Bernt Brendemoen, Elizabeth Lanza and Else Ryen (eds.)

Language encounters across time and space

Studies in language contact

Novus forlag – Oslo 1999
The dynamics of code-copying in language encounters

Lars Johanson, Johannes Gutenberg-Universität, Mainz

The Code-Copying Framework

The present article deals with various aspects of dynamics in language encounters in time and space: forces that produce linguistic movement or change in non-monolingual speech communities. The topic will be discussed within the Code-Copying Framework, which is intended as an integrated model for describing all kinds of language encounters, the different processes they trigger and the dynamics they produce.

The Code-Copying Framework is primarily a descriptive apparatus, aiming at high descriptive adequacy, which also means incorporating various kinds of deviations from canonical patterns. Since it clarifies the connections between different kinds of copying, it can serve as the basis for coherent analysis of linguistic contact phenomena in general and thus also for considerations of a more explanatory nature.


The aims of the approach may be briefly summarized as follows:
- to incorporate language contact in a broader linguistic and sociolinguistic framework,
The dynamics of code-copying in language encounters

- to describe all kinds of non-monolingual speech production in an integrated way,
- to allow dealing with linguistic contact processes as highly complex processes,
- to link different kinds of code-copying to each other by studying the co-occurrence of ‘global’ and ‘selective’ code-copying in the same framework,
- to demonstrate the creative aspects of code-copying phenomena,
- to connect synchronic and diachronic aspects in an apparatus that does not describe possible structures alone, but is dynamic in the sense of accounting for successive processes of change,
- to describe the new products that constantly emerge as a result of language encounters, through successive copying and conventionalization of copies,
- to account for the relation of individual copying events to contact-induced language change over time,
- to demonstrate the relevance of contact-induced influence to monolingual situations, and
- to typologize the processes involved to create a basis for typological comparison.

Some of these main points will be commented on in the following discussion. This presentation will not, however, allow us to elaborate on the specific relations of the Code-Copying Framework to other models of contact-induced development.

Kinds of language encounters and contact settings

What we call code-copying is a normal developmental process, a universal tendency of human language, reflected in ontogenetic development, e.g. in second language acquisition, and in diachronic change such as the emergence of contact varieties – also higher and lower varieties of one language – and other types of non-monolingual speech production, including the genesis and development of pidgins and creoles. The Code-Copying Framework is thus relevant for all the closely interconnected fields of synchronic and diachronic work represented in this volume.

To explain and possibly predict the complex phenomena of copying and development of copies, due attention must be paid to certain socio- and psycholinguistic aspects from which the structural problems cannot be separated. The developments largely depend on the environment, the contact setting in which the code interaction occurs. Social factors strongly influence the variational patterns in individuals, domains, subgroups and generations, the actual occurrence of code-copies in different varieties and the diachronic implications of contact situations. The effects of non-structural conditions on code-copying can also be described in the Code-Copying Framework outlined here.

Code-copying as a kind of code interaction

Non-monolingual contact situations allow discursive code interaction, abbreviated CI, i.e. the use of two or more linguistic codes in a single discourse. One kind of possible code interaction is intraclausal code-copying, CC, as distinguished from intraclausal code-alternation, ICA, and extraclausal code-alternation, ECA:

```
      CI
     /\  
    CA  ECA ICA CC
```

A common motivation for different kinds of code interaction is obviously the inclination to say something the way it comes most naturally or is most easily expressed. Code alternation means shifting from one code to another. Extraclausal code alternation implies an alternate choice, a change of code over clauses. Intraclausal code alternation produces what has been called 'plain code-switches', items from one code inserted as quotations into a clause of another code without being incorporated into it. Intraclausal code-copying, however, does not produce real switches in the clause and is thus not an alternational type of code interaction. The widely used term 'code-switching' is inadequate both for this phenomenon and as a general cover term for code interaction, since it suggests an alternation of elements belonging to different codes.

- Code-copying

The central concept of code-copying is that elements of one code are copied and the copies inserted into another code. Copies are not 'switched', not just juxtaposed to the elements of the other code, but adjusted to its structure in some way and to some degree, even if they are only used occasionally, as so-called 'momentary copies'. They are not 'embedded' foreign elements, but indeed part of the code they are inserted into. They are 'foreign' only in an
etymological sense, namely with respect to their origin. Degrees of structural adaptation, habituation and conventionalization are not criteria for the definition. Code-copying does not produce any fusion or amalgam of codes, nor any ‘mixing’ with two participating languages. Copies emanating from one code and appearing in another do not belong to the same system as their originals. Not even a high-copying code subject to heavy frame-changes (see below) is made up of two systems.

The term ‘interference’ for insertion of copies is inadequate because of its connotation of an unseemly development, of an undesired deviation from monolingual norms that causes impaired communication, e.g. the negative effects of an L₁ on the acquisition of an L₂. Terms of this kind, which have lost their originally neutral meaning, may contribute to prejudices against contact-influenced varieties as ‘mixed’ or ‘contaminated’. Code-copying does not imply ‘negative’ or ‘positive’ effects of non-monolingual speech production, but just refers to the insertion of elements copied from one code within the context of another without specifying the degree of acceptability at a given stage of development. For a discussion on questions of terminology, see also Skaaden, this volume.

Insertion of copies into a basic code: adoption and imposition

The code copied from is the model code, and the copying code is the basic code. The basic code may be the non-monolingual speaker’s primary code L₁ or a secondary code L₂. The primary code – often of emblematic value, but not necessarily used more than the secondary code(s) – is typically a ‘mother-tongue’. Secondary codes are normally used for external monolingual communication. When a copy is inserted into a speaker’s primary code, we say that it is taken over or adopted. When it is inserted into a speaker’s secondary code, it is carried over or imposed. The term copying is used to subsume both kinds of insertion.

Global and selective copying

One type of copied elements is referred to as units, consisting of free or bound morphemes or morpheme sequences, ‘blocks’ of material, combinational, semantic and frequentational properties. Free lexical units are typically open class items such as nouns, adjectives and verbs. Bound lexical units are derivative morphemes belonging to closed classes, e.g., Norwegian -bet and English -dom in frihet and freedom. Non-lexical units are free or bound function units of closed classes, e.g., Norwegian til ‘to’, Turkish -yA ‘to’ (dative marker). The other type of copied linguistic elements consists of properties of material, combinational, semantic or frequentational nature.

Copying of units is referred to as global copying, whereas copying of properties is called selective copying. In the case of global copying, the whole form and function of a unit is copied. Selective copying does not concern morphemes, but extraplated structural properties and patterns applied to indigenous units. This does not exclude material manifestations: the copied properties may also include phonics ones. The two kinds of copying processes are closely related and just differ in scope. The phenomena they produce are essentially similar and should be treated in the same paradigm.

It is important to keep in mind that, according to the model, neither global nor selective copies are identical with their originals. They are always adapted – phonologically, grammatically, semantically and frequently – to the system of the basic code.

Dominance relations producing linguistic dynamics

The dynamics in language encounters depends on asymmetrical dominance relations between a sociolinguistically dominated or ‘weak’ code A and a sociolinguistically dominant or ‘strong’ code B. Typically, A and B are co-area minority and majority languages, respectively, A being an intra-group primary code = Aₚ, and B an inter-group secondary code = Bₛ. Note that contact situations may be more or less asymmetrical, that they can involve more than one dominant code, and that dominance relations may fluctuate in the course of long-term contacts.

Dominance relations produce different kinds of linguistic dynamics. In normal non-monolingual speech communities – disregarding social networks that undergo fundamental reorganization and may lead to abrupt changes (see below) – a speaker of a dominated primary code Aₚ with a certain degree of competence in a dominant secondary code Bₛ may be confronted with two interacting kinds of dynamic influence with respect to the two codes, two kinds of dynamics producing movement:

- ‘B moves A’: adoption of copies. Aₚ is influenced, but maintained. The speaker’s secondary code brings about change in his/her variety of the primary code. Copies of elements in the model code B, e.g. lexical items, foreign structures, are inserted into the basic code A. In this case, the copies are, as stated above, said to be taken over, or adopted. Traditionally, this kind of copying is mostly referred to as “borrowing” and “calquing”.
Global copies

The copying is global when the copied elements are segmental – phonological-morphological-lexical-phraseological units. For this kind of copying, see, for instance, the material dealt with in the analysis of the Middle Low German lexical influence on Swedish translation literature (Schöndorf, this volume). The globally copied units may be complexes, often idiomatic units of cultural phraseology, conversational formulas, copied as unanalysed units, e.g. Turkish of Germany alles klar (= German) ‘OK’, colloquial Uzbek bozi moy (= Russian) ‘my God’. There are also collocational complexes in the form of constituents in which one copied unit (e.g. a verb) triggers copying of another unit (e.g. an object or an adposition). Compare the discussion of ‘lexical chunks’ in Backus, this volume. Lexical units may initially be part of globally copied complexes, e.g. idiomatic, stereotyped phrases, and later on copied as isolates. This order is practically always found in the case of bound derivational or relational units; they are first copied as parts of complex originals and may later be copied in isolation and used productively. Global copies are inserted into the slots of their ‘equivalents’ in the basic code. They assume the grammatical morphology of the basic code, free or bound function units. This way of putting copies into grammatical structures is neither a ‘switching’ from one code to another nor a ‘substitution’ or a ‘suspension’ of the model code grammar, since the basic code is the very frame for insertion.

The original of a global copy has certain material, semantic, combinational and frequentional properties. The copies are subject to adaptation with respect to these properties, i.e. to various modifications in the direction of the basic code. Properties of copies always deviate from those of the originals, though the differences may be more or less significant. There is material restructuring in the sense of phonological adaptation and other kinds of reshaping, and there is morphological reshaping in the sense of accommodation to let the copies fit in with the morphosyntactical structure. Thus, to be inserted as predicate cores, copies of verbs may be morphologically accommodated by some word-formation device, e.g. by means of ancillary ‘do’ + pro-verbs plus a standard nominal form of the verb, e.g. Norwegian Turkish yap- ‘do’ in kliare yap- (‘do climbing’) ‘climb’. Such devices allow verbs to be easily copied without the trouble with the idiosyncratic morphology that often goes with them. In exceptional cases, no special accommodation takes place, e.g. Norwegian Turkish kep-tim ‘I bought’, where the copy of a verbal stem is inserted directly (Brendemoen 1987). Other kinds of restructuring affect combinational properties, e.g. the internal organization of complex global copies, according to patterns of the
basic code, and may even involve rearrangement into more creative formations. Semantic restructuring is very common: not even copies filling lexical gaps will be totally identical in meaning to their originals. Finally, restructuring with respect to frequentational properties is observed.

Selective copies
The copying is selective when it does not involve all aspects of a segmental unit, but single properties of material, combinational, semantic or frequentational nature are copied onto units of the basic code. This influence produces ‘loan phonology’, ‘loan syntax’, ‘loan semantics’, etc. Again, it follows from our theoretical standpoint that original and copies are never identical.

Selective copying may be exclusively material in that phonetic aspects of model code units – segments and patterns typical of them – serve as models for imitation and are copied onto units of the basic code. One example is the occasional pronunciation of rt as a retroflex in Norwegian Turkish, e.g. in [doer] for dön‘four’ (Brendemoen 1992:19). For phonological code-copying in Karamaj, see Csató, this volume.

Semantic copying realigns the denotative or connotative content of units of the basic code, e.g. Norwegian Turkish (gazetteer) dur ‘be written (in the newspaper)’ = Norwegian står (i avisen) (lit ‘stand in the newspaper’). Similarity in shape between units of the basic code and their equivalents in the model code – i.e. between relatively homophonous ‘false friends’ – may favour semantic copying, e.g. Norwegian Turkish şans ‘luck, chance’ = Norwegian sjans ‘chance’, Swedish ladda ner (‘load down’) = English download, Swedish hardware = English hardware. Combinational copying concerns the internal constituency of a unit or its external combinability with other units of a word or a sentence, e.g. Norwegian Turkish corba ye- ‘eat soup’ = Norwegian spise suppe as against the normal corba sp- (lit. ‘drink soup’). Other Norwegian Turkish examples are bir içki al- (lit. ‘take a drink’) ‘have a drink’ = Norwegian ta en drink as against the normal bir içki sp- (lit. ‘drink a drink’). Compare the phraseological copying discussed in Schöndorf, this volume.

Semantic and combinational copies of these kinds are traditionally referred to as ‘calques’ and ‘loan translations’, but there are also other types. Semantic and combinational properties of function units can be copied onto units of the basic code. Thus, Karamaj -dəl ‘with’ displays uses resulting from copying functions of the Russian instrumental case (Csató 1998a). Copying of combinational patterns also includes word-internal and word-external changes: redistribution of morphemes in words and modified syntactic patterns such as constituent order and complement structures, e.g. Norwegian Turkish bir kimseyi sør- (with the accusative) ‘ask somebody’ = Norwegian spørre noen as against Turkish of Turkey bir kimseye sör- (with the dative). Copies of lexical units often tend to determine the phrasal choice in the basic code; thus, copies of verbs may influence case-marking, choice of adpositional phrases, etc.

Frequentational copying may lead to a shift in the use of existing structures, to increased or decreased use of units of the basic code. Interesting examples are found in the use of pronouns in Turkish Black Sea dialects under the probable influence of imposed copies of Greek patterns (Brendemoen 1993), and the use of ‘zero anaphora’ in Norwegian as used by immigrants from Turkey (Nistov, this volume). Combinational patterns, e.g. constituent orders, that already exist in the basic code but are more normal in the model code may gain ground and become less marked. Thus, shifts in word order are particularly common under foreign influence. Similarity in shape between units of the basic code and their equivalents in the model code may also favour frequentational copying.

Mixed copies
Apart from the pure types of global and selective code-copying, there are mixed types in the sense of a selective copy – e.g. of a combinational pattern – containing at least one global copy. In Norwegian Turkish examples such datasya vur ‘insert into the computer’ (lit. ‘hit into the computer’, with vur= Norwegian står ‘hit’), a combinational pattern is copied onto a construction which also contains a global copy of the lexical unit data (Türkler, this volume). This type has been called ‘loan blend’ when occurring in the lexicon, but it also comprises other kinds of copies. For example, a whole clause pattern may be copied together with a globally copied subjunction, e.g. a subordinative conjunction such as Iranian Azerbaijani vaaxti ki ‘when’ = Persian vaqi ki. Again, the pattern in question may already be present as an option in the basic code but increase its frequency under the influence of a foreign code. Mixed copies may be highly creative formations in which both the selective and the global copy differ essentially from their originals as a result of heavy restructuring.

Alpha and Beta lects
The Code-Copying Framework reckons with the emergence of functional contact varieties, new lects referred to as Alpha and Beta lects, respectively. An Alpha lect is a variety of a socially dominated code A, used for in-group communication with non-monolinguals who share the same linguistic background. Since varieties of A used in diaspora (immigrant) speech communities no longer represent
normal A', they are Alpha lects even if they are modest in code-copying. A Beta lect is a contact variety of the socially dominant code B, used by A speakers for communication with B monolinguals. Different stages of B acquisition and A development contribute to the input for Alpha and Beta lects.

Mutual influence of basic and model codes
The speakers' stages of proficiency in the basic and the model code influence each other. Different developmental stages of proficiency in the dominant code B, i.e. more or less advanced stages of B₂ acquisition, influence the copies adopted in their A₁ variety. Different stages of development of competence in the dominated code A determine their Alpha lect and also influence the copies imposed on their B₂ variety.

For questions concerning the analysis of the development of A₁ codes in B₂ environments, see Skadden, this volume. For different kinds of A₁ copies imposed on (transferred to) B₂ varieties in acquisition of Norwegian as a second language, see the contributions of Nistov and Ryen to this volume. Nortier, this volume, deals with A₁ patterns imposed on B₁ in the form of deletion of plural suffixes, prepositions and definite articles in the Dutch speech production of bilingual Moroccan Arabic/Dutch speakers.

Defining the A and B varieties involved
For comparison of originals and copies in terms of similarities and differences, the characteristic typological properties of the basic code and the model code must be defined.

As regards the basic code, it is subject to permanent change. First generations of diaspora communities speak Alpha lects that are close to the regional varieties they spoke in the communities left behind. Early Alpha lects may thus be rather heterogeneous, whereas later Alpha lects are often more levelled.

The model of the copies, the originals, and the variety they belong to must also be identified. The model for adopting copies in A is some B variety that is familiar to the speaker and usually reflected in his/her own B variety. Particularly at the beginning of a language contact, it may be a specific Beta lect used for superficial communication with B monolinguals. This stage is typical of first generations of immigrants. Their code-copies often seem to differ considerably from the originals, e.g. phonetically and semantically. The grammar seems reduced, the syntax simplified, or function units - gender and number inflections, articles, etc. - omitted. But such differences cannot be analysed in terms of adaptation unless it is clear from which B variety the copies are adopted. The very model may lack fuller structures, so that there is nothing to simplify or reduce. No element can be said to be 'lost' or 'deleted' if it has not been present in the original. Reduced Beta lects, e.g. varieties of so-called 'foreigner talk', are also the B models in formation of pidgins (see below).

With improved competence in B, speakers may copy elements from more native varieties, usually the ones reflected in more advanced stages of their B interlanguage. Speakers with a native or native-like command of B may have a complex input. The elements serving as models may belong to their own B varieties or to Alpha lects of previous generations, e.g. a 'home language' spoken by parents and/or grandparents. Early Alpha lects may thus exert influence long after they have disappeared.

Habitualization and conventionalization
Code-copying contributes actively to language change. Code-copies may develop along the two different lines of structural integration and conventionalization, i.e. (i) from structurally less integrated to more integrated copies and (ii) from momentary to habitualized and conventionalized copies. Structural integration - phonological adaptation, assumption of affixes of the basic code, etc. - and extra-structural integration - habitualization, conventionalization - are two kinds of development that should be kept apart, even in cases where they might seem to run parallel. The correlations of the degrees of development that global copies may reach along such different parameters should be determined carefully. To account for such developments, which are not essentially different from 'internal' developments of non-copied innovations, monolingual grammars with diachronic and variational components are required.

Code-copies may be momentary (MCC), i.e. sporadic, ephemeral instances of copying, the result of singular individual dynamic acts. They can also be habitualized and conventionalized. What starts as a performance phenomenon may have diachronic effects, i.e. lead to changes and new norms. More or less habitualized copies (HCC) occur frequently, regularly, or normally in individuals and/or in the speech community. Conventionalization is the integration with respect to acceptance in the speech community. More or less conventionalized code-copies (CCC) are the result of dynamic processes with more far-reaching and lasting effects on the role of the copy in the speech community.

The process of conventionalization - the way from MCC to CCC - is a continuum of changes in the sociolinguistic status with gliding transitions between degrees of acceptability. In order for our framework to be as descrip
tively adequate as possible, no absolute distinctions are posed in these terms, e.g. with respect to notions such as "borrowing".

As historical linguistics shows, conventionalization of copies is a highly normal phenomenon and does not occur under extreme conditions only. The conventionalization processes of global and selective copies follow similar rules. Deviations originally perceived as 'interference' may establish themselves, become socially 'unmarked' and even replace their indigenous equivalents. The conventionalization may be limited to small groups of speakers, or it may have a wider scope, to the extent that the copy in question becomes part of a more general bilingual norm. The final stage is reached with monolingualization of the copy, namely when its use no longer presupposes bilingual ability — proficiency in both codes —, but also occurs in utterances of monolinguals and is not perceived as 'foreign' by them.

It is important to stress the successive character of code-copying and the permanent dynamism involved in it. A conventionalized code-copy is part of a synchronic norm. Every such norm may be deviated from by new 'marked' copying. Each changed basic code is ready for further copying and provides the frame for new insertions and new conventionalization. The selection of what innovations are conventionalized is dependent on the basic frame already existing.

Thus, in principle, at every stage of development, the basic code must be defined anew, and the existence of new separate grammars must be posited. Code-copying involves rule-governed productive and creative processes, and the command of code-copying strategies is a component of the speakers' linguistic competence. Every new Alpha version has its own norms and should be dealt with as a code in its own right.

Previous and current processes

This means that knowledge of previous copying processes and their results is crucial for the analysis of current processes. Experience from language history may, for example, shed light on developments in modern diaspora varieties. Cases of established 'loans' may be useful in discussions on current code-copying. The conventionalization processes that have led to these items must once have started from momentary code-copies.

On the other hand, the results of the study of current language encounters may provide us with new insights into the complex processes of contact-induced language change and thus into the nature of essential linguistic mechanisms. The ongoing processes in immigrant languages of young diaspora speech communities offer unique chances to observe the dynamics of language encounters from the very beginning. Such observations can also be important for historical reconstruction and for the understanding of the ways synchronic systems have emerged. The study of what is more or less susceptible to copying and conventionalization has consequences for questions of reconstruction and genetic relations. Similarities in domains that are less susceptible to influence than others are arguments in favour of linguistic relatedness, e.g. within the much-debated linguistic stock of Altaic languages. Detailed knowledge of documentable processes may offer clues to the prehistory of languages and possibly to deeper-level genetic relationships.

Frame-changing developments

Global and selective copies of function units and selective copies of combinational patterns may produce considerable changes in the morphosyntactic frame of the basic code. In cases of adoption, B₃ is the model on which the morphosyntax of A₁ is restructured. In cases of imposition, A₁ is the frame-changing model imposed on a basic code Bₚ, which provides the morphosyntactic frame. But the two types of insertion differ from each other with respect to the relative importance of the types of copies.

Global copying of function units is a common phenomenon well attested in the contact history of languages. Grammatical relators particularly susceptible to copying are adpositions and clause junctors (conjunctions, relativizers), e.g. Danish *hvis* 'if' ≈ *hvæs* ≈ Middle Low German *wes*. On Turkic global copies from Iranian, Arabic and Slavic, see Johanson (1992:15, 108, 1993c, 1996, 1997, 1999). It cannot be claimed that copies of function units are impossible until they have been conventionalized as "borrowings", since all copies must once have been used prior to conventionalization.

Though the use of copied function units in the morphosyntactic frame of a basic code seems less restricted than sometimes predicted, equivalence positions are needed to insert the copies. If the two interacting codes are closely related, as in the case of English and Norwegian, there are relatively many equivalence positions, which allows a wider range of copying patterns. Again, similarity in shape between a unit of the basic code and it equivalent in the model code — a relatively homophonous 'false friend' — may favour copying, e.g. the new Swedish preposition *givet* 'knowing about, considering' (≈ English *given*) in examples such as *givet demna situation* 'given this situation' (without gender congruence between *givet* and *situation*).
Certain function units are generally easier to insert with respect to equivalence positions, e.g. coordinative relators such as and, which link relata of equal rank, without forming a constituent with any of them, and preferably occupy the position between them. Equivalence positions are easily found for discourse junctors such as and and but, whose typical insertion points are at boundaries of turns, sentences and clauses, and which are thus often copied (see Johanson 1997).

Copies of subordinating relators may be much more difficult to insert. Since they link a dependent to a head and form one constituent with one of them, they may cause problems in case of word order differences between the model code and the basic code. Their preferred position is mostly at the periphery of the dependent. There is thus a fundamental discrepancy between head-first and head-final languages. If the dependent is postpositive, the relator prefers the initial position, e.g. Norwegian for; fred, ‘for peace’, and if the dependent is prepositive, the relator prefers the final position, e.g. Turkish habırdı. In a language such as Norwegian with predominantly head-first patterns, subordinating relators are typically free units, i.e. prepositions and conjunctions. In languages of a head-final type, the subordinating relators are postpositions, case suffixes or other postposed particles or suffixes. This fact obviously causes incongruity. Prepositions display obvious similarities to case suffixes and postpositions in their use as function units, but their positions in relation to the dependent are different, e.g. English in front of the house, versus Turkish evin, ṭaninda, going back to constructions such as the front of the house and evin ṭın, in which front and on are the heads. Further examples are English because (of) = Old French par cause (de) ‘by reason of’. Norwegian i term er av = English in terms of, German in puncto ‘with respect to’. The most copiable classes include conjunctions and other units fulfilling crucial discourse functions. For copied auxiliaries see, see Johnson (1993c, 1996, 1997, 1999). The least copiable function units are the ones with highly generalized meanings (often in combination with a more reduced shape): certain case markers and adpositions, copulas and auxiliary pro-verbs (e.g. of the ‘do’ type).

Selectively copied combinational patterns
Selective copying of combinational patterns of a model code, often labelled as ‘grammatical interference’, has great impact in producing and triggering frame-changing innovations. In certain cases, however, it may be difficult to demonstrate that a given change is due to copying and not just an internal change: See the ‘general tendencies’ dealt with in Johnson (1992:114-117), and the discussion in Mønnesland, this volume, on the possible copying of an Albanian syntactic pattern in Serbian dialects (loss of the distinction motion vs. location with certain prepositions).
Selective copying of combinational patterns may affect rules of constituent ordering: head-final and head-first order, e.g. SOV versus SVO, genitive-noun (GN) versus noun-genitive (NG), adverbial placement etc. See the examples of copied German word order in early Slovenian and Swedish (Janko, Schöndorf, this volume); compare the critical discussion of the role of cross-linguistic word-order equivalence in Boeschoten, this volume.

Copying of purely combinational patterns is often preceded by global copying of function units and may start with mixed copies, i.e. copied combinational patterns including a global copy of a function unit. But direct copying of the order pattern itself - without previous global copying of function units - also seems possible. An ordering pattern typical of the model code is more easily copied if the type it represents is an alternative already present in the basic code. In this case, frequent copying may increase its use, e.g. the frequency of an SVO order as against a more normal SOV order. However, it also seems possible to copy ordering patterns that have so far been totally alien to the basic code.

Combinational copying may change the ordering patterns of the clausal frame to the effect that new equivalence positions for global copying emerge. This may even be the very reason for much combinational copying. Problems in using global copies in a given clausal frame may strengthen the speaker's ambition to create more congruency.

On the other hand, partial change of ordering rules, e.g. from head-final to head-first order and vice versa, may lead to disharmony within the basic code. For example, the change from a genitive-noun order to a noun-genitive order may remove the very production pattern for 'noun-like adpositions' of the types in front of the house, and evin, önünde. If an originally head-final language copies the reverse order noun-genitive, the old and very important grammaticalization pattern for postpositions of this type is no longer productive. For the disharmonic situation in Karaim, see Csató (1998a).

Long-term effects and convergence of the contact codes

Both global and selective copying lead to convergence of the basic and the model code. The frame of the former is progressively adapted to that of the latter with respect to morphosyntactic and semantic structures. Note, however, that structural adaptation reduces the degree of convergence and that every element that has become part of a given basic code is therewith automatically subject to the normal 'internal' processes of that code.

Language history offers innumerable examples of codes converging syntactically in the direction of another code, i.e. coming to resemble it more than they did before contact. This fact should not be ignored in synchronic studies, where syntactic convergence is often dealt with as superficial 'interference'. In some work on contact linguistics, contact-induced syntactic frame-changing is even excluded as a theoretical possibility.

Frame-changing patterns ease the tension between the codes and may reduce the speaker's planning effort in forming sentences in them. The reason may be common mental procedures for arranging information. This, on the other hand, does not mean fusion of the linguistic structures of the two systems themselves. The elements of the two frames always contribute to meaning in code-specific and idiosyncratic ways.

Permanent incorporation of new frame-changing copies may have strong long-term frame-changing effects. The emergence of new equivalence positions leads to further morphosyntactic convergence. Increasing similarity between the two codes may favour and accelerate the convergence. Successive copying processes may thus have veritable 'snowball effects'. It becomes increasingly difficult to distinguish between nativized copies and native, non-copied elements. Long series of changes produce copies which are established as part of the 'inherited' inventory and cannot easily be traced back to the original copying processes.

Consecutive changes of this kind may lead to considerable deviations from the features typical of the family of a code, properties which distinguish it significantly from the type of the genetic group it belongs to. It is also perfectly possible for a code to copy elements that appear 'typologically inconsistent' with the rest of its structure. If one particular language in a group of closely related languages shows radical changes in morphosyntactic devices, this is indicative of contact-induced frame-changing. In languages of the same type, the results of such processes may be rather similar. Thus, common tendencies are observed in different Turkic languages influenced by Indo-European languages, e.g. Karaim, Iranian Azerbaijani and Gagaüz (Csató 1994, 1996, 1998, Kiral 1997; Menz 1997).

Even if heavy copying from one model code may make the basic code very similar to it, the frame-providing code may be determined by means of the elements that are least susceptible to being replaced by copies, e.g. affixes and postpositions with generalized meanings such as case and tense-aspect markers, copulas, certain pronouns and auxiliaries such as 'do' pro-verbs. It is thus also
possible to identify the frame-providing code of texts in heavily influenced languages such as Classical Ottoman (Turkic) and Northern Tajik dialects (Iranian).

Heavy code-copying is consequently not tantamount to a shift of the basic code (code replacement). Structural changes do not change the roles of the codes with respect to their status as model code and basic code. Not even numerous copies of morphosyntactic elements – combinational properties and function units – in a high-copying basic code mean a shift of the frame-providing basic code. There is indeed no empirical evidence for a code shift through gradual structural development, e.g. for the idea that speakers with growing proficiency in the model code copy larger and larger parts of it and finally only use whole copied clauses in their speech. Code shift, which means that one basic code replaces another one, is a radical qualitative leap caused by loss of functional domains, and not the result of a gradual quantitative development within the basic code frame.

Settings
The typical setting for code-copying is a dominated non-monolingual speech community with strong external bonds to a dominant speech community, mostly a diaspora variety, which for some reason has been separated from the main bulk of speakers in the core area of the language in question. The asymmetric dominance relations cause one-directional dynamics: the non-monolingual situation leads speakers to adopt new linguistic habits. Dominated codes of this kind thus become more heterogeneous, display more variation than those developing in speech communities with weaker external bonds. The dominated speech communities may have various sociolinguistic profiles with different bearings on the outcome of the contact-induced influence. Important variables include:

- the degree of competence in the dominant code and the dominated code,
- the degree of social asymmetry in the setting, and
- the strength of internal bonds in the dominated speech community.

Young and older contact situations
The degree of competence in the dominant and dominated code often depends on the length of the language contact. The code interaction patterns change according to the speakers' command of the codes. Note, however, that several types of code interaction may coexist in one and the same individual or generation.

Diaspora speech communities in young contact situations usually exhibit unstable conditions with much variation and quick change. Members of first generations are usually 'A-dominant' speakers with a relatively low B proficiency. The early Alpha lects they speak with their fellow immigrants are based on A varieties of the regions they come from; they often serve as frames for insertion of rather numerous copies in strongly reshaped forms. Many of the copied lexical units are lexical gap fillers for new or modified concepts typical of the diaspora setting. Speakers with a somewhat higher B proficiency are sometimes observed to develop high-copying Alpha lects, often incorrectly assessed as 'fusions' of A and B.

Older contact situations, typical of generations born in diaspora settings, are usually more stabilized, displaying more balanced proficiency levels for A and B, often with advanced bilinguals, native speakers of both codes. There are two main options in such situations:

- develop towards removal of Alpha and shift to a Beta lect as the primary code;
- crystallization of a more stabilized Alpha lect as the primary code.

The way to code shift
The shift option means that groups of A speakers stop using the A code in favour of a more prestigious B code. Transitional stages are often observed from the very first generation of balanced bilinguals. There may be attrition and erosion of A, delay and stagnation in its acquisition, decrease in its use, which lets its norms fade (see, e.g. Skaaden, this volume). Increasingly fewer speakers use the language; children do not learn it at all, or only parly. There may be little code-copying, often because of a social stigma attached to it, and instead more preference for a different and less stigmatized kind of code interaction, ECA, i.e., extracausal code alternation, a switching between two basic codes yielding 'mixed discourse'. When the in-group use of Alpha is replaced by the use of ECA, the structure of the interactional pattern changes, but this is of course not tantamount to a structural change of the Alpha lect itself.

Shift is favoured by various sociolinguistic factors. There may be few social functions left for A, since B serves almost all purposes. The community of A speakers may have a relatively weak group identity and thus weak internal bonds of A solidarity. They may make few or no language maintenance efforts, and get little or no support from institutions of the 'host society'. Abandonment of Alpha lects is often a result of a relatively high degree of education, puristic awareness and strict separation of A and B.
The dynamics of code-copying in language encounters

These old copies typically consist of stabilized early selective copies of material and combinational properties, e.g. aspects of phonological and morphosyntactic organization. There may also be selective semantic copying, but global lexical copying, which is always involved in adoption of copies, is not necessarily present in cases of imposition. The fact that a given code has become more similar to another one in these properties, but not with respect to lexification, may thus be indicative of imposition.

On the other hand, there are also cases of relexification in which speakers of a code A1 have taken over a former B2 as B1, but at the same time extensively imposed lexical units of their former A1 on it by copying them globally into the new morphosyntactic frame. The Eymu varieties spoken in Eastern Turkestan represent such special cases of extremely heavy global copying. As speakers of an Iranian primary code, they have taken over the dominant Uyghur language as their new primary code and imposed a high number of lexical units of their former primary code on it.

It is important to bear in mind that, if dominance relations change, generalized Beta varieties, based on and developed from Beta lects, may also – just like generalized Alpha varieties – spread to original B1 speakers.

Areal code-copying

Areal processes of code-copying may be very complex and difficult to analyse. It may be extremely difficult to keep the results of adoption and imposition apart: on the one hand, copying from a B2 into an Alpha variety and, on the other hand, copying from an Alpha variety into a Beta variety as a 'substratum effect' due to shift. Note that widespread plurilinguism may be responsible for both kinds of influence and that both Alpha and Beta varieties may, as stated above, spread to original B1 speakers within a given area.

Serious complications are caused by the fact that dominance relations often undergo changes over time. The sociolinguistic role of a given code in a given linguistic area may change from being dominant to being dominated, and vice versa. Long-term contacts may thus imply varying dominance relations leading to combinations of adoption and imposition. When new patterns emerge due to contact-induced copying in dominant and dominated codes of a linguistic area, this convergence is in fact often the result of combinations of both processes.

This is the case with many complex plurilingual areal and local phenomena presented in this volume, e.g. the multidimensional interaction of varieties in a Cameroonian village (Engelsen, this volume) and the regional symbiotic development of Aramaic and Arabic (Wardini, this volume). Several types of

Imposition as substratum effect following code shift

When a dominated diaspora speech community shifts to the dominant code, properties of the former primary code do not necessarily leave any perceptible traces in the new primary code. But in many cases of code shift, a former restructuring of B2 on the model of A1 remains as a substratum effect in B1. This means that old copies of A1 elements once imposed on B2 continue their existence in B1.
influence are observed in the history of Turkish-Greek language contact in Anatolia as dealt with by Brendemoen (this volume) and exemplified with Cappadocian Greek, Karamanli Turkish, Pontic Greek and Pontic Turkish. At least in the case of Pontic contacts, we must reckon with long-term interaction of adoption and imposition processes under varying relations of social dominance. The role of the Greek varieties has changed from dominant to dominated. There are equally complex relations between Uzbek and Tajik in Central Asia or between Turkic, Fennic (Cheremisic, Permic, Mordvinic) and Russian in the Volga-Kama area.

Volga-Kama areal interaction
In the case of the Volga-Kama area, complex processes of mutual code-copying have led to convergence of socially dominated and dominant codes and introduced new linguistic patterns, partly typical of so-called 'Sprachbund' phenomena. The Turkic varieties involved are Chuvash, Tatar and Bashkir, the Fennic ones Mari, Votyak and Mordvin. Chuvash is an Oghur Turkic language, developed after the settlement of Oghur tribes in the region in the 9th century (at the latest). Tatar and Bashkir are Kipchak Turkic languages introduced into the region from the 14th century onwards. The Russian linguistic influence on the region increased rapidly from the middle of the 16th century on, i.e. after the fall of the Khanate of Kazan.

The Kipchak, Fennic and Russian varieties are commonly thought to be fundamentally influenced by Oghur Turkic, mainly in the phonology. Chuvash, particularly Upper Chuvash, displays phenomena due to close contact with Fennic: phonological, lexical and probably grammatical influence, mostly from Mari. The Fennic languages and Chuvash are heavily influenced by Kipchak, mostly Tatar, in lexicon and word-formation. All varieties of the area show strong Russian lexical and syntactic influence.

The linguistic nature of the language contacts has not yet been fully established. It is clear that the Kipchak lexical and morphosyntactic influence on Chuvash (most strongly Low Chuvash) and on Fennic is due to adoption of copies from a dominant code. The same is true of the Russian lexical and syntactic impact on the Fennic and Turkic varieties of the area.

However, many similarities are results of the complex processes of ethnolinguistic assimilation in the region since the 10th century, and thus due to imposition of copies following a shift to a socially dominant code. Early imposition of Fennic elements on Oghur is highly probable, since Oghur tribes early came to dominate the local Fennic groups on the left Volga bank and assimilated Votyak and Meadow Mari groups. The Fennic influence on Chuvash is a substratum influence due to assimilation of segments of a local Fennic population by Oghur immigrants. The influence is typically strongest in Upper Chuvash, especially in the Sundyr dialect spoken in the northwest of the Chuvash republic, in immediate Mari-speaking neighbourhood. The old Oghur impact is also likely to be a substratum influence, especially since Oghur groups on the left Volga bank were also assimilated by Kipchak groups from the 14th century on.

The Volga-Kama area thus displays effects of long-term contact processes of mutual influence, implying varying relations of social dominance. The socio-linguistic role of most of the codes involved has changed from dominant to dominated. The areal convergence is due to highly complex combinations of adoption and imposition processes.

A lingua franca as a dominant code
Areal interactions often imply the emergence of a koiné or a lingua franca, a common code shared by various groups of speakers of different primary codes as a medium of communication over larger geographical areas. Whereas copying has heterogenizing effects by creating more variation, the development of a lingua franca has homogenizing effects of levelling, simplification and regularization. If a given code functions as – or goes back to – a lingua franca, its structure may be expected to be relatively simple, regular and transparent, containing ‘attractive’ features that make it easier to learn and understand (Johanson 1992). A lingua franca is often readily accepted by speech communities with well-developed external ties to groups speaking other languages or varieties of the same language. It is, for instance, possible that the unusual preponderance of attractive features of the Turkic languages known to us is the result of a koinéization emerging in contact among highly mobile Turkic and non-Turkic groups over an extended period, and leading to continued reinforcement of more attractive structures. The lingua franca often plays the role of the dominant code in relation to the primary codes overlapped by it. It is then a B₁ in relation to each Aᵢ, even if it is a more widespread variety of the same historical language to which the Aᵢ belongs. It may influence the dominated primary codes in the same way as other dominant codes. And if a group of Aᵢ speakers shift to the lingua franca in question, using it as a B₁, there may of course be the same processes of imposition of Aᵢ copies as in other cases of substratum influence.
Similar relations may also exist between varieties that are considered to constitute higher and lower registers of one and the same language, e.g. the different levels represented by Modern Standard Arabic, Neo-Arabic dialects and intermediate prestige dialects such as the colloquial Arabic of Cairo. Complex interaction and mutual influence of lects may be observed within a multi-layered language of this kind (Medjell, this volume). The categories of the Code-Copying Framework are to be applied analogously to the description of the relations between such lects.

Effects of reorganization processes
Changes of special kinds may be due to reorganization processes that take place when groups using different codes are brought together in mixed social networks. Such abruptive situations may break the genetic connection, that is, cause varieties to break away from their ‘families’ and to enter new complex processes of mutual influence. The prehistory of several Altaic languages is likely to have involved such changes resulting from the coexistence of linguistically disparate groups in mixed tribal confederations.

Pidgins are one type of codes formed by different speech communities attempting to communicate in spite of mutually intelligible primary codes. In such cases, a basic code frame is provided by the dominant code – e.g. English in so-called ‘English-based’ pidgins –, and copies of elements of the dominated code are inserted into it. Seen in the perspective of the Code-Copying Framework, a pidgin is a Beta variety which exhibits a markedly reduced grammatical and lexical structure, and which is not the primary code of any of the groups using it. The creolization of a pidgin, by which it becomes the primary code of a speech community, implies a conventionalization process with decreasing variation.

The roles of social and structural factors
The interplay of structural and social factors and their influence on copying will not be discussed in detail here. Suffice it to say that social factors can overcome considerable structural obstacles to copying. Under appropriate social circumstances – particularly in intense and long-lasting contact situations – almost any feature from one code can ultimately be copied into another. The copiability of a particular structure into another code is determined in part by the prestige of the model code and in part by ‘attractive’ structural properties. Attractive properties may be copied even in the absence of strong social pressure, but the presence of such pressure can ultimately promote copying even of unattractive properties. The presence of a copy in a basic code may mean (i) that it has attractive properties, (ii) that the prestige of the model code has overcome its unattractiveness, or (iii) that both attractiveness and social influence have been at work. Examples considered in isolation often do not allow for distinguishing between these cases, but in many instances it seems possible to establish degrees of attractiveness independent of social factors (Johanson 1992; cf. Comrie 1995).

References
Brendemoen, Bernt 1987: Turkish in Europe – some aspects of semantic transfer. (Manuscript.)
Equivalence and levels of analysis

Hendrik Boeschoten, Katholieke Universiteit Brabant

In studies of language contact the notion of (inter-linguistic) equivalence plays an important role. Notions of equivalence have been put into perspective, among other things, in recent studies on code-switching (cf. below). In the present contribution, I will offer some critical notes on certain popular assumptions about the nature and explanatory power of inter-linguistic equivalence. In particular, I will address the problem of so-called “marked” structures.

I will first discuss partial word order equivalence in syntax. This matter has received much attention and has been put into perspective, for instance, in Thomason & Kaufman’s (1988) much quoted book. The argument goes as follows: Languages have basic word order patterns (such as SOV, SVO, etc.). However, at the same time languages might have “marked” alternative options; an SOV language may, for instance, allow for SVO patterns under certain conditions. If such a language is in contact with a language with an SVO pattern, these “marked” patterns are assumed to offer equivalence sites, not only facilitating code-mixing¹ (Myers-Scotton 1993), but also offering leverage for language change of the convergence type. As a mechanism for change, cases like these are akin to the types in which internal tendencies are reinforced.

¹) The terminological fog enveloping so-called “code-switching” research is slowly lifting nowadays. “Code-switching” is a very bad cover term for the linguistic phenomena associated with language mixing, among other things because, more often than not, it seems highly improbable that any switching between languages is going on either psychologically or linguistically (see, among others, Johnson 1993, Boteschoten 1998). Code-mixing is much better (as with Muyssen, forthc., who distinguishes insertion, alternation and what he calls “congruent lexicalisation” as subdivisions). Here, I will reserve Johnson’s (1993) term “code-copying” for grammatical phenomena.