On contact-induced grammaticalization*

Bernd Heine and Tania Kuteva
Universität zu Köln / Heinrich-Heine-Universität Düsseldorf

Grammaticalization is based on universal strategies of conceptual transfer. Contact-induced language change on the other hand is an areally confined process resulting from specific historical events. What this suggests is that the two constitute quite divergent phenomena and, in fact, in the relevant literature they tend to be described as mutually exclusive processes. Accordingly, this literature abounds with discussions on whether some specific grammatical change is due to the former or the latter. The position taken in this paper is that the two are in no way mutually exclusive; rather, perhaps more often than not, they jointly conspire in triggering grammatical change.

1. Introduction

Grammaticalization is a process leading from lexical to grammatical and from grammatical to more grammatical forms, and since the development of grammatical forms is shaped by constructions as well as larger context settings, the study of grammaticalization is also concerned with constructions and larger discourse units (Heine, Claudi & Hünnemeyer 1991; Hopper & Traugott 1993; Bybee, Perkins & Pagliuca 1994). A large body of knowledge has been assembled on the evolution of grammatical categories in the languages of the world (Heine and Kuteva 2002). All this work, however, has focused on language-internal grammatical change. Not only have grammaticalization processes usually been viewed as independent, language-internal changes, but it has also been claimed that the same grammatical category may re-emerge in a given language, and that this persistence or "diachronic stability" is a phenomenon entirely language-internally conditioned. Contact-induced language change on the other hand is an areally confined process resulting from specific historical events. What this suggests is that grammaticalization and language change induced by contact constitute quite divergent phenomena and, in fact, in the

Studies in Language 27:3 (2003), **529–572.**
ISSN 0378–4177 / E-ISSN 1569–9978

relevant literature the two tend to be described as mutually exclusive processes.[1] Whether some specific grammatical change is due to the former or the latter, has been the subject of some controversies.[2]

It has occasionally been argued though that language contact and grammaticalization can go together (Heine 1994; Kuteva 2000; Heine & Kuteva 2001). Dahl (2000b: 317) points out that grammaticalization processes tend to cluster not only genetically but also areally, and the terms areal grammaticalization (Kuteva 2000) and grammaticalization area (Kuteva 1998; Stolz & Stolz 2001: 1549) have been proposed to describe the effects of grammaticalization processes on the areal patterning of linguistic structures. Still, the question of whether indeed, or how, the two are interrelated has never been addressed in any detail.

In the present paper it is argued that universal processes of grammatical change tend to be influenced by historical processes leading to the areal diffusion of linguistic structures. The paper will deal with the unilateral or reciprocal influence of one language (or dialect) on another. Grossly speaking, influence manifests itself most clearly in the transfer of linguistic material from one language to another, where linguistic material can be of any of the following kinds:

(1) Kinds of linguistic transfer
   a. Forms, that is, sounds or combinations of sounds,
   b. Meanings (including grammatical meanings) or combinations of meanings,
   c. Form-meaning units or combinations of form-meaning units,
   d. Syntactic relations, that is, the order of meaningful elements,
   e. Any combination of (a) through (d).

In the present study we will be concerned with (1b), more specifically with the transfer of grammatical meaning. As we will see in section 2, the type of transfer examined is more complex,[3] though not necessarily less common, than other types of (1b), or of (1) in general. In works on language contact, the type of transfer that has received most attention is (1a). Useful proposals have been made with regard to the mechanisms by which cultural lexical borrowings (words for objects new to the culture, e.g. *espresso,* and also for new concepts, e.g. 'overtime') as well as core lexical borrowings materialize. It has been argued that core borrowed forms usually begin life in the recipient language when bilinguals introduce them as singly-occurring code-switching forms. Whereas code-switching is a plausible mechanism e.g. for the emergence of core lexical borrowings, when it comes to the transfer of grammatical meanings, code-switching does

not seem to be helpful as an explanatory factor. Neither is the notion of restructuring of a language´s frame (often referred to as convergence) that has been proposed as the mechanism bringing in grammatical outcomes from another language (cf. Myers-Scotton 2002), this notion simply being too general. Instead, we will propose below a concrete mechanism that promotes — we contend — new grammatical outcomes in language-contact situations.

Weinreich ([1953] 1964: 30–1) distinguishes three kinds of grammatical interference. One involves the transfer of morphemes from what he calls the source language to the recipient language, that is (1c). A second kind of interference concerns grammatical relations (1d), and the third involves functions, that is (1b), the latter two kinds involving what he refers to as model and replica languages. Since we will be concerned with (1b), we will adopt the terms proposed by Weinreich for both (1b) and (1d) by distinguishing between model languages (M), providing the model for transfer, and replica languages (R), making use of that model.

Transfer requires some kind of interlingual identification (Weinreich [1953] 1964: 7–8, 32), in our case some way of equating a grammatical concept Mx of language M with a grammatical concept Rx of language R. In situations of intensive language contact, speakers tend to develop some mechanism for equating "similar" concepts across languages, something that Keesing (1991) describes as "formulas of equivalence"; we will refer to them as equivalence relations. Without looking further into this issue, we will assume that such a mechanism is crucial for conceptual transfer.

Language contact of the kind looked at here presumably relates to situations typically involving large-scale bilingualism among the linguistic communities concerned, or at least in one of the linguistic communities, and it involves a larger time span, not seldom extending over three to five centuries. In many of the cases presented, however, there is no reliable information on either of these issues.

For a better understanding of the following discussion, a few general points need to be addressed. The effects of contact-induced change will be referred to as transfer of linguistic material from one language (or dialect) to another. It is usually fairly easy to establish that transfer has taken place when (1c) obtains, in particular when lexical borrowing is concerned. In the case of (1b), where transfer is confined to meaning, or in our case to grammatical meaning, it turns out to be much more difficult to understand and describe what exactly is transferred; the relevant literature is rife with controversies relating to this issue. None of the reconstructions presented in this paper is ours; rather, we adopted them from authors who can be assumed to have a fairly extensive knowledge of

the languages and the sociolinguistic situations concerned. We only selected examples that are fairly uncontroversial, that is, where no or no substantial alternative hypotheses have been voiced.

In evaluating the data collected in this way we used historical, sociolinguistic, and linguistic criteria. Plausible cases of transfer of the kind examined here can only be made if the relevant languages are known to have a shared history of contact, in particular if there is evidence to suggest that (at least some groups of) speakers of the replica language were sufficiently exposed to the model language to make transfer possible. Linguistic criteria relate to questions such as the following: (i) What evidence is there for transfer to have taken place? (ii) Could it have taken place without involving language contact? Question (i) requires that we have some knowledge of the structure of the language prior to contact. In a number of cases to be discussed, such knowledge can only be derived from comparative evidence. For example, if we discover that some Austronesian language spoken in Papua New Guinea exihibits structural properties not found in fellow Austronesian languages, but that exactly these properties are also found in a neighboring language of Papua New Guinea that is known to have been in contact with that Austronesian language, then a strong case can be made for transfer; as we will see below, a number of our examples are of this kind.

Question (ii) raises more serious problems. For example, some Slavic languages have developed use patterns that have been described as (definite or indefinite) articles. The fact that it is exactly those Slavic languages known to have had close contacts with Germanic or Romance languages that have done so would seem to suggest that this development was due to the influence of neighboring languages having full-fledged articles. However, the development of articles is a universal process that has taken place in a number of languages without there being any evidence that language contact played a role; accordingly, the possibility that these Slavic languages might have developed their articles without the influence from neighboring languages cannot be ruled out. In some cases there is additional evidence to argue that contact was a contributing factor in triggering or accelerating the process; in other cases such evidence is not provided by the author concerned and we decided not to discuss them in this paper.

2. Contact-induced grammaticalization

Contact-induced grammaticalization is a grammaticalization process that is due to the influence of one language on another.[4] There are two main types of contact-induced grammaticalization depending on whether or not there exists already a model for the process in the model language to be replicated. If no such model exists we will refer to the process as ordinary grammaticalization[5] (section 2.1); if there is a model that is transferred to the replica language we will refer to it as replica grammaticalization (section 2.2).

2.1 Ordinary grammaticalization

Contact-induced grammaticalization rests on a strategy used for transferring some grammatical concept from the model language (M) to the replica language (R). On the basis of the evidence available, this strategy involves a mechanism of the kind sketched in (2):

(2) Ordinary contact-induced grammaticalization
   a. Speakers of language **R** notice that in language **M** there is a grammatical category **Mx**.
   b. They develop an equivalent category **Rx**, using material available in their own language (**R**).
   c. To this end, they draw on universal strategies of grammaticalization, using construction **Ry** in order to develop **Rx**.
   d. They grammaticalize construction **Ry** to **Rx**.

(2) is a simplified account of the process concerned, whose exact cognitive structure is still largely unclear. (2a) concerns translational taxonomy, an issue that is complex enough to deserve separate treatment. For our purposes, the question is what exactly the nature of the categories concerned is. Similarly, (2b) also raises a number of questions, e.g. whether or to what extent it relates to a conscious or an unconscious act, or what the motivations may be for (2b) to happen. (2c) and (2d) fall within the scope of grammaticalization theory, and there are at least some ways in which they have been, or can be, dealt with (see especially Heine, Claudi, & Hünnemeyer 1991; Hopper & Traugott 1993; Bybee, Perkins, & Pagliuca 1994). Note further that the mechanism sketched in (2) relates to a gradual process that does not happen overnight and may involve several generations of speakers; (2d) in particular may extend over centuries.

(2a) and (2b) may convey the impression that contact-induced grammaticalization serves to fill a gap in the replica language by drawing on an appropriate category in the model language. This, however, is not necessarily the case. As we will see below, there are cases where replication took place even though there already existed an equivalent category in the replica language.

We may illustrate the mechanism sketched in (2) with the following example. A grammatical domain where speakers of pidgin and creole languages appear to have drawn quite often on "substrate" languages as models for new categories is that of personal pronouns. Where distinctions between inclusive and exclusive first person plural pronouns or of dual pronouns were commonly made in the languages of the people speaking pidgins or emerging creoles as second languages, such distinctions tended to be replicated in the second languages. Tayo is a French-based creole which evolved around 1860 in St Louis, New Caledonia. Drubéa and Cèmuhî (= M), the two main Melanesian languages spoken in St Louis at that time, have an obligatory semantic category of dual (= Mx). Presumably in an attempt to replicate this category in Tayo (= R), the speakers recruited the French numeral *deux* 'two' (= Ry) and grammaticalized it to a dual form *-de* (= Rx); table 1 lists the resulting personal pronouns (Corne 1995). Note, that crosslinguistically numerals for 'two' constitute the primary source for the grammaticalization of dual markers (Heine & Kuteva 2002: 303–4).

Table 1. Dual personal pronouns in Tayo (Corne 1995: 125, 128).

Tayo		**Metropolitan French**	
nu- de	'we (DU)'	nous deux	'we two'
u- de	'you (DU)'	vous deux	'ye two'
le- de	'they (DU)'	les deux	'the(y) two'

Even more dramatic cases of replication of person-number distinctions can be found in varieties of Melanesian Pidgin. Proto-Oceanic, the hypothetical ancestor of the present-day Oceanic languages, is assumed to have distinguished between four categories of number (singular, dual, trial, and plural) as well as between inclusive and exclusive first person non-singular pronouns. These distinctions have largely been retained in modern Oceanic languages (= Mx) and were apparently replicated by speakers of these pidgin varieties (= Rx) by drawing on common grammaticalization processes (see Heine & Kuteva 2002):

The numeral *tu* 'two' was grammaticalized to a dual marker (cf. above for Tayo), the numeral *tri* 'three' to a trial marker (apparently not widely distinguished), the marker *-fala* (< English *fellow*) to a non-singular marker, and the combination *yu* 'you' plus *mi* 'I' to a first person inclusive marker (= Ry > Rx)[6] (Keesing 1988: 160–1).

But rather than to specific grammatical forms, contact-induced grammaticalization may also lead to the rise of new morphological classes. The following two examples illustrate the development concerned. Traditionally, the Aztecan language Pipil (= R), like most other Mesoamerican languages, has neither prepositions nor postpositions, but it has relational nouns instead (Campbell 1987; Harris & Campbell 1995: 126–7). Under the influence of the model language Spanish (= M), Pipil speakers have more recently drawn on these relational nouns (Ry) to develop a set of Spanish-type prepositions (= Rx): Consequently, they used a crosslinguistically common process whereby relational nouns are grammaticalized to adpositions (Heine, Claudi & Hünnemeyer 1991). As table 2 suggests, the process shows all the effects of the three main mechanisms involved in grammaticalization (see Heine & Kuteva 2002: 2): *desemanticization*, e.g. the loss of nominal meaning ('possession') in favor of grammatical meaning ('of'), *decategorialization*, leading to the loss of categorial properties of nouns, such as taking modifiers and affixes, and *erosion*, i.e. loss of phonetic substance (*-(i)hpak* > *pak* 'on').

Table 2. The grammaticalization of relational nouns to prepositions in Pipil (based on Harris & Campbell 1995: 126–7).

Relational noun		Preposition	
-(i)hpak	'on, upon, over, on top of'	pak	'on'
-pal	'possession'	pal	'of'
-wan	'with'	wan	'with'

A case strikingly resembling the Pipil one has been reported from Papua New Guinea. Ross (2001) describes a situation where two genetically unrelated languages spoken on Karkar Island off the north coast of Papua New Guinea have become semantically and syntactically largely intertranslatable while each of the two has retained its own lexical material — a situation he proposes to call metatypy.[7] The model language (= M) is Waskia, a Papuan language, and the

replica language Takia (= R), a Western Oceanic language; in the process of contact,[8] "Takia speakers have increasingly come to construe the world around them in the same way as the Waskia" (Ross 2001: 144). Among the properties discussed by Ross there is a set of postpositions exhibiting a similar semantic patterning in the two languages. Western Oceanic languages commonly have prepositions but Takia speakers have lost the prepositions. In an attempt to replicate the postpositions of Waskia (= Mx), Takia speakers developed postpositions (= Rx) by grammaticalizing inalienably possessed relational nouns (Ry).[9] In this way, a Proto-Western Oceanic construction illustrated in (3) turned into a postpositional construction (4) in Takia (note that Takia has given up the possessee-possessor order of Western Oceanic and adopted the possessor-possessee order of Waskia).[10] In accordance with mechanisms associated with grammaticalization, the prepositional phrase **i lalo-ña* turned into an adposition (decategorialization) and was phonetically reduced to *lo* (erosion). Thus, in order to establish equivalence with the model language Waskia, speakers of the replica language Takia appear to have had recourse to a grammaticalization process; using relational nouns for 'inside' and developing them into locative (inessive) adpositions, for example, is a crosslinguistically widespread strategy (Heine & Kuteva 2002: 182–3).

(3) Proto-Western Oceanic (Ross 2001: 143)
**i lalo- ña a Rumaq*
PREP inside- its ART house
'inside the house'

(4) Takia (Western Oceanic; Ross 2001: 143)
ab lo
house in
'in the house'

Takia and Pipil underwent the same process from relational nouns to adpositions in the creation of a new word class; what distinguishes the two cases essentially is that this new class appears to have replaced an earlier class of prepositions in Takia, while in Pipil there was no earlier equivalent class, hence no replacement.

In (2c) we observed that speakers of the replica language draw on universal strategies of grammaticalization in order to develop a new category that is equivalent to the one they find in the model language. The following example may illustrate the procedure involved. If a given language acquires a new definite article then essentially the only way in which this happens is by

grammaticalizing a demonstrative attribute to a definite marker (see Greenberg 1978). The Slavic language Sorbian is spoken in Germany and has been affected by nearly a millennium of contact with German. One result of this contact is that Sorbian (= R) has acquired a definite article (= Rx) on the model of German (= Mx). Sorbian has done so by extending the use of the Sorbian deixis-neutral demonstrative *tón* (M)/ *ta* (F)/*te,to* (N)[11] (= Ry) to introduce an optional definite article (Boretzky 1986: 17; Lötzsch 1996: 52–3). Note that the German definite article itself arose in the same way, and in some of its uses, it has retained properties of a demonstrative. However, it is unlikely that Sorbian speakers were aware of this fact when they developed their article; rather, we assume that what Sorbians did was to replicate a category by drawing on a universal strategy[12] — the very strategy that speakers of German and many other languages had used earlier.

Blends

That it can be language-specific properties in addition to, or even instead of, universal pathways that influence the particular way grammaticalization takes can be illustrated with the following example. In the Malaita Oceanic languages (= M) there is a post-verbal particle to mark the perfect aspect (= Mx), illustrated in (5). Apparently in an attempt to replicate this category, speakers of the English-based Solomons Pijin (= R) drew on the adverb *nao* (< English *now*) (= Ry) and grammaticalized it to a perfect particle (= Rx) in the same post-verbal slot of the verb phrase as in the Malaita model languages, cf. (6), and subsequently the existing pidgin marker *bin*, used for completed actions, virtually disappeared from Solomons Pijin.

(5) Kwaio (Eastern Oceanic; Keesing 1991: 330)
(ngai) e leka no'o.
him he go PERF
'He has gone.'

(6) Solomons Pijin (English-based pidgin; Keesing 1991: 330)
hem- i go nao.
(him- he go PERF)[13]
'He has gone.'

As far as we are aware, a grammaticalization from a temporal adverb to a perfect marker is not commonly found.[14] What appears to have influenced this choice of grammaticalization is, at least to some extent, the fact that *nao* resembles phonologically the corresponding Oceanic perfect particles (Kwaio, 'Are'are *no'o*, Lau *na,*

Kwara'ae *na'a*; Keesing 1991: 329–30) — a case described by Weinreich ([1953] 1964: 39–40) as interlingual equivalence based on formal similarity.

As a rule, however, contact-induced grammaticalization influenced by formal similarity is in accordance with commonly observed patterns of grammatical evolution. Swedish speakers in the United States (R) have grammaticalized their transitive verb *bekomma* 'obtain, do somebody good or bad, make an impression on somebody' (= Ry) to an inchoative auxiliary 'become' (= Rx), replicating the English auxiliary *become* (= Mx) presumably on account of the phonological similarity it shares with the model (Ureland 1984: 308). In doing so, they have drawn on a universal strategy whereby lexical verbs tend to be recruited for the expression of tense-aspect meanings and are grammaticalized into verbal auxiliaries.

In a similar fashion, speakers of Pennsylvania German (= R) may use the locative preposition *bei* 'at' to present agents in passive constructions (Costello 1985: 117). The canonical marker for agents in passive constructions in German is the ablative preposition *von* 'from, of', but crosslinguistically a much more common strategy is to grammaticalize a locative marker to an agent marker (Heine & Kuteva 2002: 199–200). That Pennsylvanian German speakers opted for this strategy (= Ry > Rx) was most likely influenced by the fact that their second language English uses the phonologically (and etymologically) corresponding preposition *by* (= Mx) in its equivalent passive construction.

These examples suggest the following: Among all the various possible ways of adapting their language to the model language, speakers of the replica language appear to select what seemingly is a fairly complex solution. Rather than simply borrowing grammatical forms from the model language, they draw on structures that correspond neither in their morphonsyntactic form nor in their meaning to the model, and in order to achieve equivalence they have to go through a process[15] whereby lexical structures are developed into grammatical markers, adpositional phrases are gradually transformed into adpositions, etc.

In the examples presented in this section, speakers of the replica languages had a model category (= Mx), but the model language (= M) apparently did not provide any guidance as to how to replicate Mx. The situation is different in another type of contact-induced grammaticalization that we will now turn to, where the model language not only provides a model category (Mx) but also a way of how to replicate that category (My).

2.2 Replica grammaticalization

Ordinary contact-induced grammaticalization, as described in section 2.1, is probably not the most common way in which language contact leads to the development of a new grammatical category in the replica language. There is a second type that appears to be more common; what is achieved by it is that, rather than a grammatical concept, it is a grammaticalization process that is transferred from the model (M) to the replica language (R);[16] we will therefore refer to it as replica grammaticalization. The mechanism underlying the process concerned is sketched in (7); it differs from (2) only in a slightly different framing of (c).

(7) Replica grammaticalization
   a. Speakers of language **R** notice that in language **M** there is a grammatical category **Mx**.
   b. They develop an equivalent category **Rx**, using material available in their own language (**R**).
   c. To this end, they replicate a grammaticalization process they assume to have taken place in language **M**, using an analogical formula of the kind [**My** > **Mx**] = [**Ry** > **Rx**].
   d. They grammaticalize category **Ry** to **Rx**.

That the mechanism sketched in (7) is different than the one discussed in section 2.1 can be illustrated with the following example volunteeed by Weinreich:

> In Silesia, the identification of the third person plural of the local Polish dialect ("Wasserpolnisch") with the equivalent category in German has led to the unexpected use, completely unknown in Polish, of the third person plural for polite address […]. (Weinreich [1953] 1953:40).

As we will see below (especially section 2.3), the use portrayed by Weinreich is in no way "unexpected"; rather, this example appears to be a canonical instance of (7): the model language German (= M) has extended the use of the third person plural pronoun *sie* 'they' (= My) to serve also as a second person singular pronoun (*Sie* 'you') used for polite/formal reference (= Mx). Polish speakers (= R) in Silesia replicated this process by extending the use of their third person plural pronoun (= Ry) to a new function. Most likely, those Polish speakers were unfamiliar with the historical factors that were responsible for that grammaticalization in German; still, from the sociolinguistic, pragmatic, and grammatical information that was accessible to them they had enough information for replication (but see section 2.4.1).

Cases of replica grammaticalization are fairly easy to identify when the model language has developed a grammatical category by using a conceptual source that is rarely encountered crosslinguistically and where exactly the same source is used by speakers of the replica language. The Irish "hot-news" perfect[17] (= Mx) is a case in point: This category is based on what appears to be an instance of a Location Schema [X is after Y], where the aspectual notion of a hot-news perfect is encoded by means of a locative (or temporal) preposition 'after' (= My), cf. (8). Presumably around the late 17th century (Sullivan 1980: 205), the same grammaticalization appears to have taken place in Irish English (= Ry > Rx; Filppula 1986; Boretzky 1986: 25; Harris 1991: 01ff.), cf. (9). We are not aware of any other language in the world that has undergone a similar process — a fact that may be taken as compelling evidence to consider this to be an instance of replica grammaticalization.

(8) Irish (Harris 1991: 205)
Tá sí tréis an bád a dhíol.
be: NON-PAST she after the boat selling
'She has just sold the boat.'

(9) Irish English (Harris 1991: 205)
She's after selling the boat.[18]
'She has just sold the boat.'

In some cases again it turns out hard or even impossible to distinguish replica grammaticalization from internal grammaticalization. In English, the verb *do* (= My) has a grammaticalized function as a pro-verb (= Mx), that is, as a verbal anaphor referring to a previously mentioned predicate (e.g., *She likes roses, and I* do *too.*). As a result of close contact with English, speakers of Pennsylvania German appear to have replicated this use pattern, using their action verb *duh* 'do' (= Ry > Rx) for this purpose (Huffines 1986: 146–50). Now, verbs meaning 'do' are crosslinguistically probably the main source for pro-verbs (cf. Heine & Kuteva 2002: 199–200). While it is likely that the generalization of this use pattern in Pennsylvania German was triggered by language contact, it remains unclear whether the choice of *duh* 'do' as a pro-verb was due to English influence or to a universal principle, or else to a joint effect of both.

What distinguishes replica grammaticalization from ordinary grammaticalization (section 2.1) is the fact that the model language provides speakers of the replica language with guidelines as to how to replicate a grammatical category via grammaticalization. We will now illustrate the nature of this process with a few examples.

From interrogative to relative clause marker

The above example concerns a case where a new grammatical category was developed for which there existed no equivalent in the replica language. We will now see that this is by no means a requirement for replication to happen since the replica language already had an equivalent category. Our first example concerns Basque. The use of interrogative pronouns (e.g. English *who*?) as relative clause markers (English *who*) is crosslinguistically widespread, still, most instances of it are to be found within Europe (Heine & Kuteva 2002): It has happened in all Romance, all Slavic, in some Germanic languages, as well as in Modern Greek, Hungarian and Georgian (though not in Finnish); it is therefore discussed by Haspelmath (1998: 279–80; 2001) as one of the eleven properties characterizing Standard Average European (SAE). Basque is a notable exception; in this language, a finite relative clause precedes its head; there is no relative pronoun, but the verb takes a suffix marking it as subordinate. However, for some generations now, a new relativization strategy has been used in certain parts of the Basque-speaking area (= R) under the influence of Spanish (= M). What happened is that the question marker *zein* 'which?' (also 'who?' in places) was used as an equivalent (= Ry) of the Spanish interrogative pronouns (= My) and *zein* was grammaticalized to a relative clause marker (= Rx) (Trask 1998: 320; see also section 2.3.1). Relative clauses introduced by *zein* are documented in texts since the 17th century but as a rule only in translations (Hurch 1989: 21). As we noted above, the new form does not replace the earlier structure; rather, *zein* is added to the existing relative construction[19] — with the result that there is now double marking on post-nominal relative clauses.

Basque is not the only language spoken at the periphery of the Standard Average European area that has replicated the European strategy of using interrogatives as relative clause markers; another example is provided by Romani: Under the influence of Balkanic languages, the Kelderaš/Lovari variety of the Vlach dialect of Romani has grammaticalized its interrogative markers *k-* and *s-*, both inherited from its Indic past, to relative clause markers, thereby replacing the earlier Indic constructions of marking relative clauses (Matras 1996: 65–6).

With the spread of European languages to other parts of the world, this interrogative-to-relative strategy has also been extended to other continents, such as the Americas. The North Arawak language Tariana of northwestern Brazil is in close contact with Portuguese, the official language of Brazil, and has been influenced by the latter in various ways (Aikhenvald 2001). One way in which this influence is reflected concerns a process that is strikingly similar to

what happened in the Basque case just examined: Young and innovative speakers of Tariana (= R) recognize that in Portuguese (= M), interrogative pronouns (= My) are also used as relative clause markers (= Mx), and these speakers also use their interrogative pronouns (= Ry) as markers of relative clauses (Rx) on the pattern of the model language Portuguese; in doing so, they retain their own relative construction and simply add their interrogative pronoun (e.g. *kwana* 'who?') — exactly the way Basque speakers do. And The Standard Average European model of extending the use of interrogative pronouns to relative clauses also appears to have been applied among Yiddish bilinguals in the USA, even though this model had already been used earlier in Yiddish: These bilinguals identify English *who*, the interrogative pronoun for human referents (= My), with their corresponding interrogative pronoun *ver* 'who?' (= Ry) and introduced it as a relative clause marker *ver* (= Rx) on the model of the English relative clause marker *who* (= Mx), e.g. *der mentš ver is do* 'the man who is here' (Weinreich [1953] 1964: 30).

Extending the functions of case markers

A third possibility is provided by situations where there is a grammatical category in the replica language that is equivalent to the one in the model language, but where the former is replaced by the new one, replicated from the model language. Basque is a language which is rich in case distinctions marked by nominal suffixes. The Basque dialects spoken in France have had centuries of close contacts with Gascon and later on with French (see section 2.3). They have virtually lost one of their cases, the instrumental — *ez* (see Haase 1992: 67–71). While this case is confined primarily to marking instrumental participants, the model languages Gascon and French (= M), like most other European languages (see Stolz 1996a, 1996b), have a comitative-instrumental polysemy, that is, Gascon *dab*, *dambe* and French *avec* are used for both comitative and instrumental participants (My = Mx); note that while case is expressed by suffixes in Basque (= R), the model languages use prepositions instead. Now, wherever the comitative-instrumental preposition is used to present instrumental participants (= Mx) in the model languages, the replica language Basque uses its comitative case suffix — *ekin* (or — *ekilan*) (= Ry) for instrumentals (= Rx), cf. (10).

(10) Basque (Haase 1992: 67)
koller bat- ekilan
spoon one- COM
'with a spoon'

In other words, the instrumental case is gradually replaced in these Basque varieties by the comitative suffix,[20] that is, in an attempt to establish an equivalence relation with the model languages, the comitative-instrumental polysemy of the model languages is replicated in Basque by grammaticalizing the comitative to present instrument participants too.[21] However, as is to be expected in grammatical developments of this kind, the replaced category does not disappear entirely; rather, it tends to survive (at least for some time) in frozen expressions. Accordingly, the instrumental suffix — *ez* is still used in fixed/lexicalized contexts (e.g. *oin-ez* [foot-INSTR] 'on foot'), idioms, proverbs, or in more strongly grammaticalized uses as a marker of clause combining (Haase 1992: 67–8).

This Basque example is not an isolated instance in Europe. There are other languages spoken at the periphery of what Haspelmath (1998; 2001) calls Standard Average European that have undergone a similar process of replica grammaticalization. The Slavic language Sorbian is one of them: It has been in contact with German for nearly a millennium (Lötzsch 1996), and as a result of this contact, Sorbian has replicated the German comitative-instrumental polysemy.

Slavic languages such as Russian and Polish consistently distinguish between marking comitatives and instrumentals. Thus, in example (11a) from Polish, the comitative participant has double marking, an instrumental case suffix plus the preposition *z*, whereas the instrument in (11b) is marked simply by means of the instrumental case suffix. Both case roles use the instrumental case suffix, but the comitative is distinguished by adding a preposition. Sorbian is claimed to have replicated the German comitative-instrumental polysemy by grammaticalizing the comitative preposition *z(e)* 'with' to an instrumental marker — with the effect that this preposition is now used obligatorily for both comitative (12a) and instrumental participants (12b).

(11) Polish (Lötzsch 1996: 56)
  a. *Ja mówię z przyjacielem.*
     I speak with friend:INST[22]
     'I speak with my friend.'
  b. *Ja pracuję ręką.*
     (I work hand:INST)
     'I work with my hand.'

(12) Upper Sorbian (Lötzsch 1996: 56)
  a. *Ja rěču z prěćelom.*
     I speak with friend:INST)
     'I speak with my friend.'

b. *Ja dźělam z ruku.*
I work with hand:INST)
'I work with my hand.'

Sorbian is not the only Slavic language that, as a result of its contact with a non-Slavic language, has replicated the comitative-instrumental polysemy of its SAE model language. Another example is provided by Molisean, a variety of the Slavic Croatian minority in Molise, southern Italy, which has a history of 500 years of language contact with Italian (Breu 1996: 26–7). In the same way as Sorbian, this Slavic minority has also extended the use of its preposition *s* to mark instrumental participants. Accordingly, a Standard Croatian instrumental phrase as in (13a) is expressed in Molisean Croatian as in (13b) on the model of Italian, cf. (13c).

(13) Instrumental marking in Standard Croatian, Molisean Croatian, and Italian (Breu 1996: 26).
a. Standard Croatian
nožem
knife.INST
'with a knife'
b. Molisean Croatian
s nožem
with knife.INST
'with a knife'
c. Italian
con un coltello
with a knife
'with a knife'

In this way, Molisean speakers, like Sorbian speakers, established a kind of equivalence relation in a double sense: First, they introduced a comitative-instrumental polysemy identical to the one of the model language and, second, they introduced a preposition where the model language also has a preposition.

Other cases where replica grammaticalization had the effect of introducing case-marking polysemies are not hard to come by. The following example concerns language contact between Turkish and the Kartvelian language Laz in Eastern Anatolia. Ardeşen Laz, belonging to the Pazar dialect group of Laz, has drastically restructured its nominal case system as a result of heavy Turkish influence (Haig 2001: 214–6). Still, this influence had its limits. As table 3 shows, there remain differences between Ardeşen Laz and Turkish in the

marking of core arguments: The indirect object in particular takes a suffix — *(y)A* in Turkish but is zero-marked in Ardeşen Laz. Now, Haig found that the young urban Turkish/Laz bilinguals he consulted (now living in Ankara) (= R) consistently used the goal marker *-şa*[23] for indirect objects when speaking Laz, suggesting that these Laz speakers have gone one step further in matching the Turkish system of nominal core argument marking. In doing so, they drew on a common grammaticalization process whereby the use of their allative/goal marker (= Ry) was extended to mark indirect objects too (= Rx) (see Heine & Kuteva 2002) and, most likely, this process was triggered by the situation in Turkish (= M), where the goal marker -(y)A (= My) is also used for indirect objects (= Mx) (see table 3). While the process can be described as being morphological in nature, its effect is that the core case system of these Ardeşen Laz speakers is now structurally almost identical with that of the model language Turkish.

Table 3. Case marking of core arguments in Ardeşen Laz and Turkish (based on Haig 2001: 215–6).

Argument type	S	A	0	Indirect object	Goal
Ardeşen Laz	-Ø	-Ø	-Ø	-Ø	-š
Turkish	-Ø	-Ø	-Ø/-(y)I	-(y)A	-(y)A
Ardeşen Laz (young urban bilinguals when speaking Laz)	-Ø	-Ø	-Ø	-ša	-ša

Word order change

Grammaticalization is essentially a semantic process but, as we saw above, it can also have remarkable morphosyntactic consequences: Once there is a pattern of grammaticalizing items having a nominal meaning to e.g. locative markers, this may lead to a development from relational nouns to a morphosyntactic category of adpostions, as the Pipil and Takia examples in section 2.1 suggest.

One common effect language contact has is to induce people to transfer arrangements of meaningful elements from one language to another. In many cases, however, a new word order arises without actually involving word order change. This is the case, for example, when in the process of grammaticalization some structure is reinterpreted as some other structure, with the result that a seemingly new word order arises. The following is an example of such a word

order change without word order change, again relating to the Takia-Waskia situation of metatypy on Karkar Island, Papua New Guinea (Ross 2001; see section 2.1). In an attempt to assimilate their language to Waskia (= M), speakers of the Western Oceanic language Takia (= R) largely adopted the syntax of the Papuan language Waskia. For example, whereas Proto-Western Oceanic is reconstructed as having had SVO (subject-verb-object) as its unmarked clause order, Takia has adopted the SOV-order of Waskia, and while in Proto-Western Oceanic the determiner (article) preceded the head noun, cf. (14a), it follows the head noun in both Takia (14b) and Waskia (14c).

(14) Determiner — noun order (Ross 2001: 142)
  a. Proto-Western Oceanic
     *a tam^w ata
     DET man
     'the man'
  b. Takia
     tamol an
     man DET
     'the man'
  c. Waskia
     kadi mu
     man DET
     'the man'

The way Takia experienced a change from preposed to postposed determiner is the following: Proto-Western Oceanic had a set of three deictic morphemes, one of them being *a ('that, near speaker'). When one of these was used attributively, it followed the Proto-Western Oceanic adjective syntax, taking a pronominal suffix agreeing in person and number with the head noun, the result being a structure as in (15a). This structure underwent a canonical grammaticalization process (that we encountered already in our Sorbian example; section 2.1) from demonstrative to definite determiner, resulting in the Takia structure (15b) (= (14b)). In this process, the construction was subject to the usual mechanisms of grammaticalization, that is, desemanticization: loss of deictic force, decategorialization: loss of its status as an inflectable constituent, and erosion: loss of the preposed article *a and reduction of the postposed determiner *a-ña > an.

(15) Grammaticalization (Ross 2001: 142)

a. Proto-Western Oceanic
*a tam^w ata a- ña
DET man that- 3.SG
'the man'

b. Takia
tamol an
man DET
'the man'

To conclude, as the description by Ross suggests, the change in Takia from preposed to postposed determiner did not involve any change in the order of constituents;[24] rather, in order to adapt to the postposed determiner order (= Mx) of the model language Waskia (= M), Takia (= R) speakers drew on a construction that was available to them, namely a postposed deictic determiner (Ry), and they grammaticalized it to Rx.

This example illustrates a process that appears to be fairly common in language contact: In order to develop a structure that is equivalent to the one in the model language, speakers draw on some existing use pattern in the replica language that corresponds most closely to the model, frequently one that until then was more peripheral and of low frequency, and they activate it — with the effect that that peripheral pattern turns into the regular equivalent of the model, acquires a high frequency of use, and eventually may emerge as a fully grammaticalized category, equivalent to the model category.

Kupwar

Contact-induced transfer is a wide field, and we are confined here to but one of its manifestations, that is, to a sub-type of the transfers distinguished in (1). Nevertheless, replica grammaticalization appears to play a role in virtually all situations of intensive language contact we are familiar with, in particular in situations of metatypy (Ross 2001; see 2.1); the examples of Takia-Waskia metatypy in Papua New Guinea that we presented above are a case in point. Another classic example of metatypy is that of the Indian village of Kupwar.

Kupwar offers a rich laboratory for contact-induced assimilation of languages. Spoken in State in the Sangli District of Maharashtra India, Kupwar has been described as an extreme case of mutual assimilation of languages, leading to a high degree of intertranslatability (Gumperz & Wilson 1971). The languages involved belong to two genetic stocks, viz. Indo-European (Urdu, Marathi) and Dravidian

(Kannada, Telugu); Marathi and Kannada have been spoken in Kupwar for more than six centuries, and Urdu for at least three centuries.

One example is provided by the development of markers for polar ("yes-no") questions. As far as the description by Gumperz and Wilson (1971: 160) suggests, the Indo-European languages Standard Hindi-Urdu and Marathi provided the model (= M): They appear to have grammaticalized the interrogative for 'what?' (*dii* in Standard Hindi-Urdu) (= My) to a polar question particle (= Mx), while Kannada has a verbal suffix — *a* instead. Now, all languages spoken in Kupwar (= R) have used this grammaticalization as a model, replicating the Indo-European pattern in general and the Marathi one in particular: They use the interrogative for 'what?' (= Ry) as a clause-final marker for polar questions (Kupwar Urdu *kya*, Kupwar Marathi *kay*, Kupwar Kannada *yan*) (= Rx), and Kannada speakers in Kupwar reject the suffix *-a*. What this suggests is that the Dravidian language Kannada as used in Kupwar has undergone the same grammaticalization process from an interrogative pronoun for 'what?' to a polar question marker as the two Indo-European languages had done earlier.

A second example concerns case marking in Kupwar, this time involving Urdu as a replica language. Crosslinguistically, there is a grammaticalization process whereby benefactive (or dative) case markers give rise to purpose markers when instead of a human participant there is an inanimate participant (cf. English *for*; see Heine & Kuteva 2002). Standard Urdu uses the marker *ke liye* in purpose constructions, but Urdu speakers of Kupwar (= R) follow the pattern used in the other two languages spoken in Kupwar, namely that of the Indo-European language Marathi and the Dravidian language Kannada (= M), according to which purpose clauses are encoded by means of a construction [verb + oblique + dative]. What this situation suggests is that Kupwar Urdu has extended the use of its dative marker (= Ry) to express purpose too (= Rx). In doing so, it appears to have replicated a grammaticalization process that Marathi and Kannada had undergone earlier.

Kupwar provides yet another example of case syncretism. There is a not uncommon grammaticalization process according to which the use of allative/dative case markers is extended to mark patient participants too, i.e. direct objects; at the initial stage, this process tends to be confined to human and definite object referents (cf. Spanish *a*, Maltese *(li)l*; see Heine & Kuteva 2002: 38, 103:, Borg & Mifsud 2002). Such a process also appears to have occurred in Kupwar (Gumperz & Wilson 1971). Standard Kannada, a Dravidian language, distinguishes accusative and dative participants by means

of postpositions. But in Kupwar, Kannada speakers have been in contact with speakers of the Indo-European languages Urdu and Marathi (= M), which both have only a dative postposition for human objects (My = Mx). As a result of these contacts, lasting for more than three centuries, Kannada speakers (R) appear to have replicated the situation found in the two Indo-European languages by grammaticalizing the dative postposition (= Ry) to mark human direct objects too (= Rx) (Gumperz & Wilson 1971:158).

2.3 Case studies

In the preceding sections we have been dealing with a range of different languages and a diversity of grammatical categories. In the present section we will deal with two more specific issues. On the one hand, we will demonstrate how contact-induced grammaticalization may affect the structure of some particular language, the language chosen being Basque (2.3.1). On the other hand, we will look into the question of how one particular grammatical meaning may be affected by language contact; to this end, we will be concerned with a crosslinguistic survey of future tense categories (2.3.2).

2.3.1 *Basque*

In his recent work on Standard Average European (SAE), Haspelmath (1998; 2001) concludes that on the basis of the relative number of features they share with SAE, three kinds of language groupings can be distinguished: The nucleus, comprising Dutch, German, French, and northern Italian dialects, the core, which in addition includes the other Romance and Germanic languages plus West and South Slavic, and the Balkan languages, while the periphery consists of East Slavic, Baltic, Balto-Finnic, and Hungarian, "plus perhaps Basque, Maltese, Armenian and Georgian" (Haspelmath 1998:273). Basque is classified somehow as a marginal European language, showing very few Europeanisms (Haspelmath 2001:1493). In the present section we will try to show that Basque speakers in France nevertheless made considerable efforts to move towards SAE,[25] and that in doing so they used contact-induced grammaticalization as one of their main strategies. Discussion is confined essentially to Basque varieties spoken in south-western France, for which there exists a detailed analysis (Haase 1992).

For centuries, these Basque varieties have been influenced by two Romance languages, Gascon and later on French — to the extent that there is now a wide range of semantic and morphosyntactic equivalence relations, resulting in a certain degree of translational equivalence; Ross (2001:146) therefore views this as a case

of metatypy (see section 2.1). We will now provide a few examples to illustrate how contact-induced grammaticalization contributed to the rise of such equivalence relations; the reader is referred to Haase (1992) for additional examples.

A paradigm case concerns the rise of an indefinite article in Basque via replication of the Romance path of grammaticalization from numeral 'one' to indefinite article; we will return to this case in section 2.4. Another case was discussed in section 2.2, where we gave an example of how Basque evolved a comitative-instrumental case polysemy by extending its comitative case marker to also present instrumental participants on the model of the two Romance languages, thereby contributing to the gradual disappearance of their own instrumental case suffix (see also below). Note that the Romance languages use primarily prepositions for case marking while Basque has 14-plus suffixes for case marking. Contact with the Romance languages left the morphological shape unaffected but had some effects on the meaning of case suffixes in Basque, resulting from the new use patterns of these suffixes.

But semantic replication also has a morphosyntactic effect increasingly affecting the Basque case system: On the model of the Romance prepositions, Basque is gradually developing a set of postpositions via the grammaticalization of relational nouns (a process that, as we saw in section 2.1, led to the emergence of new paradigms of adpositions in Pipil and Takia). Thus, by grammaticalizing relational nouns such as *buru* 'head' or *baita* 'interior', Basque is acquiring a set of complex postpositions modelled after corresponding complex prepositions in the Romance languages; for example, the Basque postposition *-ri/-ra(t) buru-z* (DAT/DIR head-INSTR) 'in the direction of' is modelled after the Gascon preposition *(de) cap a* ((from) head at) 'in the direction of' (Haase 1992: 73–5);[26] for a paradigm of equivalences in complex adpositions, see Haase (1992: 80).

Within the verbal system, the same kinds of replication can be observed, as the following examples suggest. Gascon has grammaticalized a progressive aspect based on the Location Schema [X is at doing Y]; Basque has done the same on the pattern of Gascon, using a nominalizer (*-tze*) instead of the Gascon infinitive marker[27] (subsequently grammaticalizing the progressive further into an imperfective; Haase 1992: 93).

Furthermore, Basque is said to have replicated the Latin-Romance 'have'-perfect via contact-induced grammaticalization, using the verb *ukan* 'have' plus the transitive participle as a complement to form a recent perfect. The category is restricted to transitive main verbs, while intransitive verbs use the same periphrasis with the verb *izan* 'be' as an auxiliary (Haase 1992: 92–3).

But developments within the Basque verbal system have also had an impact on participant marking. Basque has traditionally no passive category, but in an attempt to replicate the Gascon/French passive it is grammaticalizing its resultative aspect for this purpose, adding the intransitive participle form *izan* (the agent being suppressed):

(16) Basque (Example taken from the weekly *Herria*; Haase 1992: 102)
Adolfo Villoslada … libra- tu- a izan da.
Adolfo Villoslada … free- PTCP- ART be.PTCP PRES.3.SG
'Adolfo Villoslada (a kidnapped industrialist) has been freed.'

And Basque has also taken the next stage in replicating a Romance-type passive: Following the Gascon model, which uses either of the prepositions *de* or *per* to present an agent, it has chosen the ablative-partitive case suffix *-(r)ik* to introduce an agent:

(17) Basque (Haase 1992: 132)
A(b)antxü xakür bat- eta- (r)ik ausiki iza tzün.
almost dog one- TRN- ABL/PART bite.PTCP be.PTCP PRT.3.SG.2.AL
'He was almost bitten by a dog.'

As we saw in section 2.2, language contact in Basque as spoken in Spain has also affected the marking of relative clauses, in that the Spanish strategy of using interrogative pronouns as relative clause markers has spread to certain parts of the Basque-speaking area, where the question marker *zein* 'which?' has assumed the function of a relative clause marker (Trask 1998: 320). A similar situation is found in "French" Basque, where *zoin* 'which?' has been grammaticalized to a relative clause marker in the written language, a use that goes back at least to the 18th century (Haase 1992: 152; see section 2.4 below).

To conclude, contact-induced grammaticalization was one of the main factors responsible for changes leading to an increase in intertranslatability between the replica language Basque and the two Romance model languages. Via these model languages, Basque has acquired a number of features characterizing SAE[28] even if, on the surface, this fact is not yet reflected in its overall typological profile.

2.3.2 *Future tenses*

While grammaticalization in language contact situations is determined by universal principles, we also noted above that the particular conditions obtaining in such situations also play a role. What kind of conditions these are can be illustrated by using future tense marking as an example.

Universally, the primary source of future tenses is provided by motion schemas involving goal-directed verbs for 'come to' [X comes to Y] or 'go to' [X goes to Y], or a volition schema [X wants Y] using a verb for 'want' (Bybee et al. 1991; 1994). These schemas are referred to by Dahl (2000b: 319ff.), respectively, as the de-venitive, de-andative, and de-volitive constructions. We will adopt Dahl's terms as shorthand labels for these conceptual schemas — with one exception: we propose to replace the term "de-andative" by "de-allative" since the former term is employed in a number of linguistic works for the notion 'motion away from the deictic center', while the goal-directed motion schema [X goes to Y] involved in the rise of future tenses centrally involves 'motion to(wards) some goal', that is, an allative function. Other sources are cross-linguistically less common; this applies e.g. to temporal adverbs ('then', 'afterwards', etc.), obligation markers ('have to'), or inchoative markers ('become'), which have given rise to future tense markers but clearly less frequently so than the motion or volition schemas.

With the possible exception of temporal adverbs, all these sources have been recruited in European languages and are each suggestive of areal spread across neighboring European languages (Dahl 2000b). However, in spite of the fact that the linguistic history of Europe is better known than that of any other part of the world, the discussion by Dahl suggests that there is essentially no reliable information on which of the languages provided the models, or on whether indeed language contact, as against genetic retention or coincidence, was the decisive factor for spreading future tense categories. In the case of de-venitive futures of Europe, though, one might argue that the evidence for contact-induced spread is compelling. First, these futures are confined to two areas each having a documented history of language contact. One area is mainland Scandinavia, including Danish, Norwegian, Swedish, and marginally Finnish. The second area is Switzerland, where Romansh dialects and Schwyzertütsch (Swiss German) share the grammaticalization from a verb for 'come' (*vɛɲ* 'come' and *k^{h}un* 'come', respectively) to future tense marker (Weinreich [1953] 1964: 41; see Dahl 2000b: 320–1 for details). Second, common genetic origin can be ruled out, at least in the case of Switzerland, where a Romance and a Germanic language are involved.[29]

An example of a future derived from an obligation marker ('have to') is provided by Molisean, a variety of the Slavic Croatian minority in Molise, southern Italy, which has been strongly influenced by Italian (Breu 1996: 26–7). Molisean has an inherited de-volitive future tense, but on the model of Italian it has developed a second future using *imat* 'have' as an auxiliary which, like the

corresponding Italian future, expresses both necessity and future tense, while the inherited future acquired a probability meaning as a result of the rise of the new future (Breu 1996: 30).

An instance of a de-volitive future is found e.g. in the Balkan languages,[30] in fact it is among the uncontroversial morphological features that have been adduced to define the Balkan sprachbund. Joseph describes the situation thus: "A future tense based on a reduced, often invariant, form of the verb 'want' is found in Greek, Tosk Albanian, Rumanian, Macedonian, Bulgarian, Serbo-Croatian, and Romani" (Joseph 1992: 154). That the relevant markers have a "reduced" and "invariant" form is a predictable result of grammaticalization (see Heine & Kuteva 2002: 2), whereby in the process of developing into a tense marker, the erstwhile verb 'want' loses in verbal properties, eventually turning into an invariable marker (decategorialization), and also tends to be phonetically reduced (erosion).[31] As is common with sprachbund situations, there is no reliable evidence on which language provided the model for grammaticalization but, again, there can be hardly any doubt that language contact played some role in the spread of this grammaticalization.

That de-volitive futures in Balkan languages are due to language contact is suggested e.g. by the fact that the Romani varieties spoken in the Balkans have developed a future category marked with *ka(m)-*, which derived from *kam-av* 'want, love' (Boretzky 1989: 368). The role played by areal pressure on shaping equivalence relations and grammaticalization can be demonstrated perhaps most clearly by using this language as an example. Romani offers a perhaps extreme case of how speakers can adapt their modes of grammatical categorization to those of their neighbors. In the Balkans, adaptation had the effect that Romani speakers developed a de-volitive future; elsewhere in Europe they found other models and, accordingly, developed other kinds of future categories. In the dialect of Wales they developed a de-allative future on the model of English *be going to* (Boretzky 1989: 368).

In Russian Romani dialects there is a future tense using the verb *l-av* 'take' as an auxiliary (e.g., *l-av te xav* 'I am going to eat') — a grammaticalization that is crosslinguistically quite unusual. Now, Ukrainian has two future tenses, and one of them also uses the verb 'take' as a future auxiliaby. Boretzky (1989: 369) suggests that Vlach Romani speakers acquired their 'take'-future when they crossed Ukrainian territory. Quite a different situation is found in the Sinti variety that was influenced by German: There is no formal future expression in this variety; rather, the present tense is used to refer to events in the future. This

is exactly the situation found in colloquial German, where the present tense also tends to be used for future events.

What this situation suggests is that future tense exists in Romani only to the extent that it is replicated from languages with which Romani speakers came into close contact.[32]

What we observed with respect to the situation in Europe also applies to other parts of the world: New future tenses appear to be constantly arising in language contact situations on the basis of the principles sketched at the beginning of this section. An example of a de-allative future is provided by the Aztecan language Pipil. Harris and Campbell (1995: 148–9) observe that the former future suffixes found in Pipil are extremely rare today, unused and for the most part unknown. Instead, Pipil speakers tend to use a 'go-to'-future, illustrated in (18), which these authors attribute to the influence of the local Spanish de-allative future (*lo voy a hacer* 'I'm going to do it').[33] Once again, we seem to be dealing with the replication of a grammaticalization process that has a universal base but appears to have been influenced by language contact. Note, however, that Pipil speakers used a serial schema (where both the main verb and the auxiliary are inflected for and agree in person; see Heine 1993), whereas local Spanish uses a plain motion schema (where the main verb is presented in a non-finite form as an allative/locative complement of the auxiliary). What this suggests (assuming that this is in fact an instance of contact-induced change) is that Pipil speakers did not replicate the syntax of the Spanish construction but rather a conceptual structure, that is, a motion schema.

(18) Pipil (Harris & Campbell 1995: 149)
ni- yu ni- k- chiwa.
I- go I- it- do
'I'm going to do it.'

Language contact outside Europe also provides evidence for the grammaticalization of temporal adverbs to future tense markers. A frequently cited example is the development of English *by-and-by* to a future tense marker *(baimbai >) bai* in Tok Pisin and other varieties of Melanesian Pidgin (Sankoff & Laberge 1973). Another example can be found in the Nilotic languages Luo and Kalenjin, which have developed temporal adverbs into future tense markers on the model of temporal distinctions obtaining in neighboring Bantu lanugages, grammaticalizing a common use pattern in their own languages according to which temporal adverbs constitute the primary means of expressing distinctions in deictic time (see Dimmendaal 2001 for details).

An even less common pathway has been chosen by speakers of the Bantu language Ilwana, spoken along River Tana in eastern Kenya: They have developed a future tense by means of a structure [non-past marker – nominalizer + verb], apparently on the model of neighboring Orma, a dialect of the East Cushitic language Oromo, which uses a structure [verb + nominalizer – auxiliary 'be (somewhere)'];[34] note that the two languages have a shared history of over three centuries of contact leading to a substantial restructuring of Ilwana grammar (Nurse 1994; 2000: 149–51). That we are dealing with an instance of replica grammaticalization is suggested in particular by the fact that the pattern used in the two languages is encountered in other Oromo dialects but is absent in Bantu languages, and it is rarely found elsewhere in the world.

To conclude, it is essentially the entire range of options that are universally available for developing categories of future tense that has been exploited in situations of language contact. However, which particular option is chosen appears to be determined by the particular circumstances surrounding the contact situation, such as the use patterns of both the model and the replica languages.

2.4 On polysemy copying

As we pointed out earlier (section 2.2), one might argue that there is an alternative analysis to the one proposed here, namely that, instead of a grammaticalization process, we are dealing with the replication of a polysemy pattern. In fact, there are cases of grammatical transfer where an analysis in terms of polysemy copying (or grammatical calquing) provides the most plausible hypothesis.[35] This applies to cases where speakers of the replica language, rather than replicating that process, appear to have used a shortcut by simply copying the initial and the final stages of the process. A paradigm case can be seen in the grammaticalization of interrogative pronouns to relative clause markers that we discussed in section 2.2. This is a complex grammaticalization process, requiring specific contexts and extending over long periods of time. It is unlikely that speakers of Basque, Romani, or Tariana (= R) were able to reconstruct and replicate this process, the results of which they observed in their respective model languages (= M); more likely they simply observed a polysemy pattern whereby the model languages used the same marker for presenting interrogative and relative clauses (My = Mx) , and they replicated this pattern in their own languages (Ry = Rx). Conceivably, our Silesian Polish example in section 2.2 concerning the extension of the third person plural pronoun to be also used for polite second person singular address on the model

of German involved polysemy copying rather than a gradual process of grammaticalization (Claudia Riehl; p.c.).

What these observations suggest is that grammaticalization does not provide the only strategy available; rather, polysemy copying constitutes an alternative for transferring grammatical categories from the model language to the replica language.

Still, on the basis of the data that are available it would seem that polysemy copying does not appear to constitute the primary strategy used to transfer grammatical categories from one language to another. Returning to an example that we already mentioned in section 2.3.1 may be of help to look into this issue. Basque as spoken in southwestern France has had centuries of close contact with Gascon and later on with French. As a result of this contact, Basque speakers introduced a category which they did not have previously, namely an indefinite article (Haase 1992: 59–61, 71). In doing so they appear to have used the following strategy: Being aware that speakers of the Romance languages (= M) had a grammaticalized concept of an indefinite marker, and that they used the same form (cf. French *un(e)*) for the numeral 'one' (= My) and the indefinite article (= Mx), Basque speakers developed an equivalent category in their own language (= R) by using a grammaticalization process that they may have assumed had taken place in the model languages: They developed their numeral for 'one' (= Ry), *bat*, into an indefinite article (= Rx)[36] — a distinct case of replica grammaticalization (2.2).[37]

On the basis of a polysemy-copying hypothesis one might argue that Basque speakers were aware that the relevant French (or Gascon) forms were polysemous between a numeral meaning (= My) and an indefinite article meaning (= Mx) and they simply copied this polysemy in their own language, by using their equivalent form *bat* for both meanings, Ry and Rx. While in a number of cases that we will be presenting in the following paragraphs such an analysis cannot be ruled out, in cases where we have more detailed information, an analysis in terms of polysemy copying does not seem to be supported by facts, for the following reason: If the polysemy analysis were correct then one would expect that the replica language had the same polysemy structure as the model language. This, however, does not seem to be the case: Not only is the use of the indefinite article in Basque much more contextually constrained, it also exhibits a clearly less advanced stage of grammaticalization than its equivalent in the model languages.[38]

What this means with reference to our Basque example is the following: The grammaticalization of indefinite articles normally proceeds along the following

main stages (see Heine 1997b): First, the article appears as a presentational marker, e.g. in narratives, subsequently its use is generalized to indefinite specific reference, and finally it may be extended to be used even in contexts involving generic reference. While the French indefinite article has gone essentially through all these stages, the Basque indefinite article has not. According to Haase (1992: 60), occasional uses of *bat* as a specific marker are attested already in 1782. While there are incipient uses as a non-specific marker as early as 1545, the grammaticalization as a non-specific article is clearly a recent innovation of Basque. That it has not reached the same degree of grammaticalization as e.g. the corresponding French article is suggested, first, by the fact that its use is optional in a number of cases and, second, that it is not used in some cases where in French the indefinite article is obligatory.

But it would seem that there is another piece of evidence to suggest that an analysis in terms of grammatical calquing or polysemy copying, while being appropriate in some cases (see 2.4), is inadequate in others. Consider the following example.

(19) Sranan (English-based creole; data adapted from Bruyn 1996)
na a oso baka
LOC DET house back
'behind the house/at the back of the house'

(20) Ewe (Kwa, Niger-Congo)
le xɔ- á megbé
LOC house- DET back
'behind the house/at the back of the house'

Sranan is an English-based creole of Surinam, while Ewe is a Kwa language of the Niger-Congo family, spoken in southeastern Ghana and southern Togo. Ewe and other varities of the Gbe language cluster were the native languages of many of the slaves in Surinam in the second half of the 17th centuby. Sentences (19) and (20) exhibit a number of striking similarities, in particular the following: (i) Each uses two adpositional elements, i.e., both contain a general-purpose preposition (*na* and *le*, respectively) and (ii) a postposition, (iii) the postpositions can be assumed to be derived from a noun for the body part 'back' (*baka* and *megbé*, respectively), and (iv) the "postposition" is ambiguous between a grammatical meaning ('behind') and a lexical meaning ('back part of the body'). These structural properties suggest that (19) and (20) are each part of a grammaticalization chain[39] extending from lexical noun ('back') to

grammatical marker ('behind'), where there is an intermediate stage of ambiguity between the lexical and the grammatical structure.

That this is not an isolated case of transfer is suggested by another example volunteered by Bruyn (1996: 40): Both Sranan and Ewe have a lexical verb 'give' (Sranan *gi* [< English *give*], Ewe *ná*) which is also used as a benefactive/dative marker ('for'/'to'), and once again we seem to be dealing with a chain of grammaticalization extending from lexical to grammatical structure: (i) The lexical meaning ('give') is found when the item is used as a main verb, while the grammatical meaning ('for'/'to') is confined to contexts where the item occurs as the second verb in a serial verb construction, (ii) in both languages the erstwhile verb acquires properties of a preposition, and (iii) in both languages there are uses where the item is ambiguous between the lexical and the grammatical meaning (see Lord 1993 for more details on this process).

Such similarities are unlikely to be coincidental,[40] and Bruyn (1996) argues convincingly that they are due to what she calls apparent grammaticalization, that is, to a process whereby a grammaticalization process was tranferred from a substrate language (Ewe or a related language) to the target language Sranan — in other words, Sranan modeled a grammaticalization process on a similar structure in one or more of its major substrate languages. To conclude, we are dealing with a canonical instance of replica grammaticalization (section 2.2).

What these two cases show is that in the process of grammaticalization there is an intermediate step of overlap between the earlier (non-grammaticalized) and the later (grammaticalized) meanings — with the effect that there is ambiguity between the two meanings.[41] The fact that such an overlap situation exists both in the model and the replica languages suggests that the latter has in fact undergone a grammaticalization process.

Finally, there is another piece of evidence to suggest that, rather than copying a static situation, most of the cases discussed here involve a process. In section 2.3.2 we referred to the Balkan sprachbund, which is characterized *inter alia* by the presence of a de-volitive future tense. While it is largely unclear which language or languages provided the ultimate model, it seems obvious that the languages that acquired this category did not replicate a polysemy involving the lexical meaning 'want' and the grammatical meaning future; rather, each language appears to have gone through a process where the verb gradually lost in lexical properties and gained in grammatical properties. This process is reflected in the present-day structure of the future tense markers which, as Joseph (1992: 154) observes, tend to exhibit a reduced and often invariant form — all characteristics of an extended process of grammaticalization. The situation in

Balkan Romani that we mentioned in section 2.3.2 illustrates this situation, where the future marker *ka(m)-* (e.g., *ka-džav* 'I will go') represents the grammaticalized product of the verb *kam-av* 'want, love' (e.g., *kamav te džav* 'I want to go'; Boretzky 1989: 368).

It would seem in fact that the majority of cases described in this paper cannot be accounted for satisfactorily with reference to notions such as calquing, loan translation, or polysemy copying. The evidence we presented was of four kinds. First, we saw that in ordinary contact-induced grammaticalization (2.1) the model language (= M) provides a model for a category (= Mx) but not for how to develop an equivalent category (= Rx) in the replica language (= R). In other words, there is no possible polysemy that could be copied, and the only way of acquiring Rx is via grammaticalization of some other category Ry.

Second, it turns out that in the majority of cases examined the replica construction is less grammaticalized than the corresponding model construction.[42] For example, in the initial stage of grammaticalization, the new category tends to be ambiguous between its literal and its grammaticalized meaning, it tends to be confined to few contexts, and its use is optional — a situation that has not seldom led to controversies among grammarians on whether or not the relevant category really exists in the language concerned. Such properties are commonly encountered in replicated categories. For example, we noted above that both the definite article in Sorbian, replicated on the model of German, and the indefinite article of Basque, replicated on the model of Romance languages, exhibit properties of categories in the early stages of grammaticalization, as is suggested e.g. by the fact that their use is contextually restricted and optional to some extent. They thus differ from the corresponding categories in the model languages, which both are fully grammaticalized articles.

Third, the presence of overlap situations in the replica language suggests that what is copied is not simply a polysemy pattern but rather a full process including the intermediate stage of gradual transition from source to target category where both coexist side by side, giving rise to semantic and morphosyntactic ambiguity. And finally, replicated categories tend to exhibit properties that bear witness to their grammaticalization history, that is, desemanticization (loss of meaning components), decategorialization (loss of morphosyntactic properties) and erosion (phonetic reduction or simplification).

2.5 On directionality

Grammaticalization has been shown to be essentially unidirectional, that is, at least ninety per cent of all instances of grammatical change can be assumed to be in accordance with principles of grammaticalization (Newmeyer 1998: 275–6, 278; Haspelmath 1999: 1046). The question that arises here is whether this also applies to grammaticalization occurring in or being triggered by language contact.

It would seem that the answer is in the affirmative: Essentially all cases that we found are in accordance with the unidirectionality hypothesis. There are exceptions, however. One case concerns the post-verbal perfect markers of the Malaita Oceanic languages that we dealt with in section 2.1. As Keesing (1991: 331) argues, these markers have been further grammaticalized to topicalizing particles.[43] This process appears to have been replicated in Solomons Pijin,[44] where the replica language exhibits the same structure as the model language Kwaio or other Oceanic languages of Malaita. Now, to our knowledge, the development from a verbal aspect marker to a topic (or focus) marker has so far not been documented, nor does it seem conceptually plausible — hence we might be dealing with a counter-example to grammaticalization. Another case of a violation of the unidirectionality principle concerns a situation that appears to be specific to language change in contact situations. When establishing equivalence relations with the patterns of categorization they observe in the model language, it may happen that speakers equate some category of their language with a less grammaticalized category of the model language, with the possible effect that a more grammaticalized category changes to a less grammaticalized one in the replica language. For example, while the grammaticalization of verbs as prepositions is well documented, a reversed directionality is hard to find.[45] Still, there is an example in at least one creole: what are prepositions in Dutch and English (e.g. Dutch *door* 'through') end up as items having verbal uses in the English-based creole Sranan (e.g. Sranan *doro* 'put through, go through'). Presumably under the substrate influence of West African languages, speakers of Sranan appear to have equated their prepositions (or adverbs) with verbs in the creole replica language (Bruyn 1996).

Another possible example is provided by Basque as spoken in southwestern France, which, as we saw above, has undergone major restructuring under the influence of language contact with two Romance languages, Gascon and French. Basque has a closed word class consisting of nine "modal operators", serving modal, evidential, and aspectual functions; they are invariable and

immediately precede the auxiliary (Haase 1992: 111–24; see section 2.3.1). The model languages Gascon and French do not have an equivalent word class. Now, in an attempt to establish equivalence with the model languages, Basque speakers appear to have equated their modal operators with auxiliary verbs in these Romance languages. The result is that Basque speakers in contact with these Romance languages are losing their category of modal operators in that the latter gradually develop into a new class of modal auxiliaries. Two of these items, *ahal* 'be able to' and *nahi* 'want', have acquired verbal properties and two others, *behar* 'must' and *ari* progressive, have already developed uses of full verbs (Haase 1992: 123–4).

To summarize, directionality in contact-induced language change is on the whole in line with principles of grammatical change to be observed elsewhere, but it may sometimes take unexpected directions when speakers aspire to achieve intertranslatability between the model and the replica languages.

3. Conclusions

The present paper raises more questions than it is able to answer. A key question that we were unable to answer is the following: To what extent are the developments sketched here the result of universal principles of grammaticalization and to what extent are they due to the specific factors obtaining in situations of language contact? On the basis of the scanty data that are available we can say no more than that there must be some "conspiracy" of both factors (cf. Plag forthc.), but what exactly this means is open to further research. Related to this is the question of whether language contact as observed in pidgins and creoles is qualitatively different than that observed in other languages. The data available suggest that contact-induced grammaticalization does not behave essentially differently in these two contrasting cases, but more data is needed for a more qualified assessment.

What the observations made in the preceding sections seem to suggest is that in addition to the kinds of motivation that have been defined so far (see e.g. Heine, Claudi, & Hünnemeyer 1991; Hopper & Traugott 1993) there are others that have to be taken into account in the study of grammaticalization. One appears to be communicative in nature, viz. to make the categories existing in the languages that are in contact mutually compatible and more readily intertranslatable. Other motivations may be to talk like one's neighbors for social, political, psychological, economic, or other reasons.

These observations furthermore suggest that contact-induced change as studied here is gradual rather than abrupt, typically involving a process of the following kind (see section 2.2): In order to develop a structure that is equivalent to the one in the model language, speakers choose among the use patterns that are available in the replica language the one that corresponds most closely to the model, frequently one that until then was more peripheral and of low frequency of use, and they activate it — with the effect that a peripheral pattern gradually turns into the regular equivalent of the model, acquires higher frequency of use, and eventually it may emerge as a full-fledged grammatical category.

It has been a common practice to classify instances of language change in terms of whether they are the result of language-internal and/or historically definable factors or else whether they are due to universal cognitive, communicative, or other constraints. What the discussion in this paper may have shown is that these are in no way mutually exclusive alternatives;[46] on the basis of our data on contact-induced grammaticalization both can be expected to be potentially present in some way in virtually any given case of language change.

Notes

* We wish to express our gratitude to Bernard Comrie, Martin Haspelmath, Claudia Riehl, and Thomas Stolz, and two anonymous referees for valuable comments on an earlier version of this paper. We also thank the DFG for the generous financial support.

1. For example, the fact that in many European and other languages of the Old World there is a striking similarity between possessive constructions (e.g. *He has a dog.*) and perfect constructions (*He has gone.*) has given rise to controversies on whether this similarity is due to historical/conctact-induced or to universal cognitive factors (see e.g. Jacob 1998: 106).

2. For example, in an important paper on grammatical development in creole languages, Plag concludes: "I will argue that only by differentiating internal and contact-induced changes can we reach important new insights into the nature of both creolization and grammaticalization." (Plag forthc.:5)

3. Note that we do not wish to claim that the processes summarized in (1) are not complex (see for example Ramat & Bernini 1990: 25–6).

4. Contact-induced grammaticalization coincides with Kuteva's (2000) term areal grammaticalization and comes close to but is more complex than what is referred to as "grammaticalizing metatypy" in Heine and Kuteva (2001: 402, 408).

5. For a similar use of this term, see Bruyn (1995).

6. The dual and trial markers of Proto-Oceanic appear to be grammaticalized forms of the numerals **dua* 'two' and **tolu* 'three', respectively, and in modern Oceanic languages, the dual and plural markers tend to exhibit a form similar to that of the respective numerals (see

Keesing 1988: 71ff.). Conceivably this similarity provided a basis for grammaticalization by pidgin speakers; we will return to this issue below.

7. As far as we are aware, Ross does not provide a definition of metatypy; in accordance with his descriptions it would seem that metatypy can be defined as the wholesale restructuring of a language's semantic and syntactic structures on the model of another language, leading to a new typological profile in the replica language on the one hand, and to a large degree of direct intertranslatability with the model language on the other.

8. At present, bilingualism in Takia and Waskia is not particuarly common (Ross 2001: 147); there does not appear to be any clear information on the nature of bilingualism in the past.

9. The evidence for this reconstruction is confined to two of the eight Takia postpositions presented by Ross (2001: 143).

10. Concerning word order change in Takia, see section 2.2.

11. These forms are taken from Upper Sorbian; Lower Sorbian has slightly different forms for the demonstrative (see Lötzsch 1996: 52).

12. The same strategy was employed by speakers of other Slavic languages in the Balkans, who also drew on demonstratives to develop definite articles (Breu 1996: 30).

13. Glosses are ours; there are no glosses in the original.

14. Rather than perfect markers, adverbs meaning 'now' are more likely to give rise to aspectual constructions denoting that an action is still going on (Heine & Kuteva 2002: 218).

15. Jacob (1994: 53) makes an important observation in this connection: "Wichtig ist aber, dass Lehnbildungen immer eine gewisse intuitive Suggestivität, d.h. "Natürlichkeit" der Ausgangsstruktur als Motivation für die Nachbildung durch die entlehnenden Sprecher voraussetzen."

16. Instead of languages, the strategy may as well be applied between different dialects of the same language.

17. What this aspectual category expresses is an event that is located at a point that is separated from but temporally close to the moment of speaking, for example *John has just arrived* (see Harris 1991: 201ff.).

18. Note that the object ('boat') precedes the participle in Irish but follows it in Irish English. We will not deal with this issue here; it is a complex one that has received several contrasting interpretations. A possible intermediate stage between the Location Schema and the hot-news perfect in Irish English can be seen in a construction where instead of a verbal complement the preposition *after* has a nominal complement, e.g. *He's after the flu* 'He just had the flu' (Sullivan 1980: 205).

19. But see alsoTrask (1998: 320).

20. In some cases also by the inessive and absolutive cases (Haase 1992: 69). The fact that in certain contexts the instrumental suffix is not replaced by the comitative but rather by the absolutive, which has zero expression, might suggest that the present case of contact-induced grammaticalization is not the only factor responsible for the decline of the instrumental case.

21. Concerning the grammaticalization comitative > instrumental, see Luraghi (2001), Stolz (2001), Heine & Kuteva (2002).

22. Glosses are ours; no glosses are provided by the author.

23. Haig (2001: 215–6) uses the form *-şa* in the running text but *-ša* in his table 7.

24. It goes without saying that this is not the only way in which contact-induced word order change may occur (Bernard Comrie; p.c.); our interest here is confined to changes resulting from grammaticalization.

25. Trask offers an alternative view of this issue: "There is virtually no observable tendency for IE morphological features to be transferred into Basque" (Trask 1998: 319), and Bernard Comrie (p.c.) observes that "Basque certainly doesn't seem to borrow many bound morphemes from Romance, nor to change its overall morphological typology (i.e. it retains a rich nominal morphology, also a rich verb morphology but virtually restricted to a few auxiliary verbs). In the nominal system, there is some shifting of the content of morphological categories under contact. And in the verb, it is largely the development of periphrastic constructions."

26. Haase (1992: 81) refers to this process as the creation of translational equivalents via loan translation.

27. That infinitives of model languages tend to be replicated by means of nominalization markers is possibly a more general phenomenon; for example, Johanson (1998: 332) observes that combinatorial properties of the Persian infinitive are often copied onto the Southern Azerbaijanian verbal noun in — *mAK*, that is, a category of the model language Persian is replicated to some extent in a language that has been strongly influenced by Persian.

28. Among the features we discussed in this section there are the presence or emergence of an indefinite article, a 'have'-perfect, and of relative clauses with an introducing relative pronoun. These are all features that Haspelmath (2001) proposes to be diagnostic of Standard Average European but finds to be absent in Basque.

29. We are ignoring here Finnish on account of its marginal status. Dahl (2000b: 320) says that the Finnish construction "has been looked upon with some suspicion as being a Swedish calque."

30. Other European languages are English (*will*), Danish (*vil*), Norwegian Bokmål (*vil*), Faroese (*vil*), and Frisian *wal*; Dahl 2000b: 322).

31. With regard to their relative degree of grammaticalization, two types of Balkan languages can be distinguished: one where there is still an inflected auxiliary (Romanian, Serbian, and Croatian), and another one where the resulting future marker ends up as an uninflected particle (Modern Greek, Bulgarian, Macedonian, and Albanian (cf. Dahl 2000b: 323).

32. Boretzky (1989: 369) concludes that on their way to Europe, Romani speakers did not have a future tense in their language.

33. In addition to Spanish, a de-allative future is found in other European languages: Portuguese, French, Dutch, and English (Dahl 2000b: 319).

34. Note that the linear arrangement in Ilwana is the mirror image of that found in Orma, for obvious reasons: While the SVO-language Ilwana places auxiliaries before the nominalizer/infinitive marker and the main verb, the SOV-language Orma has the reverse order.

35. We are grateful to Martin Haspelmath and Claudia Riehl for having drawn our attention to this fact.

36. Haase describes this process thus: "Anders ausgedrückt: *bat* und frz./gask. *un* sind im Bereich der Zahlwörter *Übersetzungsäquivalente*. Wie in anderen Fällen von Übersetzungsäquivalenz, die wir später noch kennenlernen werden, kann sich nun der Funktionsbereich von *bat* auf alle die Fälle ausbreiten, in denen in den Modellsprachen *un* gebraucht wird, also auch auf die Signalisierung von Indefinitheit." (Haase 1992: 59)

37. Since the grammaticalization of a numeral 'one' into an indefinite marker is commonly attested across languages, a language-internal development cannot be ruled out entirely here.

38. It goes without saying that in the course of time the replica category may attain the same degree of grammaticalization as the model category, but this usually takes centuries to happen.

39. Concerning the term "grammaticalization chain", see Heine 1992.

40. With the epithet "are unlikely" we wish to draw attention to the fact that there are also alternative ways of accounting for the properties of these two cases. First, the grammaticalization of a body part noun 'back' as a locative adposition and of a verb for 'give' to a benefactive/dative adposition are universally attested (Heine & Kuteva 2002), and this would also account for the fact that there are intermediate stages characterized by ambiguity between the lexical and the grammatical uses. Second, that Sranan developed postpositions such as *baka* 'behind' is not surprising, considering the fact that one of the major Sranan constructions of attributive possession consists of the juxtaposition possessor and the possessee, e.g. *a datra oso* (the doctor house) 'the doctor's house' (Bruyn 1996: 33). On the basis of this construction one could predict that a grammaticalization of this construction to an adpositional construction would yield postpositions rather than prepositions. Third, that there is a preposition in addition, viz. the multi-purpose preposition *na*, is in no way surprising since it could be argued that this is a retention of a lexifier pattern. And finally, that the grammaticalization of a verb for 'give' to a preposition-like item is confined to uses as the second verb in verb serialization is also commonly encountered crosslinguistically (see Lord 1993). Nevertheless, while internal development cannot entirely be ruled out, the overall description volunteered by Bruyn suggests that substrate influence is clearly a more plausible hypothesis and we adopt it here.

41. This situation is described in terms of what is called the *overlap model*, taking the form A > A/B > B (Heine 1993).

42. It goes without saying that, given enough time, the replica construction may also develop into a fully grammaticalized category.

43. On the basis of the translations volunteered by Keesing it would seem that the "topicalizing particles" are perhaps more appropriately described as focus markers.

44. Phonological similarity may have contributed to the equation of the Oceanic perfect markers (e.g., *no'o, na'a, na*) with Solomons Pijin *nao* (Keesing 1991: 329; see section 2.1).

45. We are ignoring here cases of lexicalization; see Heine 2002.

46. See Mufwene (1986) on a similar view concerning pidgins and creoles.

Abbreviations

A	subject of transitive clauses
ABL	ablative
AL	allocutive
ART	article
BEN	benefactive
CAU	causative verb
COM	comitative
CONJ	conjugational verb
DAT	dative
DEM	demonstrative
DET	determiner
DIR	directive
DU	dual
DUR	durative
EXCL	exclusive
F	feminine
FUT	future
IN	inessive
INCL	inclusive
INF	infinitive
INSTR	instrumental
M	model language, masculine
N	neuter
NF	non-feminine
O	object
PART	partitive
PERF	perfect
PL	plural
PREP	preposition
PRES	present
PROG	progressive
PRT	preterit
PTCP	participle
R	replica language
S	subject of intransitive clauses
SG	singular
SRP	subject-referencing pronoun
TOP	topic
TRN	transnumeral
TRS	transitive suffix
1, 2, 3	first, second, third person

References

Aikhenvald, Alexandra Y. 1996. "Areal diffusion in northwest Amazonia: The case of Tariana". *Anthropological Linguistics* 38:73–116.

Aikhenvald, Alexandra Y.; and R.M.W. Dixon, R.M.W. (eds.) 2001. *Areal diffusion and genetic inheritance: Problems in comparative linguistics*. Oxford: Oxford University Press.

Baker, Philip; and Anand Syea (eds.) 1996. *Changing meanings, changing functions: Papers relating to grammaticalization in contact languages*. London: University of Westminster Press [Westminster Creolistics Series 2].

Bakker, Peter; and Mous, Maarten (eds.) 1994. *Mixed languages: 15 case studies in language intertwining*. Amsterdam: Institute for Functional Research into Language and Language Use (IFOTT).

Bechert, Johannes; Bernini, Giuliano; and Buridant, Claude (eds.) 1990. *Toward a typology of European languages*. Berlin: Mouton: De Gruyter [Empirical Approaches to Language Typology 8].

Bernini, Giuliano; and Ramat, Paolo. 1996. *Negative sentences in the languages of Europe: A typological approach*. Berlin: Mouton de Gruyter [Empirical Approaches to Language Typology 16].

Bisang, Walter 1991. Verb serialization, grammaticalization and attractor positions in Chinese, Hmong, Vietnamese, Thai and Khmer. In Seiler and Premper 1991: 509–62.

———. 1996. "Areal typology and grammaticalization: Processes of grammaticalization based on nouns and verbs in East and Mainland South East Asian languages". *Studies in Language* 20(3):519–97.

———. 1998a. "Grammaticalization and language contact, constructions and positions". In: Giacalone Ramat & Hopper 1998: 13- 58.

Borg, Albert: and Mifsud, Manwel. 2002. "Maltese object marking in a Mediterranean context". In Ramat and Stolz 2002: 33–46.

Boretzky, Norbert 1986. Regelentlehnung und Substrateinfluss in Kreolsprachen. In: Boretzky et al. 1986: 9–39.

———. 1989. "Zum Interferenzverhalten des Romani". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42(3): 357–74.

Boretzky, Norbert; Enninger,Werner; and Stolz, Thomas (eds.) 1986. *Beiträge zum 2. Essener Kolloquium über "Kreolsprachen und Sprachkontakte"*. Bochum: Brockmeyer [Essener Beiträge zur Sprachwandelforschung 2].

Boretzky, Norbert, Werner Enninger, & Thomas Stolz (eds.) 1989. *Beiträge zum 5. Essener Kolloquium über "Grammatikalisierung: Natürlichkeit und Systemökonomie"*. Volume 1. (Essener Beiträge zur Sprachwandelforschung, 7.) Bochum: Brockmeyer.

Boretzky, Norbert; and Igla, Birgit 1994. Romani mixed dialects. In: Bakker and Mous 1994: 35–68.

Breu, Walter (ed.) 1990a. *Slavistische Linguistik 1989*. Munich: Otto Sagner.

———. 1990b. "Sprache und Sprachverhalten in den slavischen Dörfern des Molise (Süditalien)". In Breu 1990a: 35–65.

———. 1992. "Das italokroatische Verbsystem zwischen slavischem Erbe und kontaktbedingter Entwicklung". In Reuther 1992: 93–122.
———. 1994. "Der Faktor Sprachkontakt in einer dynamischen Typologie des Slavischen". In Mehlig 1994: 41–64.
———. 1996. "Überlegungen zu einer Klassifizierung des grammatischen Wandels im Sprachkontakt (am Beispiel slavischer Kontaktfälle)". *Sprachtypologie und Universalienforschung* 49(1): 21–38.
Bright, William (ed.) 1992. *International encyclopedia of linguistics*. Volume 1. Oxford: Oxford University Press.
Bruyn, Adrienne 1996. "On identifying instances of grammaticalization in Creole languages". In: Baker and Syea 1996: 29–46.
Bybee, Joan L.; Pagliuca, William; and Perkins, Revere D. 1991. "Back to the future". In: Traugott and Heine 1991b: 17–58.
Bybee, Joan L.; Perkins, Revere D. and Pagliuca, William. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.
Campbell, Lyle 1987. Syntactic change in Pipil. *International Journal of American Linguistics* 53 (3): 253–80.
Campbell, Lyle; Terrence Kaufman; and Smith-Stark, Thomas C. 1986. "Meso-America as a linguistic area". *Language* 62 (3): 530–70.
Corne, Chris 1995. "A contact-induced and vernacularized language: How Melanesian is Tayo?" In: Baker 1995: 121–148.
Dahl, Östen 1990. "Standard Average European as an exotic language". In: Bechert, Bernini and Buridant 1990: 3–8.
———. 1996. "Das Tempussystem des Deutschen im typologischen Vergleich". In Lang and Zifonun 1996: 359–68.
———, (ed.) 2000a. *Tense and aspect in the languages of Europe*. Berlin, New York: Mouton de Gruyter [Empirical Approaches to Language Typology, Eurotyp 20–6].
———. 2000b. "The grammar of future time reference in European languages". In: Dahl 2000a: 309–28.
———. 2000c. "Verbs of becoming as future copulas". In: Dahl 2000a: 351–61.
Denison, N. 1968. "Sauris: A trilingual community in diatypic perspective". *Man*, n.s., (3):578–92.
———. 1988. "Language contact and language norm". *Folia Linguistica* 22:11–35.
Dihoff, Ivan R. (ed.) 1983. *Current approaches to African linguistics (vol. 1)*. Dordrecht: Foris Publications.
Dimmendaal, Gerrit J. 1995. "The emergence of tense marking in the Nilotic-Bantu borderland as an instance of areal adaptation". In: Zima 1995: 29–43.
———. 2001. "Language shift and morphological convergence in the Nilotic area". *SUGIA* 16–17: 83–124.
Drewes, A. J. 1994. "Borrowing in Maltese". In Bakker and Mous 1994: 83–111.
Durie, Mark and Ross, Malcom. (eds.) 1996. *The comparative method reviewed: Regularity and irregularity in language change*. New York: Oxford University Press.
Enfield, Nick J. 2001. "On genetic and areal linguistics in mainland south-east Asia: Parallel polyfunctionality of 'acquire'". In: Aikhenvald and Dixon 2001: 255–90.

Filppula, Markku 1986. *Some aspects of Hiberno-English in a functional sentence perspective.* Joensuu: University of Joensuu [University of Joensuu Publications in the Humanities 7].

Gilbert, Glenn (ed.) Forthcoming. *Pidgin and creole linguistics in the 21st century.* New York: Lang.

Givón, Talmy 1971. "Historical syntax and synchronic morphology: An archaeologist's field trip". *Chicago Linguistic Society* 7: 394–415.

Greenberg, Joseph H. 1978. "How does a language acquire gender markers?" In: Greenberg, Ferguson and Moravcsik 1978: 47–82.

Greenberg, Joseph H.; Ferguson, Charles A.; and Moravcsik, Edith (eds.) 1978. *Universals of human language.* Volume 3: *Word structure.* Stanford: Stanford University Press.

Gumperz, John J.; and Wilson, Robert. 1971. "Convergence and creolization: A case from the Indo-Aryan/Dravidian border in India". In Hymes 1971: 151–67.

Haase, Martin 1992. *Sprachkontakt und Sprachwandel im Baskenland: die Einflüsse des Gaskognischen und Französischen auf das Baskische.* Hamburg: Buske.

Haase, Martin; and Nau, Nicole (eds.) 1996. *Sprachkontakt und Grammatikalisierung. Sprachtypologie und Universalienforschung.* Berlin: Akademie-Verlag [Special issue, 49, 1].

Haig, Geoffrey 2001. "Linguistic diffusion in present-day East Anatolia: From top to bottom". In: Aikhenvald and Dixon 2001: 195–224.

Harris, Alice C.; and Campbell, Lyle. 1995. *Historical syntax in cross-linguistic perspective.* Cambridge: Cambridge University Press.

Harris, John 1991. "Conservatism versus substratal transfer in Irish English". In: Trudgill and Chambers 1991: 191–212.

Harris, Martin; and Ramat, Paolo (eds.) 1987. *Historical development of auxiliaries.* Berlin: Mouton de Gruyter [Trends in Linguistics, 35].

Haspelmath, Martin. 1998. "How young is Standard Average European?" *Language Sciences* 20(3):271–87.

———. 1999. "Why is grammaticalization irreversible?" *Linguistics* 37(6):1043–68.

———. 2001. "The European linguistic area: Standard Average European". In Haspelmath et al. 2001: 1492–1510.

Haspelmath, Martin; König, Ekkehard; Oesterreicher, Wulf; and Raible, Wolfgang (eds) 2001. *Language typology and language universals: An international handbook.* Volume 2. Berlin, New York: Walter de Gruyter.

Heath, Jeffrey 1978. *Linguistic diffusion in Arnhem Land.* Canberra: Australian Institute of Aboriginal Studies [Australian Aboriginal Studies Research and Regional Studies, 13].

Heine, Bernd 1992. Grammaticalization chains. *Studies in Language* 16(2): 335–68.

———. 1993. *Auxiliaries: Cognitive forces and grammaticalization.* Oxford: Oxford University Press.

———. 1994. "Areal influence on grammaticalization". In :Pütz 1994: 56–68.

Heine, Bernd; Claudi, Ulrike; and Hünnemeyer, Friederike 1991. *Grammaticalization: A conceptual framework.* Chicago: University of Chicago Press.

Heine, Bernd; Kuteva, Tania. 2001. "Convergence and divergence in the development of African languages". In: Aikhenvald and Dixon 2001: 393–411.

———. 2002. *World lexicon of grammaticalization.* Cambridge: Cambridge University Press.

Heine, Bernd and Reh, Mechtild 1984. *Grammaticalization and reanalysis in African languages.* Hamburg: Buske.

Hopper, Paul J. and Traugott, Elizabeth C. 1993. *Grammaticalization.* Cambridge: Cambridge University Press.

Hurch, Bernhard 1989. "Hispanisierung im Baskischen". In: Boretzky, Enninger, and Stolz 1989: 11–35.

Huffines, Marion Lois 1986. "The function of aspect in Pennsylvania German and the impact of English". *Yearbook of German-American Studies* 21:137–54.

———. 1991. "Pennsylvania German: Convergence and change as strategies of discourse". In: Seliger and Vago 1991: 125–37.

Hymes, Dell (ed.) 1971. *Pidginization and creolization of languages: Proceedings of a conference held at the University of the West Indies, Mona, Jamaica, April 1968.* Cambridge: Cambridge University Press.

Jacob, Daniel 1994. Die Auxiliarisierung von habere und die Entstehung des romanischen periphrastischen Perfekts, dargestellt an der Entwicklung vom Latein zum Spanischen. Habilitationsschrift, University of Freiburg.

Johanson, Lars 1992. Strukturelle Faktoren in türkischen Sprachkontakten. In: *Sitzungsberichte der wissenschaftlichen Gesellschaft an der Johann Wolfgang Goethe-Universität Frankfurt am Main*, Volume 29(5): 169–299. Stuttgart: Franz Steiner.

———. 1996. "Kopierte Satzjunktoren im Türkischen". *Sprachtypologie und Universalienforschung* 49 (1): 39–49.

———. 1998. "Code-copying in Irano-Turkic". *Language Sciences* 20(3): 325–37.

Joseph, Brian D. 1992. "The Balkan languages". In: Bright 1992: 153–55.

Keesing, Roger M. 1988. *Melanesian Pidgin and the Oceanic substrate.* Stanford: Stanford University Press.

Keesing, Roger M. 1991. "Substrates, calquing and grammaticalization in Melanesian Pidgin". In: Traugott and Heine 1991a: 315–42.

König, Ekkehard; and Haspelmath, Martin. 1999. "Der europäische Sprachbund". In: Reiter 1999: 111–127.

Kostov, Kiril 1973. "Zur Bedeutung des Zigeunerischen für die Erforschung grammatischer Interferenzerscheinungen". *Linguistique Balkanique* 16: 2:99–113.

Kuteva, Tania 1998. "Large linguistic areas in grammaticalization: Auxiliation in Europe". *Language Sciences* 20(3): 289–311.

———. 2000. "Areal grammaticalization: The case of the Bantu-Nilotic borderland". *Folia Linguistica* 34(3–4): 27–83.

———. 2001. *Auxiliation: An enquiry into the nature of grammaticalization.* Oxford: Oxford University Press.

Lang, Ewald and Zifonun, G. (eds.) 1996. *Deutsch — Typologisch.* Berlin: Mouton de Gruyter.

Li, Charles N. (ed.) 1976. *Subject and topic.* New York: Academic Press.

Lord, Carol. 1993. *Historical change in serial verb constructions.* Amsterdam/ Philadelphia: John Benjamins [Typological Studies in Language 26].

Lötzsch, Ronald 1996. "Interferenzbedingte grammatische Konvergenzen und Divergenzen zwischen Sorbisch und Jiddisch". *Sprachtypologie und Universalienforschung* 49(1): 50–59.

Luraghi, Silvia 2001. "Some remarks on instrument, comitative, and agent in Indo-European". *Sprachtypologie und Universalienforschung* 54(4): 385–401.

Matisoff, James A. 1991. "Areal and universal dimensions of grammaticalization in Lahu". In Traugott and Heine 1991b: 383–453.

Matras, Yaron 1996. Prozedurale Fusion: Grammatische Interferenzschichten im Romanes. *Sprachtypologie und Universalienforschung* 49(1): 60–78.

———. 1998. "Utterance modifiers and universals of grammatical borrowing". *Linguistics* 36 (2): 281–331.

Mehlig, Hans Robert (ed.) 1994. *Slavistische Linguistik 1993*. Munich: Otto Sagner.

Mélanges de linguistique et de philologie offerts à Jacq. van Ginneken à l'occasion du soixantième anniversaire de sa naissance. Paris: Librairie C. Klincksieck.

Michelena, L. 1978. Miscelánea filológica vasca, I. *Fontes Linguae Vasconum* 10(29): 205–28.

Mufwene, Salikoko S. 1986. "The universalist and substrate hypotheses complement one another". In Muysken and Smith 1986: 129–62.

Muysken, Pieter and Smith, Norval. (eds.) 1986. *Substrata versus universals in Creole genesis*. Amsterdam/ Philadelphia: John Benjamins [Creole Language Library 1].

Myers-Scotton, Carol 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.

Newmeyer, Frederick J. 1998. *Language form and language function*. Cambridge MA: The MIT Press.

Nurse, Derek. 1994. "South meets north: Ilwana = Bantu + Cushitic on Kenya's Tana River". In: Bakker and Mous 1994: 213–22.

———. 2000. *Inheritance, contact, and change in two East African languages*. Cologne: Köppe [Sprachkontakt in Afrika, 4].

Plag, Ingo. Forthcoming. "On the role of grammaticalization in creolization". In Gilbert Forthcoming.

Press, Ian. 1986. *A grammar of Modern Breton*. Berlin: Mouton de Gruyter [Mouton Grammar Library 2].

Pütz, Martin. (ed.) 1994. *Language contact and language conflict*. Amsterdam/Philadelphia: John Benjamins.

Ramat, Paolo. 1990. "Area influence versus typological drift in Western Europe: The case of negation". In: Bechert et al. 1990: 25–46.

———. 1998. "Typological comparison and linguistic areas: Some introductory remarks". *Language Sciences* 20 (3): 227–40.

Ramat, Paolo; and Bernini, Giuliano. 1990. "Area influence versus typological drift in western Europe: The case of negation". In: Bechert et al. 1990: 25–46.

Ramat, Paolo: and Stolz, Thomas (eds). 2002. *Mediterranean languages: Papers from the MEDITYP workshop, Tirrenia, June 2000*. Bochum: Brockmeyer [Diversitas Linguarum 1].

Reuther, Tilmann (ed.). 1990. *Slavistische Linguistik 1989*. Munich: Otto Sagner.

Ross, Malcolm. 1996. Contact-induced change and the comparative method: cases from Papua New Guinea. In: Durie and Ross 1996: 180–217.

———. 2001. Contact-induced change in Oceanic languages in North-West Melanesia. In: Aikhenvald and Dixon 2001: 134–66.

Sankoff, Gillian 1980. *The social life of language.* Philadelphia: University of Pennsylvania Press.
Sankoff, Gillian; and Laberge, Suzanne. 1973. "On the acquisition of native speakers by a language". *Kivung* 6:32–47.
Seliger, Herbert W.: and Vago, Robert M. (eds.). 1991. *First language attrition.* Cambridge: Cambridge University Press.
Stolz, Thomas 1996a. Komitativ-Typologie: MIT- und OHNE-Relationen im crosslinguistischen Überblick. *Papiere zur Linguistik* 51(1): 3–65.
———. 1996b. "Some instruments are really good companions — some are not: On syncretism and the typology of instrumentals and comitatives". *Theoretical Linguistics* 23(1/2): 113–200.
———. 1998. "UND, MIT und/oder UND/MIT? Koordination, Instrumental und Komitativ — kymrisch, typologisch und universell". *STUF* 51(2): 107–30.
———. 2001. "On circum-Baltic instrumentals and comitatives". In: Dahl and Koptjevskaja Tamm 2001: 591–612.
———. 2002. Not quite the right mixture: Chamorro and Malti as candidates for the status of mixed languages. Typescript, University of Bremen, Bremen.
Stolz, Christel; and Stolz, Thomas. 2001. "Mesoamerica as a linguistic area". In: Haspelmath et al. 2001: 1542–53.
Trask, R.L. 1998. "The typological position of Basque: Then and now". *Language Sciences* 20(3): 313–24.
Traugott, Elizabeth C.; and Heine, Bernd (eds). 1991a. *Approaches to grammaticalization.* Volume 1. Amsterdam/Philadelphia: John Benjamins.
———, (eds). 1991b. *Approaches to grammaticalization.* Volume 2. Amsterdam/ Philadelphia: John Benjamins.
Ureland, P. Sture. 1984. "The influence of American English on American Swedish — a case study on the nature of interference". In: Ureland and Clarkson 1984: 281–324.
Ureland, P. Sture; and Clarkson, Iain (eds). 1984. *Scandinavian language contacts.* Cambridge: Cambridge University Press.
Weinreich, Uriel [1953] 1964. *Languages in contact.* The Hague: Mouton.

Authors' addresses

Bernd Heine
Institut für Afrikanistik
Universität zu Köln
Meister-Ekkehartstrasse 7
50 923 Köln

bernd.heine@uni-koeln.de

Tania Kuteva
Anglistik III
Heinrich-Heine-Universität Düsseldorf
Geb. 23.21.01
Universitätsstrasse 1
40 225 Düsseldorf

tania.kouteva@phil-fak.uni-duesseldorf.de