no longer identifiable as a separate language'. The author adds: 'This is in reality an extreme form of borrowing. The language concerned seems to commit suicide. It slowly demolishes itself by bringing in more and more forms from the prestige language, until it destroys its own identity' (Aitchison 1991: 198).

However, not even extensive structural copying has been shown to change the status of a Basic Code, i.e. to cause its replacement by a new Basic Code. It is not obvious that an A code characterized by extensive copying from B turns into B as a next stage. No-one has presented solid empirical evidence to support a scenario of code shift through gradual structural development. Heavy code-copying as such has not been shown to be the reason for shift, nor an unequivocal omen foreboding code replacement.

### *Positive effects of copying*

Contrary to claims concerning 'death by borrowing', etc., there are also ideas to the effect that codes which do not accept any copies—including structural copies—are more likely to die sooner than the ones that accept loans of all sorts; see, for example, the conclusions in Hamp (1989). It is difficult to believe that languages that do not accept borrowings are more prone to death; at least it cannot

mean that isolated languages are predetermined to die due to their isolation. On the other hand, the negative results of purism in communities speaking endangered languages is a well-established fact. Strongly dominated codes in areas of intense communication with dominant codes may survive owing to their openness and susceptibility to copying, which enhances their functionality (Csató 1999). Copying in general fulfills the communicative needs of the speakers. It is a means to create what is needed for easy communication, i.e. in the sense of an internal need, not as an external constraint. Note also that a strongly dominated language usually has varieties exhibiting different degrees of copying according to the nature of the given communicative situation.

According to Slobin (1986), contact-induced syntactic influence has helped the Turkic languages enrich their available means of expression. Compare the recent literature on how languages may become richer in grammatical categories and other structures as the result of intensive copying (Schmidt 1985, Dal Negro 1998, Dorian 1999, Aikhenvald 2002a and 2002b). The question whether foreign influence may have positive effects is a complex issue which cannot be pursued further in the present contribution.

### *Basic Code shift*

The 'structuritis' fallacy is still encountered in more or less subtle forms. In the discussion on code-switching, i.e. the use of more than one code within the same flow of discourse, there have been attempts to predict several constraints on 'switches'. It has been argued that a shift of the Basic Code may be accomplished by gradual acquisition of more material from a Model Code. In particular, switching certain function markers from A to B has been thought to imply a shift of the Basic Code from A to B. For example, Myers-Scotton (1993 and later publications) has established the so-called 'System Morpheme Principle' in code-switching, predicting that all active 'system morphemes' will come from the 'matrix language' (= Basic Code). The intraclausal use of 'system morphemes' from a foreign code would thus mean a shift to that code, that is, in Myers-Scotton's terms, a 'turnover of the matrix language'. If, for instance, German and English alternate within the same flow of discourse, the use of certain German system morphemes in a given clause would make German the matrix language of that clause, even if its other morphemes are English. The title of one of Myers-Scotton's early contri-

butions to the so-called Matrix Language Frame Model is particularly telling: 'One way to dusty death: the Matrix Language Turnover Hypothesis' (1995).

The first versions of Myers-Scotton's model (1993) presented some problems with respect to the definition of 'system morpheme', which made it easy to produce counterexamples to the hypothesis. The author has recently refined her model, redefining the crucial notion of 'system morpheme'. The new version just predicts one specific type of 'system morpheme' to come from the matrix language (Jake & Myers-Scotton 1997, Myers-Scotton & Jake 2000). The refined hypothesis is much more compatible with empirical data and as such of utmost interest for future studies. The model is now also applicable in the same way both to code-switch data and to other contact data.

Nevertheless, it is an unproven shift scenario that an increase in copied materials itself may lead to a shift of the Basic Code. Several contact linguists have tried to draw a line of development from a monolingual variety of a code A to the monolingual variety of a code B.

In one of his innovative contributions to contact linguistics, Backus (1993) sketches an imaginable scenario in which a speaker of Turkish successively inserts copies of Dutch elements into a sentence meaning 'He also talks to women'. First, the verb *praten* may be copied when it has been established in the speaker's Dutch grammar. It may be accommodated and inserted as *praten yap-* 'to talk' ('to do talking'), where *yap-* is a Turkish verb meaning 'to do'. Since *praten* usually co-occurs with the Dutch preposition *met* 'with', the collocation *praten met* 'to talk with' may also be established. This in turn leads to copying a further lexeme as a complement: *vrouwen* 'women'. Next, the sentence *Ook met vrouwen praten yapıyor*, containing just one Turkish word, *yapıyor* 'does', is produced. The endpoint of the process would be the monolingual Dutch sentence *Hij praat ook met vrouwen*.

This is an interesting tentative scenario. Backus later mentions it as a hypothesis which he is, however, inclined to abandon, since he does not find the data to prove it. In a later publication (Backus 1996: 323-325), he prefers to talk about a sudden move from content-word insertion to Dutch clauses rather than a gradual shift with inserted constituents as an intermediate step ('a rather sudden jump from predominantly insertional to predominantly alternational C[ode] S[witching]', Backus 1996: 324).

There does not seem to be any empirical evidence for a Basic

Code shift through gradual structural development, e.g. for the idea that speakers with growing proficiency in a Model Code copy larger and larger parts of it and finally use whole copied clauses in their speech.

A code does not develop into the variety it has extensively copied from. Quantitative criteria such as the amount of copies in a clause are not decisive for the Basic Code assignment. Even drastic morphosyntactic frame-changes, including insertion of copied function markers, do not lead to Basic Code shift.

### *Heavy structural copying*

A code may change considerably by successive copying. Even features originally alien to it may habitualize, conventionalize and eventually oust older indogenous features or leave them with modified functions. Even high-copying codes, displaying excessive global and selective copying, may become established varieties. Language history offers innumerable examples of codes converging structurally in the direction of another code, i.e. coming to resemble it more than they did before contact. Various kinds of combined copying may produce considerable morphosyntactic changes. Permanent incorporation of new frame-changing copies may have far-reaching long-term results. Successive copying processes may have 'snowball effects', leading to considerable deviation from the original typological profile of the copying codes and to clear convergence with their Model Codes. This fact should not be ignored in synchronic studies, where morphosyntactic convergence is often dealt with as a superficial phenomenon.

In some work on contact linguistics, contact-induced morphosyntactic frame-changing has even been excluded as a theoretical possibility (Johanson 2002a: 35-37). Strong positions with regard to constraints on copying function markers may still be found. As the contact history of many languages shows, it is, however, possible for codes to copy function markers without turning into the Model Codes copied from. Does this well-attested fact hold for integrated 'borrowings' only, whereas it is irrelevant for code-switching? Certainly not. An element that may not occur in code-switching can hardly stabilize as a 'borrowing'. In order to be conventionalized, copies of function markers must have been used prior to conventionalization as more or less ephemeral 'momentary' or 'non-integrated' copies, i.e. as 'switches'.

When it comes to function markers, certain complex types with elaborated material shapes are relatively susceptible to copying. Less elaborated relators such as case markers and simple adpositions are usually old and less prone to copying. In many cases, the lexical sources from which the grammaticalization processes started are unknown. But in simple prepositions such as French *chez* and Swedish *hos* 'at, with, beside', the nominal cores, *casa* and *hus* 'house' are still identifiable, the original pattern being 'in the house of'.

### 'Mixing'

Copying may strongly affect the structural characteristics of the Basic Code. In many cases, contacts have been sufficiently intensive and long-lasting to produce grammars that seem to be almost totally modeled on foreign patterns. This heavy copying is often inadequately referred to as 'language mixing', which can easily cause misunderstandings, since it suggests an unarranged or disarranged mixture, contamination or fusion. A distinction between 'hybrid languages' and 'pure languages' is scientifically untenable. As Schuchardt (1884: 5) already stated, there is no totally 'unmixed' language. Foreign structural influence has occurred in the history of all languages.

Haugen (1972: 317) claims that in the world of the bilingual 'anything is possible, from virtually complete separation of the two codes to their virtual coalescence'. The latter may be true of mental operations that lead to cross-linguistic fusion of certain procedures, when bilingual speakers choose to apply the same communicative strategies in both codes. It might even imply inability of individual speakers to keep the linguistic systems apart. The two codes themselves, however, are normally still identifiable as such and do not grow into one body.

### *Preserved genealogical affiliation*

This also means that the development of contact-influenced codes, as a rule, follows a single genealogical line. Even in spite of the sizeable influence of copies, each code can be identified as genealogical descendant of one specific 'parent', that is, a single 'mother' code, and not as descendants of the influencing codes.

Claims to the effect that the traditional family tree (*Stammbaum*) model of linguistic relationship cannot be upheld at all, since

all languages have been subject to considerable influence in the course of their history, suffer from 'a confusion about the nature of genetic relationship as a historical hypothesis' (Thomason & Kaufman 1988: 3).

It is true that copying may lead to considerable divergence from genetically close 'sister' codes which do not copy from the same Model Codes. A code may be influenced to such an extent that its genealogical origins are difficult to recognize. It may also become increasingly difficult to distinguish between nativized copies and non-copied native elements. There may be complicated cases of 'intertwining' which make the identification particularly difficult. But even if a code may become very similar to a foreign code by heavy copying, its identity is mostly maintained. Even massive morphosyntactic copying of combinational properties and function units is not believed to have led to genealogical mutation. Codes may be replaced by the strong codes from which they have extensively copied from, but this is then, as stated above, due to sociolinguistic circumstances that cause the community to abandon its language rather than a result of a gradual quantitative development.

### *Possible exceptions*

Nevertheless, the traditional model of genealogical relatedness and change is often claimed to be inapplicable to certain languages. These are said to be 'mixed' in the sense that they cannot be genealogically classified, because they are not derived from a single code. Since their elements can be traced back almost equally from two source languages, they are thought to have two 'parents'.

Examples include pidgins, creoles and in-group codes created to set their speakers apart from other groups. Thus, according to Thomason (2001: 227), languages such as Tok Pisin and Jamaican Creole English, which go back to the speech of semi-speakers who have not learned 'the full language', cannot reasonably be claimed to be varieties of English.

The possibility of 'double parenthood' as an exceptional linguistic relationship will not be discussed here. Several cases of an allegedly 'mixed' origin do not seem quite convincing. For example, Ma'a, a variety of the Mbugu language spoken in Tanzania, has been classified as 'mixed' since it displays a predominantly Bantu grammar but has copied a considerable part of its lexicon from neighboring Cushitic languages. The special lexical items of this in-group code do

not seem to justify the claim regarding a double ‘parenthood’ (see Mous 1994 and Dixon 1997: 11-12).

It should, however, be stressed that the *argument* presented for the exclusion from a family tree—an origin in groups that have not learned ‘the full language’—also disqualifies many genuine daughter codes (genealogical dialects) which have developed and reached stability in spite of a structure deviating from their mother codes.

### *High-copying Turkic varieties*

Examples of high-copying codes include a number of Turkic varieties, strongly influenced by Indo-European languages, e.g. Karaim in Lithuania and Ukraine, Gagauz in and around Moldova, Azerbaijanian, Khalaj, and other Turkic languages spoken in Iran, Salar and Yellow Uyghur in Western China, etc. All are geographically peripheral Turkic languages, displaying high degrees of innovation through code-copying from genetically unrelated and typologically different languages.

On the other hand, several non-Turkic varieties have been strongly influenced by Turkic. The impact of southeastern Turkic on the eastern variety of Modern Persian, nowadays called Tajik, has been so substantial that the latter has been claimed to be developing into a Turkic language (Doerfer 1967: 57). So far this has not occurred. While the predecessors of Uzbek and Tajik have interacted intensely for centuries in terms of copying, the modern languages are still not mixed or fused.

Substantial parts of the lexicon may be copied without code shift. A well-known case of extreme lexical copying is Classical Ottoman Turkish with its abundance of elements going back to Persian-Arabic originals. Its grammar was almost entirely Turkic, but it copied most of its vocabulary from unrelated languages. A corresponding case is found in the highly Romanized registers of English, which rival Ottoman in their wealth of vocabulary.

There are also cases of ‘relexification’ in which speakers of a code  $A_1$  have taken over their former  $B_2$  as  $B_1$ , but at the same time extensively imposed copies of lexical units of their former  $A_1$  on the new morphosyntactic frame  $B_1$ . The Eynu varieties, spoken in Eastern Turkestan, represent such special cases of extremely heavy global lexical copying (Hayasi 2000). As speakers of an Iranian primary code, they have taken over the dominant Uyghur language as their new primary code, imposing a high number of lexical units of their former primary code on it. Nevertheless, there

is no reason not to regard the Eynu variety as genealogically belonging to Uyghur.

*The least copiable elements*

The genealogical affiliation of a high-copying code may be determined by means of the elements that are generally least susceptible to being replaced. The least copiable function units seem to be grammatical markers with highly generalized, schematic, non-specific semantics and with reduced shapes, e.g. case and aspect-tense markers and adpositions with generalized meanings, copulas, certain pronouns and auxiliary pro-verbs such as 'to do'. Such elements allow us to identify texts in high-copying codes as representing a code A rather than a variety of a dominant code B. For example, in spite of the extreme degree of Arabic and Persian lexical impact on Ottoman Turkish, it is possible to identify any Classical Ottoman text as Turkic—and not as Arabic or Persian. Susan Gal (1979: 81) states that a sentence in a certain Hungarian variety spoken in Austria continues to function as a Hungarian sentence regardless of the number of foreign words contained in it, as long as the affixed grammatical elements—the number and case markers, person and tense markers, articles—remain Hungarian.

*Structural attrition*

Attrition, another linguistic corollary of decline, is a problematic point. It may manifest itself by stagnation in the acquisition of the A<sub>1</sub> code, failure to acquire and use certain linguistic devices, loss of elements, and fading norms. Sarah G. Thomason claims that attrition is the only type of change that is exclusive to language death (Thomason 2001: 230).

Structural attrition in the sense of *loss of structure* in phonology, morphology and syntax may certainly belong to the symptoms preceding and accompanying the cessation of the vital functions of a code. Typical features of attrition are reduction of rule-governed alternation by analogous generalization of one variant, regularization of paradigms such as irregular plural formation or verb inflection, merging or elimination of morphosyntactic categories and tendencies to replace morphologically complex constructions with analytic constructions (Thomason 2001: 230).

*Loss of structure in viable codes*

Loss of structure in this sense is, however, far from unique to 'language death' and cannot be said to forebode or forecast it. Changes of these kinds are found in many viable languages.

For instance, we find loss of structure in early Modern Persian as compared to Middle Persian (Utas 1991). The earliest Modern Persian differs from Sasanid Middle Persian, among other things, in its morphological regularization, e.g. a simplifying restructuring of the verbal system. The leveling in question may have been connected with increasingly close contacts with Turkic, though the exact Central Asian Turkic influence on the older forms of Modern Persian is still unknown (Johanson 2002a: 151).

German influence on Swedish is an instructive example of radical simplification. In the thirteenth century, the north German Hansa league established contacts in Sweden, including a considerable immigration of Germans to several Swedish towns. Their language, Middle Low German, had a strong impact on Later Old Swedish, which led to considerable changes towards the end of the fourteenth century, reaching a peak in the fifteenth century. There was massive copying of Low German elements, particularly innumerable lexical items.

But the influence also led to dissolution of native systems of grammatical forms. Classical Old Swedish had possessed a rich inflection of nouns, adjectives, pronouns and certain numerals: two numbers, three genders, four cases. The fourteenth century saw a development towards a simpler system. Less common case endings disappeared: for example, the ending *-s* was used instead of *-ar* in genitive singular, and the endings of dative singular (*-i* for masculine, and *-u* for feminine) were omitted. The simplification accelerated in the fifteenth century. Most old case endings disappeared. Nouns were declined in nominative and genitive only. In the adjective inflection, most case endings disappeared as well; there was a similar development in pronouns.

Most researchers agree that these developments were a result of the strong Low German influence in the population of the Swedish towns from the middle of the fourteenth century on. The outcome has been depicted as a 'mixed language', and therefore as a case of 'dissolution', typical of a period of cultural decay. But the loss of structure obviously did not forecast or lead to 'language death'.

*'Borrowing', 'import', 'transfer'*

How would it be possible for foreign structures to enter into a code and infect it with 'structuritis'? My answer is that it is not possible. I strongly reject the kind of realism that views linguistic contact as buying and selling, import and export. I claim that ideas of allegedly detrimental, code-killing effects of contact-induced influence are only the logical consequence of viewing the influence as 'borrowing', 'import', 'transfer', etc. These metaphors of traditional contact linguistics are highly misleading. They suggest that elements of 'donor' codes are taken over as such in 'recipient' codes.

However, given the structural economy of the individual codes, extensive 'borrowing' or 'import' of original elements from foreign—also typologically different—sources would not be possible, except as quotations. It would indeed lead to 'structuritis' and structural death. The objection to the traditional metaphors are by no means trivial, since their use often has undesirable consequences, obscuring the fact that there is no identity involved.

*Copying*

In the Code-Copying Framework, which cannot be presented in detail here, contact-induced linguistic change is conceived of in a different way. The concept of copying is claimed to account more adequately for the phenomena traditionally described as 'borrowing', etc.

The problems sketched above disappear when the elements inserted into a given contact-affected code are not taken to be *identical* with the originals, but just *copies* of them. Copies, even the most 'faithful' ones, can be distinguished from their originals. Thomason remarks that I am right in arguing for the replacement of 'borrowing' with 'code-copying', since "the source language does not give anything up, and the receiving language does not give a 'borrowed' item back" (Thomason 2001: 96). It is certainly true that the elements copied are not lent or given away. They remain in the Model Code, which is thus not deprived of anything. This would, however, be a most trivial and pedantic reason for suggesting a new term. The decisive reason is, of course, the need to stress that no elements are simply handed over from one code to another. The Basic Code does not take over anything identical with an element in the Model Code. The copy is never the same as the original. What is inserted into the Basic Code is an imit-

ation of an element of the Model Code, and this imitation is adapted to the Basic code.

These views are based on the conviction that each code possesses structures that are not directly compatible with others. This does not mean, as the vulgar polemics against structural principles will have it, coherence in the sense of Antoine Meillet's known dictum 'un système où tout se tient'. It does not imply 'a rigidity which is not characteristic of human behavior' (Haugen 1972: 303).

The assumption is a variability that allows extensive copying of new elements into a given system. The categories of each system are defined by largely idiosyncratic relations. In principle, no element corresponds directly to an element of another code. Similarities are approximate. Even relatively congruent elements exhibit intercodal differences. Each contact-induced change has consequences for the system of the Basic Code. Each developmental stage has its specific structural properties and regularities.

### *Code-Copying*

To make the argumentation more comprehensible, some main tenets of the Code-Copying Framework will be summarized here. For details, see Johanson 1992, 1993, 1999a, 1999b, 2002a, and 2002b.

As in all work published in this framework for the last decade, the term 'code' is not simply used instead of 'language', but as a neutral label for various kinds of linguistic systems, specific creative techniques of linguistic production. The term thus covers 'language' in various senses of this term as well as 'variety' and 'dialect' (social, spatial, situational, etc.). It is not, however, used in general meanings such as 'mode of speaking', 'mode of interaction', or 'communicative style'. Compare analogous terms such as 'code-switching', 'code-mixing'. The process by which a selected code translates mental contents into messages is called 'encoding'. The process of recovering the content conveyed from its coded form is referred to as 'decoding'.

The central concept of the framework is that copies of lexical, phonic or grammatical elements of a Model Code are inserted into clauses of a Basic Code. Copies can be more or less habitualized and conventionalized, thus ranging from ephemeral insertional switches to established linguistic changes.

### *Global copying*

GLOBAL COPYING means insertion of a copy of whole morphemes or morpheme sequences, whole 'globes' of material, semantic, combinational and frequential properties, into the Basic Code.

Examples: Copies of lexical units: English *edible* ← Late Latin *edibilis*, Swahili *kiplefiti* 'traffic island' ← English 'keep left', Zulu *ithisipunu* 'teaspoon' ← English *teaspoon*, German *live* (adverb) ← English *live* (adjective, adverb), German *joggen*, Swedish *jogga* ← English *jog* (verb), German *downloaden* ← English *download*. Derivational morphemes: English *-able* (in *eatable*, etc.) ← Old French *habile*, *-able* < Latin *-abilis*, English *-er* (in *driver*, etc.) ← Latin *-arius*, English *-ee* (in *employee*, etc.) ← Old French *-é* < Latin *-atus*. Function marker: German *in puncto* + genitive ('in the point of') 'as regards'.

Global copies are subject to adaptation (or accommodation) with respect to their material, semantic, combinational and frequential properties. Examples of material (phonic or phonological) adaptation: German *live* [laɪf] ← English *live* [laɪv], German *joggen* [ˈtʰɔgən], Swedish *jogga* [ˈjɔga] ← English *jog* [dʒɔg]. Speakers of Pashto replace [v] in global copies by [w], etc.

Morphological adaptation allows the global copies to fit in with the morphosyntactical frame of the Basic Code, e.g. prepares them to assume grammatical morphology. Copies of verbs are often accommodated by means of word formation devices, e.g. Turkish of the Netherlands *kijken yap-* 'to look' ('to do looking') ← Dutch *kijken* 'to look', with the Turkish verb *yap-* 'to do', which integrates the copied unit in the Turkish verbal morphology. Other forms of morphological reshaping are illustrated by, e.g. English *handicapped* 'suffering from a handicap, disabled' ⇒ German *ge-handicap-t*, Swedish *handikappad*, German *downgeloaded* ← *downloaded*, Zulu *ithisipunu* ← *teaspoon* with *i-* for singular and *amathisipunu* 'teaspoons' with *ama-* for plural.

Semantic adaptation means that certain semantic properties of the original are replaced. No global copies will be totally identical in meaning to their originals, probably not even copies filling lexical gaps. Examples: German *joggen*, Swedish *jogga* 'to run at a jogtrot' ← English *jog* 'to push, stir up, shake [etc.]', German *live* 'presented at the time it is happening, actually performed, not filmed or taped' ← English *live* 'being alive, vivid, full of life, fresh, flowing freely, etc.'

Combinational adaption means that certain combinational properties of the original are replaced. This restructuring affects the com-

binability of copies and the internal organization of complex copies. Examples of combinability: English plural *cakes* ⇒ German *Keks*, Swedish *kex* 'biscuit, cookie' (singular). German *live* is only used as adverb, whereas English *live* is also used as an adjective.

Frequential adaptation means that certain frequential properties of the original are replaced, i.e. there is increase or decrease in frequency of occurrence. For example, some Macedonian derivational suffixes (such as *-davi* as a nomen agentis suffix, expressing professionals) are more productive than their Turkish originals.

### *Selective copying*

SELECTIVE COPYING means insertion of copies of selected structural properties—material, semantic, combinational and frequential properties—onto units of the Basic Code. Only parts of 'globes' are selected for copying. These features owe their existence to foreign patterns, although they do not occur in globally copied units. Properties typical of units of the Model Code are copied onto equivalent units of the Basic Code. This kind of contact-induced influence produces 'loan phonology', 'loan syntax', 'loan semantics', etc. Again, there is no identity between originals and copies, but adaptation to the system of the basic code.

In the case of material copying, phonic or phonological structures (sounds, phonotactic patterns, accent patterns of stress and pitch, etc.) of Model Code units serve as models. For example: contact-induced change of *r*-sounds: from lingual [r] to uvular [ʀ], an innovation once spreading from Parisian French to numerous European vernaculars.

Selective copying can also be restricted to nonmaterial aspects: semantic, combinational and frequential aspects.

Semantic copying causes changes in content at the denotative or connotative level. The semantics of units of a Model Code is copied onto equivalent units of the Basic Code. This influence manifests itself as so-called 'loan semantics' and produces semantic calques and 'loan translations', e.g. German *Stern* 'famous actor, singer, etc.' ← English *star*.

Combinational copying causes changes in combinational properties of units of the Basic Code. This influence manifests itself as 'loan syntax' and also as 'loan morphology', since it may affect the internal constituency of units. It produces changes in the organization of grammar, alters grammatical structures by modifying construction

types and categories or by creating new ones. It may create new placement patterns, phrase structure rules, rules for lexical subcategorization, valency patterns, word-internal morphemic patterns, etc. The internal combinational pattern of a complex of the Model Code may be copied onto equivalent units of the Basic Code. Head-initial patterns may be copied onto head-final constructions, and vice versa. Dependents that are normally placed in front of or after their heads may become postpositive or prepositive, respectively, e.g. OV > VO. A new Swedish combinational pattern *när det kommer till...* (instead of *när det gäller*) has been copied from English *when it comes to...* Swedish singular nouns referring to collectives may combine with plural adjectives, e.g. *Polisen är effektiva* (instead of *effektiv*) 'The police are efficient', again as a selective copy from English.

Semantic-combinational copies are traditionally called 'calques' or (in the lexicon) 'loan translations'. Examples: English *beforehand* ← Old French *avant main*, Swedish *ladda ner* ('load down') ← English *download*, *uppdatera* ('date up') ← *update*.

Combinational and semantic properties of specific units of the Model Code can be copied onto corresponding units of the Basic Code in order for them to serve as function markers. One frequent type consists of a nominal core conveying a specific content plus a case marker or a simple adposition functioning as a syntactic anchor, e.g. Old French *par cause de* ('by reason of') ⇒ Middle English *by cause of* > Modern English *because of*, English *in terms of* ⇒ Dutch *in termen van*.

Frequential copying causes changes in frequency patterns. The frequency of use peculiar to an element of a Model Code is copied onto equivalent elements of the Basic Code. The influence manifests itself in increase or decrease in frequency of occurrence. A certain word, sound, combinational pattern, etc. which already exists in the Basic Code may gain or lose ground and become less or more marked. The German adverb *mittlerweile* seems to be making progress under the influence of English *meanwhile*, which has a similar shape.

### *Copying directions*

There are two possible copying directions between a socially dominated primary code  $A_1$  and a socially dominant secondary code  $B_2$ :

- (i) ADOPTION or TAKE-OVER. When a copy from  $B_2$  is inserted into a

speaker's  $A_1$ , it is said to be *taken over* or *adopted*.  $A_1$  is the Basic Code, and  $B_2$  is the Model Code.

Example: As a native speaker of Swedish, I take over global and selective copies (e.g. lexical and syntactic copies) from the dominant language English in my variety of Swedish.

(ii) IMPOSITION or CARRY-OVER. When a copy from  $A_1$  is inserted into a speaker's  $B_2$ , it is said to be *carried over* or *imposed*.  $B_2$  is the Basic Code, and  $A_1$  is the Model Code.

Example: As a native speaker of Swedish, I carry over global and selective copies (e.g. lexicon, pronunciation, syntax) from my Swedish into my variety of the dominant language English.

### *Imposition without shift and death*

In the literature on historical linguistics and contact-linguistics, imposition is mostly dealt with as 'substratum influence' (e.g. Thomason & Kaufman 1988) or 'shift-induced interference' (Thomason 2001). However, imposition is not necessarily connected with shift. It does not only occur in situations where speakers have abandoned a dominated code and shifted to a dominant code. Imposition is also common under stable bi- or multi-codal conditions, with parallel use of codes. Furthermore, code shift may in principle occur without any appreciable imposition from the abandoned code.

There are often long-term bidirectional copying processes in which codes converge at increasing rates due to combinations of adoption and imposition. An example of this is the relationship between Uzbek and Tajik in Central Asia. Imposition is indeed possible without shift and 'language death'.

Moreover, even if a code causing carry-over influence has vanished, the result is not necessarily substratal. 'Stratum' notions (sub-, super-, ad-) concern socio-political stratification. An example of historical superstratal carry-over copying is the Frankish imposition on Gallo-Romance.

### *Conclusions*

In sum, copies are by definition never identical with their originals. Neither global nor selective copies are true replicas of their models, but are always adapted to the system of the Basic Code. Similarities are always partial. Only in this way is it possible for

codes to copy a wide range of elements, even those that seem to be typologically inconsistent with the rest of their structures.

Lexical copies differ more or less from their originals, a natural consequence of the fact that languages organize their lexica in different ways. Copies of grammatical units and combinational patterns are often integrated in ways quite different from their behavior in the Model Code. Conventionalized copies are subject to internal processes of the Basic Code, which may lead to further differences. The outcome of copying can thus be highly creative. This is especially true of high-copying codes displaying strong snowball effects with respect to structural changes.

It is difficult to discern 'mechanisms' according to which language contact and multilingualism as such would necessarily threaten and endanger dominated codes. Languages do not cease to exist for structural reasons, e.g. because they have become overwhelmed by foreign elements. Structures are not 'borrowed', 'imported' or 'transferred' in contact situations; this would indeed lead to 'structuritis'. Elements of the Model Code are just copied, and the copies are adapted to the Basic Code. This also allows copying of typologically rather different elements. Even heavy contact-induced code-copying is thus not tantamount to 'mixing'.

Nor does heavy code-copying cause code replacement and death. Structural code-copying is not a linguistic factor on which rules for maintenance and shift can be based. If there is a connection between multilingualism and language endangerment, which is certainly the case, the decisive factors are social in nature: weakening under the pressure of strong codes, negative attitudes, restriction of domains, weakened interest in the acquisition.

Codes 'die' because they are no longer acquired by new generations. Even considerable degrees of structural change do not prevent them from being passed on from one generation to the next. Languages do not die of 'structuritis'.

*Address of the Author:*

Johannes Gutenberg University, Mainz <johanson@mail.uni-mainz.de>

*Notes*

\* The present article is a revised version of a talk given at the *Workshop on Multilingualism and Language Endangerment*, convened by Dafydd Gibbon,

Geoffrey Haig and Claudia-Maria Riehl at the Annual Conference of the German Linguistics Society (Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft), Mannheim, Germany, 27 February – 1 March 2002. Work on the article was carried out while the author was Visiting Fellow at the Research Centre for Linguistic Typology, Institute for Advanced Studies, La Trobe University, Melbourne, and Guest Scientist at the Max Planck Institute for Evolutionary Anthropology, Leipzig. I am grateful to Alexandra Y. Aikhenvald for most insightful comments. I also thank two anonymous referees for valuable suggestions.

### *Bibliographical References*

- AIKHENVALD, Alexandra Y. (2002a), "Traditional multilingualism and language endangerment", in BRADLEY & BRADLEY (2002: 24-33).
- AIKHENVALD, Alexandra Y. (2002b), "Language obsolescence: Progress or decay? The emergence of new grammatical categories in 'language death'", in BRADLEY & BRADLEY (2002: 144-155).
- AITCHISON, Jean (1991), *Language change: progress or decay?*, Cambridge, Cambridge University Press.
- BACKUS, Ad (1993), "Bilingual norms and linguistic change", unpublished conference paper.
- BACKUS, Ad (1996), *Two in one. Bilingual speech of Turkish immigrants in the Netherlands* (Studies in Multilingualism 1), Tilburg, Tilburg University Press.
- BRADLEY, David & Maya BRADLEY, eds. (2002), *Language endangerment and language maintenance*, London, RoutledgeCurzon.
- CAMPBELL, Lyle & Martha MUNTZEL (1989), "The structural consequences of language death", in DORIAN (1989: 181-196).
- CSATÓ, Éva Á. (1999), "Should Karaim be 'purer' than other European languages?", *Studia Turcologica Cracoviensia* 5: 81-89.
- DAL NEGRO, Sylvia (1998), "Spracherhaltung in der Beiz – das Überleben von der Walsersprache zu Pomatt/Formazza", *Wir Walser* 36: 13-16.
- DIXON, Robert M. W. (1997), *The rise and fall of languages*, Cambridge, Cambridge University Press.
- DOERFER, Gerhard (1967), *Türkische Lehnwörter im Tadschikischen* (Abhandlungen für die Kunde des Morgenlandes 37, 3), Wiesbaden, Steiner.
- DORIAN, Nancy C. (1981), *Language death: The life cycle of a Scottish Gaelic dialect*, Philadelphia, University of Pennsylvania Press.
- DORIAN, Nancy C., ed. (1989), *Investigating obsolescence. Studies in language contraction and death* (Studies in the Social and Cultural Foundations of Language 7), Cambridge, Cambridge University Press.
- DORIAN, Nancy C. (1999), "The study of language obsolescence: stages, surprises, challenges", *Languages and Linguistics* 3: 99-122.
- GAL, Susan (1979), *Language shift. Social determinants of linguistic change in bilingual Austria*, New York, Academic Press.

- HAMP, Eric P. (1989), "On signs of health and death", in DORIAN (1989: 197-210).
- HAUGEN, Einar (1972), *The ecology of language*. Essays by Einar Haugen. Selected and introduced by A.S. DIL, Stanford.
- HAYASI, Toru (2000), "Lexical copying in Turkish: The case of Eynu", in A. GÖKSEL & C. KERSLAKE (eds.), *Studies on Turkish and Turkic languages* (Turcologica 46), Wiesbaden, Harrassowitz: 433-439.
- JAKE, Janice L. & Carol MYERS-SCOTTON (1997), "Codeswitching and compromise strategies: Implications for lexical structure", *International Journal of Bilingualism* 1: 25-39.
- JOHANSON, Lars (1992), *Strukturelle Faktoren in türkischen Sprachkontakten*. (Sitzungsberichte der Wissenschaftlichen Gesellschaft an der J. W. Goethe-Universität Frankfurt am Main 29, 5), Stuttgart, Steiner.
- JOHANSON, Lars (1993), "Code-copying in immigrant Turkish", in Guus EXTRA & Ludo VERHOEVEN (eds.), *Immigrant languages in Europe*, Clevedon / Philadelphia / Adelaide, Multilingual Matters: 197-221.
- JOHANSON, Lars (1999a), "Frame-changing code-copying in immigrant varieties", in Guus EXTRA & Ludo VERHOEVEN (eds.), *Bilingualism and migration* (Studies on Language Acquisition 14), Berlin/New York, Mouton de Gruyter: 247-260.
- JOHANSON, Lars (1999b), "The dynamics of code-copying in language encounters", in Bernt BRENDAMOEN, Elizabeth LANZA & Else RYEN (eds.), *Language encounters across time and space*, Oslo, Novus Press: 37-62.
- JOHANSON, Lars (2002a), *Structural factors in Turkic language contacts*, London, Curzon.
- JOHANSON, Lars (2002b), "Contact-induced linguistic change in a code-copying framework", in Mari C. JONES & Edith ESCH (eds.), *Language change: The interplay of internal, external and extra-linguistic factors* (Contributions to the Sociology of Language 86), Berlin, Mouton de Gruyter: 285-313.
- MOUS, Maarten (1994), "Ma'a or Mbugu", in Peter BAKKER & Maarten MOUS (eds.), *Mixed languages: 15 case studies in language intertwining*, Amsterdam, Institute for Functional Research into Language and Language Use: 175-200.
- MYERS-SCOTTON, Carol (1993) [1997], *Duelling languages. Grammatical structure in codeswitching*, Oxford, Clarendon Press for Oxford University Press.
- MYERS-SCOTTON, Carol (1995), "One way to dusty death: the Matrix Language Turnover Hypothesis", Paper presented at the Dartmouth Conference on Endangered languages, February 1995.
- MYERS-SCOTTON, Carol & Janice L. JAKE (2000), "Testing the 4-M model: An introduction", *International Journal of Bilingualism* 4: 1-8.
- SASSE, Hans-Jürgen (1992), "Language decay and contact-induced change: similarities and differences", in Matthias BRENZINGER (ed.), *Language death. Factual and theoretical explorations with special reference to East Africa*, Berlin, Mouton de Gruyter: 59-80.
- SCHMIDT, Annette (1985), *Young people's Dyirbal. An example of language death from Australia*, Cambridge, Cambridge University Press.
- SCHUCHARDT, Hugo (1884), *Slawo-Deutsches und Slawo-Italienisches*, Graz.

- SLOBIN, Dan I. (1986), "The acquisition and use of relative clauses in Turkic and Indo-European languages", in Dan I. SLOBIN & Karl ZIMMER (eds.), *Studies in Turkish linguistics*, Amsterdam/Philadelphia, Benjamins: 273-294.
- THOMASON, Sarah G. (2001), *Language contact*, Edinburgh, Edinburgh University Press.
- THOMASON, Sarah G. & Terrence KAUFMAN (1988), *Language contact, creolization, and genetic linguistics*, Berkeley/Los Angeles, University of California Press.
- UTAS, Bo (1991), "New Persian as an interethnic medium", in Ingvar SVANBERG (ed.), *Ethnicity, minorities and cultural encounters* (Uppsala Multiethnic Papers 25), Uppsala: 103-111.