3 Grammaticalization

Summarizing some of his findings on language contact in the Balkans, Victor Friedman observes that

the structural convergences called Balkanisms, among which grammaticalized status can be counted, must have begun as discourse-bound variations that resulted in part from communicative needs and desires of multilingual speakers and in part from competing grammatical systems. Balkanisms began as variation when speakers of different languages attempted to communicate more effectively and mediated between the languages of their interlocutors and the structures of their native languages. The place of any given Balkanism in the systems of the various languages can be described in terms of a continuum from pragmatically conditioned variation to grammaticalization, which in turn suggests that discourse functions are not merely subject to borrowing but actually serve as entry points for the development of structural change. (Friedman 2003: 110)

Discourse-bound, pragmatically conditioned variation was discussed in chapter 2 in terms of use patterns. The present chapter highlights the second phase of the process leading to fixed grammatical templates: It is concerned with the emergence of new functional categories and constructions (see also Heine & Kuteva 2003).

3.1 The mechanism

As the preceding chapter may have shown, the transfer of grammatical information from one language to another without involving any linguistic forms is perhaps more widespread than has previously been thought. In its initial stages, replication tends to involve pieces of discourse that acquire higher text frequency, are extended to new contexts, and gradually come to be associated with new grammatical functions. We described this development with reference to the gradual transition from minor to major use patterns, and we observed that this development may lead to the crystallization of new grammatical structures. What the observations made seem to suggest is that virtually any minor use pattern can be activated in language contact and give rise to new functional categories – even if most of these use patterns never succeed in acquiring category status. We will now describe the mechanism giving rise to such categories
in situations of language contact, which we propose to call contact-induced grammaticalization.

Contact-induced grammaticalization is a grammaticalization process that is due to the influence of one language on another. We will distinguish two main types of contact-induced grammaticalization depending on whether or not there exists already a model source-to-target grammaticalization process to be replicated. If no such model exists we will refer to the process as ordinary (contact-induced) grammaticalization. Replication in this case is confined to creating a category in the replica language that corresponds to the model category—that is, it does not affect the manner in which that category is created (section 3.1.2). If the model language provides a model for both a category and the way that category is replicated, we will refer to it as replica grammaticalization (section 3.1.3).

3.1.1 Grammaticalization theory

In section 1.3 we presented an outline of grammaticalization theory, whose primary goal it is to describe how grammatical forms and constructions arise and develop through space and time, and to explain why they are structured the way they are. To this end, we proposed a catalogue of parameters for describing and understanding the evolution of grammatical use patterns and categories (see section 1.3, (5)):2

(1) Parameters of grammaticalization

a. extension, i.e. the rise of novel grammatical meanings when linguistic expressions are extended to new contexts (context-induced reinterpretation),
b. desemanticization (or "semantic bleaching"), i.e. loss (or generalization) in meaning content,
c. decategorization, i.e. loss in morphosyntactic properties characteristic of lexical or other less grammaticalized forms, and
d. erosion (or "phonetic reduction"), i.e. loss in phonetic substance.

In the present chapter we will be concerned with data on contact-induced language change from a wide range of languages, with a view to establishing that grammatical replication is a cross-linguistically fairly regular process that can be accounted for within the framework of grammaticalization theory.

3.1.2 Ordinary contact-induced grammaticalization

As observed in Heine and Kuteva (2003: 533), contact-induced grammaticalization rests on a strategy used for transferring some grammatical concept or structure from the model language (M) to the replica language (R), involving the following mechanism:

(2) Ordinary contact-induced grammaticalization

a. Speakers notice that in language M there is a grammatical category Mx.
b. They create an equivalent category Rx in language R on the basis of the use patterns available in R.
c. To this end, they draw on universal strategies of grammaticalization, using construction Ry in order to develop Rx.
d. They grammaticalize Ry to Rx.

We saw in chapter 2 that contact-induced grammatical replication involves two distinct though interrelated kinds of structures, which are use patterns on the one hand and grammatical categories on the other. The mechanism sketched in (2) is phrased in terms of the latter, but it applies in much the same way to the former. Furthermore, instead of languages, the mechanism may involve different varieties of one and the same language.

The mechanism raises a number of problems, some of which were discussed in Heine and Kuteva (2003). One problem relates to the nature of categories and the more general issue of cross-linguistic equivalence, which we will return to in section 6.1. Another problem concerns the question of whether, or to what extent, the mechanism is based on conscious or unconscious linguistic activity, and what the ultimate motivations may be for (2b) to happen. The remaining questions (2c) and (2d) can be answered with reference to grammaticalization theory, and there are at least some ways in which they have been, or can be, dealt with (see especially Heine, Claudi & Hünemeyer 1991; Hopper & Traugott 1993; Bybee, Perkins & Pagliuca 1994; Lehmann [1982]1995). But there remains a fundamental problem: grammaticalization is a gradual process that may involve generations of speakers and extend over centuries. This implies that there is likely to be an asymmetrical relation between the two languages, in that there is a category (Mx) in the model language corresponding to a process (Ry > Rx) in the replica language. What this means is that equivalence in grammatical replication is a highly complex notion (see section 6.1).

With the term "create" in (2b) we are drawing attention to the fact that the process concerned is essentially a creative act: Speakers of R do not simply produce a copy of material they find in language M; rather, they tend to create new categories using the resources available in R. As we saw in section 1.5, however, creativity is constrained in a number of ways: first, it is shaped in particular by the nature of the model category Mx, second, it is constrained by universal principles of grammaticalization, and third, it is also influenced by the structural outfit of both the model and the replica languages.
The mechanism sketched in (2) can be illustrated with the following example. Eastern Oceanic languages of northern and central Vanuatu (= M, the model languages) commonly distinguish a durative aspect indicating that an act is in progress (= Mx). Apparently in an attempt to find an equivalent for such a category (= Rx) in Bislama, an English-based pidgin of Vanuatu (= R), speakers used an expression commonly recruited cross-linguistically to develop progressive and durative aspect markers (Heine & Kuteva 2002: 127, 198): They chose a use pattern involving their verb *stap* 'stay, be present, exist' (= Ry) to develop a durative aspect marker (= Rx), which appears in the same syntactic slot as the durative markers (Mx) in the model languages (Keesing 1991: 328; cf. the following sentences, where (3) illustrates the replica and (4) the model structure.

(3) Bislama (English-based pidgin; Keesing 1991: 328)
em i stap pik-im yam.
he he- DUR dig- TRS yam
'He's in the process of digging yams.'

(4) Vetmbao (Malekula, Oceanic; Keesing 1991: 328)
naji ng- u- xoe dram.
he he- DUR- dig yam
'He's in the process of digging yams.'

That speakers of replica languages draw on universal principles of grammaticalization in order to develop a category that is equivalent to the one they find in the model language can be illustrated with another example from Oceanic languages. Kwaio (= M), a Malaita language of Eastern Oceanic, has a grammatical particle *me'e* used before action verbs to express what could be paraphrased in English as 'went ahead and'; Keesing (1988: 217) calls it a narrative discourse marker introducing consecutive, new events (= Mx). Now, one common way this discourse function is expressed cross-linguistically is by using a verb for 'come' or 'go' as an auxiliary (Heine & Kuteva 2002: 68-9), and Kwaio speakers of Solomons Pijin (= R) have grammaticalized the pidgin verb *kam* 'come' (= Ry) to replicate the *me'e*-category, creating a grammatical marker (= Rx), illustrated in (5), that is both morphosyntactically and semantically equivalent to *me'e* (= Mx), cf. (6).

(5) Solomons Pijin (English-based pidgin; Keesing 1988: 217)
olketa-i kam goap long loiakeni.
(they) AUX ascend LOC rattan: vine
'And then they went ahead and climbed that rattan vine.'

(6) Kwaio (Eastern Oceanic; Keesing 1988: 217)
gila me'e fane naa 'ue la'akau.
(they) AUX ascend LOC vine DEM
'And then they went ahead and climbed that rattan vine.'

These examples involve cases where contact-induced grammaticalization introduced a verbal category in the replica language for which previously there does not appear to have been any equivalent. 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 Cémuhé (= 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) (Corne 1995; see Heine & Kuteva 2003: 534 for an example).

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 or -pela* (< English fellow) to a non-singular marker, and the combination *yu 'you* plus *mi 'I* to a first-person inclusive marker (= Ry > Rx). These grammaticalizations have given rise to a system of personal pronouns in Tok Pisin that appears to reflect the corresponding systems of the Austronesian model languages (see, e.g., Keesing 1988: 160-1). Table 3.1 gives a list of personal pronouns of Tigak, an Austronesian language spoken on the island New Ireland of Papua New Guinea, and table 3.2 provides the corresponding Tok Pisin pronouns.

In a number of the languages of the world, it can be observed that universal principles of grammaticalization can be held responsible for the rise of new markers for personal deixis in situations of language contact. A paradigm case can be seen in a process whereby either impersonal pronouns or nouns meaning
Table 3.1 Independent personal pronouns of Tigak

<table>
<thead>
<tr>
<th>Number and person</th>
<th>Singular</th>
<th>Dual</th>
<th>Trial</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 inclusive</td>
<td>nakarag</td>
<td>nakaratul</td>
<td>nakara</td>
<td></td>
</tr>
<tr>
<td>1 exclusive</td>
<td>naniu</td>
<td>namentul</td>
<td>namem</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>naru</td>
<td>namug</td>
<td>nami</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>nane</td>
<td>nareg</td>
<td>nari</td>
<td></td>
</tr>
</tbody>
</table>

Source: Austronesian; Jenkins 2002: 216.

Table 3.2 Tok Pisin personal pronouns

<table>
<thead>
<tr>
<th>Number and person</th>
<th>Singular</th>
<th>Dual</th>
<th>Trial</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 inclusive</td>
<td>yumi(tupela)</td>
<td>yumipela</td>
<td>yumi(pela)</td>
<td></td>
</tr>
<tr>
<td>1 exclusive</td>
<td>mi</td>
<td>mitripela</td>
<td>mipelal</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>yu</td>
<td>yutripela</td>
<td>yupela</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>em</td>
<td>ol(tupela)</td>
<td>ol(tripela)</td>
<td></td>
</tr>
</tbody>
</table>

Source: Jenkins 2002: 216.

'person' or 'people' are grammaticalized to first-person plural pronouns (Heine & Kuteva 2002: 233–4). This process can be contact-induced, as demonstrated, for example, by Aikhenvald (2003: 18) with an example from the North Arawak language Tariana. In an attempt to account with the structures of Tucanoan model languages, Tariana speakers are grammaticalizing their impersonal prefix and pronoun to a first-person inclusive marker ('we, including you').

On the rise of adpositions

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.

The first example concerns Pipil, an Aztecan language of El Salvador which is now nearing extinction, being replaced by Spanish. Traditionally, Pipil (= R), like most other Meso-American languages, has neither prepositions nor postpositions, but it has relational nouns instead (Campbell 1987; Harris & Campbell 1995: 126–7; see also Heine & Kuteva 2003: 535). 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 cross-linguistically common process whereby relational nouns are grammaticalized to adpositions (Heine, Claudi & Himmelmayr 1991). As table 3.3 suggests, the process shows all the effects of the three main mechanisms involved in grammaticalization (see section 1.3): 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)h)pak > pak 'on').

Pipil is not an isolated example of a Central American language having redefined its structure of adpositions as a result of contact with Spanish. Another example is provided by Mexican as spoken in the Malinche Volcano towns of Central Mexico: Malinche Mexicanos speakers are nearly all bilingual in Mexican and Spanish, and their language has been shaped by Spanish to the extent that Hill and Hill (1986: 1) call it a “syncretic language.” These authors observe that Classical Mexican (Aztec, Nahua) had postpositions marking locatives, although the repertoire of true postpositional elements was somewhat limited (Hill & Hill 1986: 247–8). What happened under Spanish influence is, first, that Malinche Mexicanos speakers borrowed the Spanish possessive/genitive particle de, second, that they replicated the prepositional structure of Spanish by turning what appear to be relational nouns having a possessive prefix coreferential with the possessor into preposition-like entities and, third, that they replaced that possessive prefix with the indefinite element la-

What distinguishes the Malinche Mexicanos example from the Pipil one is most conspicuously that, in addition to replication, it also involved borrowing. Furthermore, contact-induced change appears to have been one from minor to major use pattern (see section 2.2), rather than from one categorial structure to another; there was already a preposition-like use pattern in Classical Mexican, e.g. f-pan in cama (its-on in bed) 'on the bed,' which under Spanish influence turned into what Hill and Hill (1986: 248) call the absolutely “preferred way to render prepositional relationships in Malinche usage.” And whereas in Pipil the possessive prefix was eliminated on the way from relational...
nound to preposition, it was replaced by an indefinite marker in Malinche Mexican, e.g. *ta-čuulapan den čarcet (ta=behind de. in prison) 'behind the prison.' Finally, Malinche Mexican and Pipi also appear to differ from another in the fact that the old postpositions of Classical Mexican, while having 'largely given way to a more Hispanicized way of speaking,' have not disappeared in modern Malinche Mexican, even if they are rarely used.

That the same kind of contact-induced grammaticalization process can occur anywhere is suggested by the following example from Papua New Guinea. Ross (1996; 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. The model language (= M) is Waskia, a Papuan language of the Trans New Guinea type, and the replica language Takia (= R), a Western Oceanic language of the Bel family of the North New Guinea cluster. 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, or no longer use them productively. In an attempt to replicate the postpositions of Waskia (= Mx), Takia speakers developed postpositions (= Rx) by grammaticalizing inalienably possessed relational nouns (Ry). In this way, a Proto-Western Oceanic construction illustrated in (7) turned into a postpositional construction (8) in Takia (note that Takia has given up the possessee-possessor order of Western Oceanic and adopted the possessor-possessee order of Waskia). In accordance with the parameters of grammaticalization, the prepositional phrase *i lalo-ña lost its nominal structure, turned into an adposition (decategorialization), and was phonetically reduced to lo (erosion). Thus, speakers of the replica language Takia took recourse to a grammaticalization process; using relational nouns for 'inside' and developing them into locative (inessive) adpositions, for example, is a cross-linguistically widespread strategy (Heine & Kuteva 2002; see also Heine & Kuteva 2003: 536).

(7) Proto-Western Oceanic (Ross 1996: 189; 2001: 143)
*i lalo-ña a Rumaq
PREP inside- its ART house
'inside the house'

(8) Takia (Western Oceanic; Ross 1996: 190; Ross 2001: 143)
ab lo
house in
'in the house'

Another example illustrating the same grammaticalization process in Takia suggests that this is not an isolated case.

(9) Proto-Western Oceanic (Ross 1996: 189)
*i papo-ña a Rumaq
PREP top- its ART house
'on the top of the house'

(10) Takia (Western Oceanic; Ross 1996: 190)
ab [fulfo
house on
'on top of the house'

Once again, a prepositional phrase with a possessive noun phrase as its modifier, cf. (9), turned into a postposition, as can be seen in (10), involving the decategorialization of a nominal construction and the erosion of a complex form to a postposition fulfo or fo.

The data provided by Ross (1996: 188–90) make it possible to reconstruct this grammaticalization process in more detail. First, they show that, predictably, the process did not lead straight, e.g. from a prepositional phrase *i lalo-ña to a monosyllabic postposition lo, but rather involved a number of intermediate stages. One such stage is reflected in the Takia form i-lo-n, where the erstwhile preposition i- is fossilized. In Gedaged, another Bel language closely related to Takia, the form has been further eroded, the postposition for 'in' being lon. The final-stage form lo is not only found in Takia but also in the other western Bel languages. To summarize, the development from prepositional phrase to postposition in the Bel languages appears to have involved a series of decategorializations and erosions of the following kind:

(11) Takia and other Bel languages (Ross 1996: 189–90)

From prepositional phrase to postposition
*i lalo-ña > i-lo-n > lon > lo

Note that two of these stages of grammaticalization have been retained in Takia, and the two can cooccur in one and the same phrase: when a Takia speaker wishes to be precise about the location, s/he can say the following:

(12) Takia (Western Oceanic; Ross 1996: 190)
ab ilo- n lo
house inside- its in
'in the house’s inside'
Second, these data also allow us to understand why Takia underwent a change from preposition to postposition.10 One cross-linguistically fairly widespread conceptual schema leading to the rise of new possessive constructions is referred to by Heine (1997a: 144) as the Topic Schema, taking the form [(As for) X, X’s Y] (e.g. something like John, his car), where the possessee follows the possessor and agrees with the latter in the form of a possessive attribute. Predictably, in languages using this schema the possessee follows, rather than precedes, the possessor — irrespective of any word-order constraints that may characterize the language concerned.11 It would seem that Takia speakers have drawn on this schema, with the effect that the possessee follows the possessor. Now, when the possessee phrase *i lalo-ha was grammaticalized, the expected result was a postposition on the erstwhile possessor rather than a preposition. This suggests that Takia’s history from prepositions to postpositions did not involve any word-order change; rather, the prepositions were lost and the postpositions arose via the creation of a new possessive construction. While this was seemingly a language-internal process, we concur with Ross that it was at the same time contact-induced, in that the choice of the Topic Schema provided Takia speakers (or their ancestors) with a strategy to match the postpositional structure of the model language Waskia.

That this hypothesis is correct is also suggested by the following. It would seem that this process is not an instance of ordinary grammaticalization but rather of replica grammaticalization (see below), in that there is evidence to the effect that Takia speakers replicated a grammaticalization process that they observed in Waskia. Ross reconstructs the corresponding process in Waskia thus:

In a Waskia phrase like kar kuali ‘on the fence’, kuali ‘on’ is derived from *k<k>al + i, where the infix <k> marks a third person possessor (corresponding to PWOc *-ña, Takia -n), -i is the location postposition (corresponding functionally to the PWOc preposition *-i — the formal correspondence is chance), and *kal was a now-lost part noun meaning ‘top’ (corresponding to PWOc *papo).12 (Ross 1996: 190)

Accordingly, a Waskia phrase such as (13b) is hypothesized to be historically derived from a structure as sketched in (13a).

(13) Waskia (Trans Guinea Type; Ross 1996: 190)

a. *X — k<k>al-i
   **X on its top
b. X — kuali
   ‘on top of X’

This analysis suggests that Waskia speakers also used the Topic Schema or, more precisely, Takia speakers appear to have replicated the grammaticalization from Topic Schema to postposition on the model of Waskia.

To conclude, 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.

Extension

Ideally, grammaticalization — whether contact-induced or not — involves all of the four parameters listed in (1). As a matter of fact, however, this is not always so; in quite a number of cases it is confined to desemanticization. But there are also cases where it appears to be restricted to one of the other parameters. We will now illustrate the effect of the other three parameters.

A number of examples were provided in section 2.2.2 exhibiting the effects of extension. That extension tends to go hand in hand with desemanticization is suggested by cases such as the following. In Standard German, one of the main ways in which purpose clauses are formed is by means of the complex conjunction um zu ‘in order to, so that.’ This conjunction is not normally used to introduce nominal or adjectival modifiers; still, there is a widely accepted use pattern where um zu can be used, e.g. ein Messer um Brot zu schneiden (a knife for bread to cut) ‘a knife for slicing bread.’ Among German speakers of South Tyrol (Südtirol) in northern Italy, Riehl (2001: 255) found that, under the influence of Italian, this use pattern has been extended to contexts where a purpose interpretation no longer makes sense, and where a new function appears to be emerging, namely that of presenting infinitival phrases as modifiers of nouns and adjectives (see section 2.2.3). Thus, the Tyrolean German expression Schwierigkeiten, um Freundschaften zu knüpfen (difficulties to form friendships) ‘difficulties to form friendships,’ which is not well-formed in Standard German, is modelled on Italian difficoltà a farsi delle amicizie. This process is in accordance with a cross-linguistically widespread grammaticalization of markers for purpose clauses to infinitival markers (see Haspelmath 1989). In this way, the German marker um zu was extended to new contexts, at the same time acquiring a more general function, namely presenting both purposive and infinitival participants.

To conclude, context extension is one of the most salient forces characterizing contact-induced grammaticalization, and it is inextricably associated with semantic generalization. But there remains a problem that is still unresolved: there is no conclusive evidence to determine which of the two forces,
extension and desemanticization, provides the primary motivation for grammatical change.

Decategorialization

The following example suggests that contact-induced grammaticalization can be confined to decategorialization. Contact between Bantu and Nilotic languages in Kenya resulted in a number of grammatical changes, which have been described in detail by Dimmendaal (2001b). Our example concerns another case where Bantu languages appear to have acted as model languages. Bantu languages are known for their rich paradigms of verbal derivational extensions marked by suffixes. There is nothing comparable in the Nilotic language Luo or its closest relatives, the Southern Lwoo languages of Uganda and the Sudan: verbal derivation is limited, mainly involving internal morphology in the verb root. Now, apparently on the model of neighboring Bantu languages (M), Luo (= R) speakers have developed a set of what look like verbal suffixes (= Rx), resembling structurally the Bantu verbal suffixes (Mx), expressing functions typically encoded by the Bantu derivational applied suffix */l-l/ (‘for, to, with reference to, on behalf of’). This typological innovation has been described by Heine and Reh (1984: 50–3) as “verbal attraction,” a grammaticalization process whereby adpositions lose their independent status and turn into verbal derivational clitics and affixes. Luo speakers used the prepositions ne (or nl) benefactive, e locative, and gi instrumental (= Ry) in order to develop verbal enclitics or suffixes (= Rx); the following example is confined to the benefactive preposition ne, where (14a) illustrates the prepositional use and (14b), where Juma is topicalized, the use as a verbal suffix (see also Dimmendaal 2001b: 101–2).

(14) Luo (Western Nilotic, Nilo-Saharan; Heine & Reh 1984: 51)

a. joni nego diel ne juma.
   John is.killing goat for Juma
   ‘John is killing a goat for Juma.’

b. juma joni nego- ne diel.
   Juma John is.killing- for goat
   ‘John is killing a goat for Juma.’

These examples might 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 morphosyntactic form nor in their meaning to the model, and in order to achieve equivalence they have to go through a complex process whereby lexical structures are developed into grammatical markers, adpositional phrases are gradually transformed into adpositions, free forms are grammaticalized to clitics and affixes, etc.

Erosion

That contact-induced grammaticalization can be restricted to erosion is suggested by the following example. The Austronesian language Tigak of New Ireland, Papua New Guinea, has replicated a number of structures from Tok Pisin, spoken by the Tigak as their primary L2 (Jenkins 2002: 243–55). Tigak speakers have developed the phrase lo tang gaan ‘at the time’ into a temporal subordinating conjunction ‘when.’ Whereas traditional, conservative Tigak speakers use this full phrase, as in (15a), younger Tigak have reduced the phrase to the noun gaan ‘time,’ cf. (15b). As Jenkins (2002: 254–5) argues, this reduction process is suggestive of contact-induced influence by Tok Pisin, which uses the noun taim ‘time’ as an equivalent subordinating conjunction, as can be seen in (16).

(15) Tigak (Austronesian; Jenkins 2002: 254–5)

a. lo tang gaan tang vuul gi me sang vo
   LOC ART time ART canoe 3.SG.S MT arrive IRR
   nag kos.
   1.SG.S board
   ‘When the canoe arrives I will go on.’ (Lit.: ‘at the time . . .’)

b. gaan tang vuul gi me sang vo nag kos.
   time ART canoe 3.SG.S MT arrive IRR 1.SG.S board
   ‘When the canoe arrives I will go on.’

(16) Tok Pisin (English-based pidgin/creole; Jenkins 2002: 255)

taim mun i kam bai mi kis-im.
   time canoe PM come FUT 1.SG get- TRS
   ‘When the canoe arrives I will go on.’

When an adpositional phrase, like Tigak lo tang gaan, is grammaticalized to an adpositional or clause subordinator this is likely to trigger erosion, whereby that phrase is phonologically and/or morphologically reduced (cf. English by cause of > because (> coz)). This process need not, and most often is not, induced by language contact. In the case of Tigak, however, it seems plausible that reduction was triggered or accelerated by the model provided by Tok Pisin. Note that this case involves morphological rather than phonetic erosion (see
Heine & Reh 1984), in that reduction led to a loss of morphological items rather than of phonetic substance of morphemes.

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 provides not only a model category (Mx) but also a way of how to replicate that category (My).

3.1.3 Replica grammaticalization

As was observed in Heine and Kutova (2003: 539), there is a second type of process that appears to be even more common than the one described in section 3.1.2. In this process, it is not a grammatical concept but rather a grammaticalization process that is transferred from the model (M) to the replica language (R) (see also Nau 1995: 96 for an insightful discussion). The mechanism underlying this process, which they call replica grammaticalization, differs from the one sketched in (2) only in the fact that (17c) replaces (2c).

(17) Replica grammaticalization
a. Speakers notice that in language M there is a grammatical category Mx.  
b. They create an equivalent category Rx in language R, using material available in 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 Ry to Rx.

This mechanism corresponds to some extent to what Matras proposes to call mutual isomorphism in his discussion of contact-induced changes in Balkan Turkish and Romani.

Mutual isomorphism taken in the metaphorical sense suggested here is the tendency of languages in an areal to syncretize their operational procedures in such a way that would enable a corresponding structure to initiate a corresponding processing operation. [. . .] Individual languages select suitable candidates among their inherited stock of grammatical items, to which the necessary functions can be assigned most naturally. As a result, language internal forms: function correlation is extended, leading ultimately to functional scope-enhancement. (Matras 1998b: 100–1)

There is virtually no information on what conceptual clues speakers may have to reconstruct a process presumed to have taken place in the model language, and it is not always possible on the basis of the evidence available to distinguish neatly between ordinary and replica grammaticalization. Nevertheless, as we will see below, there are frequently clues that make it possible to discriminate between the two (see, e.g., our discussion in 3.3 on the future tense in varieties of Pennsylvania German and Yiddish).

The mechanism sketched in (17) is different from the one discussed in section 3.1.2, as can be illustrated with an example reported by Weinreich (1953: 1953: 40). In German, the third-person plural pronoun sie ‘they’ is formally the same as the polite/formal second-person singular pronoun Sie ‘you,’ and the same situation is found in the local Polish dialect of Silesia, a contact region between German and Polish speakers. This example appears to be a typical case of (17): the model language German (= M) has undergone a grammaticalization process whereby the third-person plural pronoun (= My) was grammaticalized to a second-person singular pronoun to be used for polite/formal reference (= Mx). Polish (= R) speakers in Silesia replicated this process by extending the use of their third-person plural pronoun (= R = Ry) to a new function (Rx). 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. Obviously, replication did not mean that the Polish speakers repeated the history of the German Sie-construction; however, replication was not confined to simply copying a polysemy pattern (see section 3.2) that they found in the model language but rather involved a process that was structurally not unlike the one speakers of the model language had undergone centuries earlier.

Verbs expressing predicative possession (‘have’) are cross-linguistically a not uncommon source for existential copulas, where an expression of the form ‘it there has’ is grammaticalized to an existential copula (‘there is’; for details, see Heine & Kutova 2002: 241–2), cf. French il y a (lit. ‘it there has’ > ‘there is’). Apparently on the basis of this pathway of grammaticalization, speakers of the colloquial variety of Singaporean English (= R) are using possession verbs (= Ry) to denote existence (= Rx). In doing so, they rely on the model provided by their L1 Chinese (= M), thereby replicating the grammaticalization process from possession verb to existential marker (= My > Mx) (Ho & Platt 1993: 18; Ziegeler 2000: 90).

It is fairly easy to discover cases of replica grammaticalization when the model language has developed a grammatical category by using a conceptual source that is rarely encountered cross-linguistically and where exactly the same source is used by speakers of the replica language. The Irish “hot-news” perfect (= Mx) is based on what appears to be an instance of a spatial or temporal schema of the kind [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). Presumably around the late seventeenth century (Sullivan 1980: 205),
But essentially the same process has also occurred in the Indo-Aryan languages, for example. Konkani (1973), in fact, demonstrates that Konkani has a relatively high level of grammaticalization, similar to English, in terms of the degree of phonological change in the language. The effects of this are seen particularly in the area of the morphological system. In Konkani, the use of the possessive determiner has been replaced by a relative pronoun, which is followed by a nominal phrase. This has led to the development of a relative clause system that is similar to that found in English. Konkani has also undergone a number of other changes, such as the development of a new word order, the loss of certain case markers, and the restructuring of the nominal and verbal systems. All of these changes have contributed to the grammaticalization of the language, and have led to a greater degree of phonological and morphological change.

The mechanism of grammaticalization that has been observed in Konkani is also relevant to English. The process of grammaticalization involves the development of new grammatical structures from non-grammatical elements. In Konkani, this process has led to the development of a relative clause system, which is similar to that found in English. The development of this system has been due to the loss of certain case markers, and the restructuring of the nominal and verbal systems. All of these changes have contributed to the grammaticalization of the language, and have led to a greater degree of phonological and morphological change.

Most of the above examples concern cases where a new grammatical category has been developed for which there existed no equivalent in the language before. However, there are also cases where a new category is developed for which there already existed an equivalent category. In such cases, the new category may be grammaticalized in a different way than the old category. For example, in English, the use of the possessive determiner has been replaced by a relative pronoun. This has led to the development of a relative clause system that is similar to that found in Konkani. However, the development of this system has been due to the loss of certain case markers, and the restructuring of the nominal and verbal systems. All of these changes have contributed to the grammaticalization of the language, and have led to a greater degree of phonological and morphological change.
To conclude, replica grammaticalization has resulted in total structural isomorphism (see section 6.1) in the expression of this construction between two genetically unrelated languages; neither any other dialect of Konkani nor any of the other Indo-Aryan languages of India has such a construction. As Nadkarni (1975: 677, 679) suggests, there was no structural motivation for this replication. Konkani already had a well-functioning relative construction of the Indo-Aryan type, described as a phrasal relative structure using a relative participle instead of a finite verb. This old category, which is also found in Kannada, differs from the replicated category in being more versatile than the new one, in that it allows for extraposition while the new category is restricted to occurrence before the main clause, nor did the new category add any desirable stylistic variant to the already existing situation. Nevertheless, there are indications that the new category is expanding at the expense of the old one—a process that Nadkarni refers to as the Dravidianization of this Konkani dialect. For another effect this contact situation had, see section 4.3.3.4.

From adverb to clause subordinator

A not uncommon grammaticalization concerns the extension of adverbial modifiers to introduce subordinate clauses; for example, locative adverbs may turn into subordinators of locative, temporal, causal, or other adverbial clauses (see Heine & Kuteva 2002). Such a process can also be observed in contact-induced replication. Jenkins (2002: 269) observes that the Austroasiatic language Tigak of New Ireland, Papua New Guinea, and the English-based Tok Pisin have a form that functions both as a restrictive modifier (‘only/just’) and as an adversative (‘contrastive’) conjunction (‘but’), the forms being kisang in Tigak and tasol (< English that’s all) in Tok Pisin. Jenkins argues that this functional equivalence is due to substrate influence of Austroasiatic languages on Tok Pisin; we interpret this as a case of replication where Austroasiatic languages such as Tigak provided the model that was replicated in Tok Pisin. It would seem that underlying this equivalence relation there appears to be a grammaticalization process, so far undocumented, whereby the use of restrictive adverbs (or nominal modifiers) is extended to serve as adversative clause markers. For example, the German particle nur ‘only/just’ serves as a restrictive adverb in (19a) but as an adversative conjunction ‘but’ in (19b):

(19) German
a. Er arbeitet nur.  
   ‘He works only’

In a similar fashion, Tok Pisin tasol occurs as a clause-final adverb in (20a) but as a clause-initial subordinator in (20b)—exactly like Tigak kisang does, cf. (21).

(20) Tok-Pisin (English-based pidgin/creole; Jenkins 2002: 269)

   a. ol i pilai tasol.  
      3.PL PM play only  
      ‘They’re only playing.’

   b. em i gat mani tasol em i no givim pe  
      3.SG PM have money but 3.SG PM NEG give pay long mi.  
      to 1.SG  
      ‘He has money, but he did not pay me.’

(21) Tigak (Austronesian; Jenkins 2002: 269)

   a. nari rig karau kisang.  
      3.PL 3.PL.S.AGR play only  
      ‘They’re only playing.’

   b. name gi togi ta mani kisang  
      3.SG.PRN 3.SG.S.AGR have ART money but  
      gi veko lisani ta pulup su-gug.  
      3.SG.S.AGR NEG give ART pay to 1.SG  
      ‘He has money, but he did not pay me.’

Since Tigak provides both the source and the target uses of the relevant marker, there is reason to assume that this is an instance of replica grammaticalization.

Creating text markers

Contact-induced grammaticalization manifests itself perhaps most frequently in text structure. Unfortunately, this area has not yet been studied in great detail; still, there are a few findings that suggest that the way texts, in particular narrative texts, are organized is determined to some extent by grammaticalization. Paradigm cases concern markers of boundaries, in particular the beginning and the end of a text, significant units within the text such as paragraphs and topic change, but also of continuity of narrative discourse. There is a not uncommon pattern whereby transparent expressions such as clausal propositions are grammaticalized to markers of text organization. As a result of this process, they tend
to lose their erstwhile semantic content in favor of some discourse function, to be decategorialized to unanalyzable markers, and to undergo erosion, that is, to be reduced in their morphophonological shape. Almost invariably, the process involved is replica grammaticalization.

Further examples of replica grammaticalization serving the creation of boundary markers in spoken discourse come from the contact situation in the Vaupés area of northwest Amazonia, as a result of which the North Arawak language Tariana has replicated a wide range of grammatical structures from East Tucanoan languages in general and Tucano in particular (Aikhenvald 2002). Among these structures there is a recapitulating clause meaning ‘so s/he did’ (where ‘s/he’ refers to the subject of the previous clause). This clause, which is *di-ni* ‘he did’ in Tariana and *wee* ‘(the) did’ in Tucano, appears to have been grammaticalized to a discourse marker introducing new paragraphs. Another example of a boundary marker stemming from this contact situation concerns a marker signaling the end of a speech held in a centre house at gatherings such as religious or political meetings. In Tariana there is an expression *matia* ‘good, OK’ which corresponds to Tucano *ayi* ‘good, OK.’ In both languages, this expression has given rise to a marker signaling the end of a speech held in such meetings (Aikhenvald 2002: 169–71).

For another discourse marker, replicated in the English-based pidgin Bislama of Vanuatu on the model of Oceanic languages, see section 3.2.1 below.

**From optional to obligatory category**

One salient characteristic of grammaticalization concerns what Lehmann (1982) 1995 describes as obligatorification, whereby categories whose use is optional come to be used more frequently and may turn into obligatory parts of words or word groups. Not infrequently, this process concerns pragmatically determined categories gradually losing their pragmatic functions and acquiring the properties of automatic syntactic constituents.

Language contact provides a not uncommon trigger for obligatorification. We mentioned two cases in the preceding chapter, both involving independent personal pronouns which gradually develop into pragmatically non-required (subject and object) pronouns on the model of another language. One example concerned language contact in Sauris, northeastern Italy, between speakers of German and the Rhaeto-Romance language Friulian (section 2.1); the second example dealt with immigrant communities in the USA speaking pro-drop languages such as Spanish, Serbian, Russian, or Hungarian, who tend to redefine their pro-drop convention on the model of English (section 2.3.4).

The following observations suggest that this case of pro-drop “resetting” or “decline,” as it has been called, constitutes a canonical process of grammaticalization. As we pointed out above (section 3.1.1), grammaticalization involves four basic mechanisms, and three of these mechanisms can be said to be at work when a language changes from a pro-drop to a non-pro-drop structure. One of these mechanisms is desemanticization, whereby linguistic expressions lose in semantic content. In our case, loss concerns the fact that pronouns that are said to be used for pragmatic purposes such as presenting new information, marking topic or theme, or expressing “emphasis” lose their pragmatic meaning. Context generalization can be seen in the fact that in pro-drop languages the use of independent personal pronouns is severely restricted, being determined by pragmatic variables. As a result of language contact, the use of these pronouns is generalized, being extended to contexts where the pronouns were not normally used previously. Finally, decategorialization has the effect that the personal pronouns lose their categorial status as independent constituents, turning into largely predictable subject (or object) markers.

To conclude, “pro-drop resetting” among these immigrant communities in the USA has the characteristics of a process leading from minor to major use pattern (see section 2.3.4), and from optional to obligatorily marked category. That one of the mechanisms associated with grammaticalization, namely erosion (or phonetic reduction), does not appear to have been involved is probably due to the following: Erosion is usually, even if not always, the last process to come in grammaticalization, and it may take quite some time before its effects can be observed. All the cases discussed here concern periods of contact that were simply too short to have had a distinct effect on the phonetic substance of the pronouns undergoing grammaticalization.

**Extending lexical use patterns**

Replica grammaticalization is also involved in a number of cases which tend to be described as instances of lexical polysemy but which, nevertheless, appear to involve a development from lexical to grammatical uses of linguistic items.

The following example concerns the verb *(la)schar* of the Rhaeto-Romance language Surselvian (see section 4.1.2): The German verb *lassen* ‘let’ has a number of grammatical uses; in particular, it is used as a causative auxiliary ‘(cause to do)’ and as a permissive auxiliary ‘(allow someone to do)’. Romance languages express these two functions by using different verbs, e.g. French *faire* vs. *laisser*, Italian *fare* vs. *lasciare*, Spanish *hacer* vs. *dejar*. Similarly, the Rhaeto-Romance languages traditionally make such a distinction: *far* vs. *(la)schar*. Under German influence, however, Stümm (1984) argues, Surselvian tends to generalize *(la)schar* to express both a causative and a permissive meaning—with the effect that there now appear to be two alternative ways of expressing causativity. In this way, the Surselvian verb has undergone replica
grammaticalization on the model of the German development from lexical verb to causative auxiliary.

A similar example can be seen in the Swedish verb ha 'have': Some Swedish long-term immigrants to the USA who left Sweden in the 1920s were found in the 1980s using this Swedish verb in ways parallel to English have, which has developed a grammatical use pattern as a causative auxiliary. This pattern was replicated by the Swedish immigrants, grammaticalizing ha to a causative use pattern on the model of English have in contexts where in Standard Swedish the verb få 'cause' would be required (see Myers-Scotton 2002: 104).

3.2 Polysemy copying

We noted earlier in this chapter that 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, or of calquing or loan translation. In fact, there are cases of grammatical transfer where an analysis in terms of polysemy copying (or grammatical calquing) provides the most plausible hypothesis.18

Such an analysis offers itself, for example, in cases where, rather than the meaning of the basic form, it is that of a derived form that is replicated, as in the following example, where Oceanic languages serve as models and Bislama, the English-based pidgin of Vanuatu, as the replica language (Keesing 1991). Many northern and central Vanuatu languages (as well as other Oceanic languages; = M) have grammaticalized a reduplicated form of the verb for 'go' to a marker indicating the passage of time in a narrative (= Mx). Thus, in Epi, the reduplicated form bababa of the verb of motion ba 'go' expresses the passage of time in a narrative. This situation appears to have been replicated in Bislama (= R), where the corresponding verb of motion go (< English go) was reduplicated as go-go-go and used as a discourse marker also expressing the passage of time in narratives (= Rx).19

This is not an isolated example, as suggested by the following case, which also involves reduplication: Serbian/Croatian has developed a productive pattern of forming absolute superlatives on the model of Turkish, e.g. beli 'truly' > bezbeli 'quite truly,' ravnio 'even' > ravrnivo 'entirely horizontal' (Weinreich 1953) 1964: 42). As in the preceding example, replication in this case appears to have concerned the semantic relation between the simple and the reduplicated form, rather than a gradual process from less to more grammatical forms.

While such cases suggest that polysemy copying constitutes an alternative to grammaticalization for transferring grammatical categories from the model language to the replica language, most cases of grammatical replication that have been reported cannot be described appropriately in terms of notions such as polysemy copying, calquing, or loan translation. The reader is referred to Heine and Kuteva (2003: 555–9) for detailed discussion; we are confined here to the main points substantiating this position.

One point concerns the observation that most cases that are seemingly instances of polysemy copying can also be interpreted as grammaticalization processes. First, they are unidirectional, leading from less to more grammatical use patterns, and we are not aware of any case of reversed directionality. For example, a number of languages that have been in contact with Standard Average European languages such as French and German have grammaticalized question words to markers of relative and complement clauses; but we are not aware of any example suggesting that any language in Europe, or elsewhere, has undergone a development from relative or complement clause marker to question word.

Second, wherever there is sufficient evidence, it turns out that the replica construction is less grammaticalized than the corresponding model construction (see section 3.4.5 below). 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, both the definite article in Sorbian, replicated on the model of German (see section 2.4.1), 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, for example, 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 (Heine & Kuteva 2003: 556–7).

Another example is provided by the evolution of the possessive perfect (or 'have'-perfect) [X has done Y] based on the grammaticalization of a possessive schema [X has Y] in Romance, Germanic, and some other European languages, that is, in what Haspelmath (1998; 2001) calls Standard Average European. Some languages spoken at the periphery of the Standard Average European area have replicated this construction; almost invariably, however, the replica constructions have not reached the same advanced stage of grammaticalization characterizing that of Standard Average European languages. Slavic languages such as Sorbian, Czech, and Slovenian, which have a long history of contact with German, are cases in point; Breu summarizes the situation in these languages thus:
Einige slavische Sprachen an der Peripherie haben unter Kontaktinfuß ein analytisches HABEN-Perfekt entwickelt, das allerdings in den Asprävaisionen nicht den grammatischen Status der Gebiersprachen erreicht hat. Der Aufbau hat aber durch die Nachbildung der präphrasischen Konstruktion begonnen. [Some Slavic languages at the periphery have developed an analytic 'have'-perfect under contact influence, but in these adstrate situations it has not reached the grammatical status it has in the donor languages. Nevertheless, the construction has been initiated via the imitation of the präphrasic construction.] (Breu 1996: 31)

Third, what distinguishes polysemy copying from canonical instances of grammaticalization is that the former does not appear to involve intermediate stages of evolution. Frequently it is not possible to reconstruct the grammaticalization process in every detail since the data available are for the most part not sufficient to do so. But in a few cases it is possible, and such cases suggest that there is in fact an intermediate stage between the initial and the final stage of grammaticalization. Several examples are presented in Heine and Kuteva (2003: 555–9); the Irish “hot-news” perfect provides an additional example. As we mentioned in section 3.1.3, this category is based on what appears to be an instance of a spatial or temporal schema [X is after Y], where the aspeclul nature of a perfect is encoded by means of a locative (or temporal) preposition ‘after’ (Filippula 1986; Boretzky 1986: 25; Harris 1991: 201ff.). The same grammaticalization process, modeled on the Irish perfect, can be assumed to have taken place in Irish English presumably around the late seventeenth century (Sullivan 1980: 205), cf. (22).

(22) Irish English (Harris 1991: 205)

She’s after selling the boat.
‘She has just sold the boat.’

That replication did not involve polysemy copying is suggested by the following observation. When the Locaction Schema is grammaticalized to a construction for verbal aspect, as illustrated in (22), this involves an intermediate step where the schema has a nominal complement [X is after NP], before the construction is extended to non-finite verbal complements [X is after VP]. Such an intermediate stage appears to exist in the present case of Irish English, where instead of a verbal complement the preposition after has a nominal complement, cf. (23).

(23) Irish English (Sullivan 1980: 205)

He’s after the flu.
‘He just had the flu.’

And finally, replicated categories tend to exhibit properties that bear witness to their grammaticalization history, such as decategorialization (loss of morphosyntactic properties) and erosion (phonic reduction or simplification; see section 3.1.1 (1)).

To conclude, while the data available in the relevant literature are in most cases too scanty to reconstruct the exact process of replication, there appears to be sufficient evidence to suggest that conceptual transfer of grammatical meanings tends to require grammaticalization; for similar observations, see Nau (1995: 114–22). This is most obvious in the case of ordinary contact-induced grammaticalization (section 3.1.2), where 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).

3.3 Future tenses

In the preceding sections we have been dealing with a range of different grammatical meanings. We will now try to explore how one particular grammatical meaning may be affected by language contact; to this end, we will be restricted in this section to a cross-linguistic survey of future tense categories (for some examples, see Heine & Kuteva 2003: 551–5). The reason for dealing with this grammatical meaning is that among all grammatical categories it is future tense that appears to be the most likely to be replicated in situations of language contact – hence, it provides quite some information on how contact-induced grammatical categories arise.

The main source for the grammaticalization 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, Pagliuca & Perkins 1991; 1994). We will refer to these schemas, respectively, as the de-ventive, de-allative, and de-volitive schemas or constructions (cf. Dahl 2000b: 319ff.; Heine & Kuteva 2003: 552). Other sources for future tenses are cross-linguistically clearly less common.

Europe

All these main sources have been recruited in European languages and are each suggestive of areal spread across neighboring European languages (Dahl 2000b). De-ventive 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)
developed a de-allative future on the model of English *be going to* (Boretzky 1989: 368), cf. (24).

(24) Romani of Wales (Boretzky 1989: 369)

   brisindo džala te del.
   (rain goes to give)
   'It is going to rain.'

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 cross-linguistically quite unusual. Now, Ukrainian has two future tenses, and one of them also uses the verb ‘take’ as a future auxiliary. 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 also used to refer to events in the future. This is exactly the situation found in colloquial German, where the present tense is also used in appropriate contexts for future events; conceivably, what this suggests is that absence of a salient model prevented grammaticalization from taking place.

That de-volitive futures found in Balkan languages are in fact due to language contact is suggested, for example, by the fact that the Romani varieties spoken in the Balkans have developed a future category marked with *kam(m)*, which is derived from the verb *kam-av* ‘want’, love (Boretzky 1989: 368). Still, in view of the cross-linguistically widespread distribution of de-volitive futures, a development based exclusively on universal principles of grammaticalization, unaffected by language contact, can never be entirely ruled out. There is, however, additional evidence to suggest that this grammaticalization was contact-induced: while the various Romani varieties developed a number of different future tenses, it is only the varieties spoken in the Balkans that acquired a de-volitive future (Boretzky & Igla 1999: 729); note that the presence of a de-volitive future is widely assumed to be a salient property of the Balkan linguistic area (see section 6.2.3.1).

But the Balkan varieties of Romani provide further evidence that the rise of new future tenses has been shaped primarily by contact. We observed above that there is a less common source for future tense markers based on the grammaticalization of obligation markers typically involving a possession schema, i.e. a possessive verb ‘have’ plus a non-finite main verb – French and other Romance languages providing paradigm examples of this pathway of evolution. The Geg dialect of Albanian and varieties of Macedonian and Bulgarian have drawn on this pathway to develop future tenses, and Romani

---

**Romani**

The role played by areal pressure on shaping equivalence relations and grammaticalization can be demonstrated perhaps most clearly by using the Indo-Aryan language Romani 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 (see above); elsewhere in Europe they found other models and, accordingly, developed different kinds of future categories. In the dialect of Wales they


A case of a de-allative future has been reported from Luxembourgeois, where speakers of this German dialect are said to have replicated the French de-allative construction *aller faire* (‘go to’) by developing their verb *goen* ‘go’ into a future auxiliary (Alanne 1972).

Contact-induced futures derived from modal use patterns are reported from Molisean, a variety of the Slavic Croatian minority in Molise, southern Italy (Breu 1996: 26–7; Heine & Kuteva 2003: 552–3) and from early written Finnish (appr. 1540–1820), where Bible translators are said to have developed weakly grammaticalized futures (based on the modal verbs *pitää* ‘should, must’ and *tahdoa* ‘want, intend’) using German and Swedish modals as models. These future tense use patterns were discontinued around or after 1820 (Nau 1995: 99–104).

An instance of a de-volitive future is found, for example, in the Balkan languages – in fact it is among the uncontroversial morphological features that have been adduced to define the Balkan sprachbund (see 5.2.3.1). 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, Romanian, Macedonian, Bulgarian, Serbo-Croatian, and Romani” (Joseph 1992: 154). That the relevant markers have a “reduced” and/or “invariant” form is a predictable result of grammaticalization: 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). As is common with sprachbund situations, there is no reliable evidence on which Balkan language provided the ultimate model for grammaticalization but, again, there can be hardly any doubt that language contact played some role in the spread of the process.
speakers in contact with these Balkan languages appear to have replicated this pathway. Now, the Romani equivalent of a ‘have’-verb is a construction literally meaning ‘is me’ (si man) to designate ‘I have’; accordingly, Romani speakers replicated the possession schema to create a future tense, illustrated in (25).

(25) Romani, Prizren dialect (Boretzky & Igla 1999: 719)

Hi ma te džav lesa.

is me that go.1.SG he.INSTR

‘I will go with him.’

This kind of replication of grammaticalization processes is perhaps most pronounced in Bulgarian varieties of Romani. Bulgarian has a system of future tense marking where the positive future has a de-volitive source (using a verb meaning ‘want’) and the negative future using the possession schema (using a verb meaning ‘have’). Exactly this split structure of future marking was replicated by Romani speakers influenced by Bulgarian, as we saw in section 1.3.

What this situation suggests is that a future tense exists in Romani essentially only to the extent that it is replicated from languages with which Romani speakers came into contact.25

**English-induced de-allative futures**

English is presumably the most widespread contact language of the world, and wherever speakers of other languages have been in close contact with English, it is probable that they would replicate the English de-allative future in some way or other. A case in point are some immigrant communities in the USA and Canada, which used the English be-going-to future as a model for developing a use pattern involving goal-oriented motion as a conceptual source.

One example is provided by Old Order Mennonites in Waterloo County, Canada. In their variety of Pennsylvania German there is an emerging de-allative immediate future tense involving the verb geh ‘to go’ (Burridge 1995: 61ff.), very likely the result of intense language contact with surrounding English-speaking communities (but see section 1.4.6). Another example can be seen in the speech of a community of Yiddish speakers in Venice along the coast next to Los Angeles, which is strongly bilingual in English although Yiddish predominates in everyday conversations (Rayfield 1970). As a result of intensive contact with English, Yiddish speakers have created a future tense on the model of the English be-going-to future, illustrated in (26).

(26) Yiddish of Venice, California (Rayfield 1970: 69; quoted from Myers-Scotton 2002: 216–7)

All right, ge ikh kum-en bald.

all right go 1.SG come-INF soon

‘All right, I’m going to come in a minute.’

The last two examples, both taken from German-based varieties, are remarkable in one respect. What they have in common is that they are instances of replica grammaticalization rather than of ordinary grammaticalization. In the latter case we would have expected that the relevant speakers of Pennsylvania German and Yiddish had drawn on the pattern provided universally for de-allative futures, taking the format [X goes to Y] (see above). Since both examples lack an allative marker (encoded in the corresponding English construction by the erstwhile preposition to), it would seem that these speakers conceived English *to be going to* as an infinitive rather than an allative (or purpose) marker,26 and that they drew on an isomorphic use pattern existing in German, whereby the verb gehen ‘to go’ takes infinitival verbal complements without an allative (or purpose) marker, e.g. in Standard German *Ich gehe schlafen* (I go sleep.INF) ‘I go to sleep.’ Accordingly, the basis for grammaticalization was not the universally prescribed de-allative pattern, but rather a language-specific use pattern of the form [‘go’ + VERB.INF] that these speakers appear to have treated as an immediate equivalent of the English model.

**Conclusions**

What we observed on the situation in Europe also applies to other parts of the world: New future tenses are constantly arising in language contact situations on the basis of the principles sketched at the beginning of this section (for more examples, see Heine & Kuteva 2003: 551–5; see also section 6.1 for an example from Pipil). And 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 is determined by the particular circumstances surrounding the contact situation, such as the use patterns characterizing the replica language and the kind of category provided by the model language. For example, speakers in southeastern Europe can be predicted with a certain degree of probability to develop a new future tense by using a verb for ‘want’ since the de-volitive schema offers the primary conceptual choice for speakers in the Balkans, while immigrants in the USA are most likely to draw on the de-allative schema to create a new future construction in their L1 because the English be-going-to future provides the most immediately available model for replication. However,
the data available are not sufficient to determine what exactly it is that makes a certain conceptual model more attractive than alternative models in a given situation.

3.4 Some general issues

The analysis proposed above raises a number of questions, such as the following:

a. Is contact-induced grammaticalization unidirectional?
b. Are the conceptual sources used in contact-induced grammaticalization different from the ones to be found elsewhere?

There is yet another question that the preceding discussion raises but that does not seem to require further treatment, namely: Are there any limits to which kinds of grammatical structures can be replicated? Our data provide no basis for saying that there are limits; contact-induced grammaticalization affects in the same way verbal, nominal, and clausal morphosyntax, including clause combining, it affects ideational, textual, as well as interpersonal functions, and it affects derivational and inflectional forms, as well as free forms. It may well be, however, that on the basis of a larger corpus of data this observation needs to be modified.

In the present section we will deal with the two questions raised above in more detail.

3.4.1 Directionality

Grammaticalization is essentially unidirectional: At least 90 per cent of all instances of grammatical change can be assumed to be in accordance with principles of grammaticalization (Heine 2003b). And the same applies to grammaticalization occurring in, or being triggered by, language contact (Heine & Kuteva 2003: 560). For example, we saw in section 3.3 that language contact may give rise to new future tense categories via the grammaticalization of a verb meaning ‘want,’ while we are not aware of any example suggesting a reversed directionality, whereby language contact induced speakers to turn a future tense marker into a verb meaning ‘want.’

But, as shown in Heine and Kuteva (2003: 560–1), there are possible exceptions, such as the following case concerning post-verbal perfect markers of the Malaita Oceanic languages. As Keesing (1991: 331) argues, these markers have been further grammaticalized to topicalizing particles. In the following examples from the model language Kwaioi, (27a) illustrates the perfect use of no'o, while (27b) is an example of the use of the same marker as a topicalizing particle.

Some general issues

(27) Kwaioi (Eastern Oceanic; Keesing 1991: 330–1)
   a. e ‘akwa no’o.
      he run.away PERF
      ‘He has run away.’
   b. gila no’o la age-a.
      them TOP they do-it
      ‘They’re the ones who did it.’

This process appears to have been replicated in Solomons Pijin. As (28) suggests, the replica language exhibits the same structure as the model language Kwaio or other Oceanic languages of Malaita:

   a. hem- i ranawe nao.
      (him- he run.away PERF)29
      ‘He has run away.’
   b. hem nao i save.
      him TOP he know
      ‘He’s the one who knows.’

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 counterexample to the unidirectionality of grammaticalization. Still, such examples are rare; on the whole, contact-induced language change is in accordance with principles of grammatical change to be observed elsewhere, even if there may be specific circumstances triggering a violation of the unidirectionality principle.

3.4.2 Conceptual sources

Finally, we raised the question of whether the conceptual sources used in contact-induced grammaticalization differ from those to be found elsewhere. The answer is essentially in the negative; most cases of grammaticalization discussed in the previous sections were canonical instances also to be observed in situations where no language contact has been claimed. Thus, we witnessed common processes such as the following (see Heine & Kuteva 2002 for cross-linguistic evidence):

(i) Concrete nouns are grammaticalized to markers expressing case relations. Perhaps one of the most widespread developments concerns concrete nouns, such as nouns for body parts and other relational nouns, that are recruited to express locative, temporal, and other grammatical relations, and language contact provides a number of examples for this grammaticalization.
A different kind of this overall process can be illustrated with an example from Southeast Asia. A number of languages spoken in various parts of the world have grammaticalized a noun for ‘thing’ to a possessive marker linking two noun phrases. The construction that is used for this process takes the form [possessee, thing(-of) possessor] (Heine 1997a; Heine & Kuteva 2002). Instances of this schema can be found in Vietnamese, Khmer, Thai, and Lao, where a noun meaning ‘thing, object, stuff’ has been used with this schema and was grammaticalized to a possessive marker; note that the nouns employed appear to be etymologically unrelated. While there is no information on which language provided the model, there can be hardly any doubt that the spread of this grammaticalization was contact-induced (Matisoff 1991: 391; Enfield 2001: 260).

(ii) Verbs give rise to a wide range of grammatical functions, relating to case marking, clause embedding, tense, aspect, and modality, etc. Examples will be given, e.g. in section 6.2.2 involving the Southeast Asian languages Chinese, Hmong, Vietnamese, Thai, and Cambodian, where verbs for ‘give’ and ‘get’ have undergone a number of contact-induced grammaticalizations.

(iii) Allative and benefactive markers are commonly grammaticalized to indirect object markers, which again may turn into direct object markers. We will have examples in section 4.3.1 from Turkish-Laz contact, where the use of an allative/goal marker was extended to marking indirect objects, and in the Indian village Kupwar, Kannada speakers appear to have grammaticalized their dative postposition to also mark human direct objects.

(iv) Comitatives are a common source for other case functions including clause combining. One common line of grammaticalization concerns the extension of comitative markers to express instruments too; Basque provides an example of such a development: Its comitative case suffix -ekin is derived from an earlier postpositional phrase meaning ‘in the company of’ (Trask 1998: 318) and as a result of language contact acquired an instrumental function. Another line of development leads from comitatives to noun phrase-conjoining markers (‘and’) and eventually to clause-combining markers (Stoltz 1998; Stassen 2000; Heine & Kuteva 2002). That this line has been exploited in a language-contact situation is suggested by the following example. The Aztec language Pipil has no formal means for coordinating clauses, that is, clause-conjoining (‘and’) is not formally marked. Under the influence of Spanish, Pipil is said to have grammaticalized the relational noun -wan ‘with’ to a preposition wan ‘with,’ and wan has further developed into a noun phrase-conjoining conjunction ‘and.’ Finally, the use of wan has been extended to conjon clauses as well (Harris & Campbell 1995: 130). Thus, we seem to be dealing with an entire chain of development from relational noun to preposition (see section 3.1.2), from comitative ‘with’ to nominal conjunction, and finally to a clause-combining conjunction. Note that, if Harris and Campbell are right, this extended development must have happened within a fairly short time period, that is, only after the introduction of Spanish in the Americas (Harris & Campbell 1995: 127, 129–30, 147).

(v) Independent words turn into clitics and affixes. This process relates to the morphosyntactic component of grammaticalization (decategorialization), of which some examples were presented earlier (see especially section 3.1.2). A widely discussed example concerns the development of the English singular pronoun him into a transitive/causative marker in Melanesian Pidgin varieties (Sankoff & Brown 1976; Keesing 1991: 318–19). This process might have been triggered by the fact that in many of the substrate languages, that is, languages spoken by people using these pidgin varieties as a second language (e.g. Eastern Oceanic languages), there was a verbal suffix used for transitivizing or causativizing verbs (Keesing 1991: 318–19), and the speakers of these languages replicated this category in the pidgin varieties by grammaticalizing the personal pronoun him. The result of the process was that a self-standing pronoun assumed the grammatical function of a derivational marker and, consequently, lost its status as an independent word, turning into a verbal clitic and affix or, alternatively, becoming a lexicalized appendix of the verb. As we will see in chapter 4, contact-induced decategorialization may have more drastic results, giving rise to a new typological profile in the replica language.

To summarize, the kinds of conceptual sources that people select for grammaticalization in situations of language contact do not normally differ from those found elsewhere. A question that arises in this connection is the following: To what extent are the developments sketched here the result of language contact as opposed to universal principles of grammaticalization? 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, but what exactly this means is open to further research.

3.4.3 Restructuring

Grammaticalization accounts for much of what happens in grammatical replication leading to what is commonly described as increased “isomorphism” among languages, but it does not cover everything. For want of a better term, we will refer to processes of grammatical replication that cannot be accounted for with reference to grammaticalization as restructuring. This term has been discussed in various schools of linguistics; especially in pidgin and creole studies where it has received extensive treatment and has been related to a number of different phenomena (see especially Neumann-Holzschuh & Schneider 2000 and the contributions therein). The reader is referred to Aikhenvald (2002) for a wealth of examples suggesting that grammatical replication may lead to structural
changes in the replica language without involving principles of grammaticalization.

A paradigm case of restructuring concerns word order, where people rearrange the order of meaningful units in one language on the model of another language. The literature on language contact abounds with examples, hence there is no need to illustrate this process. Note, however, that many cases of presumed contact-induced word-order change are not really cases of restructuring but rather the by-product of grammaticalization; we saw examples of such cases in preceding sections (e.g. section 3.1.3).

Aikhenvald’s (2002: 60) detailed analysis suggests that on the one hand, restructuring neither corroborates nor contradicts principles of grammaticalization; on the other hand, it also suggests that some cases that have been treated as instances of restructuring can perhaps more profitably be reinterpreted as epiphenomenal effects of grammaticalization. The following examples may illustrate this. Aikhenvald (2002) discusses a wide range of data illustrating how East Tucanoan languages of the Vaupés basin in northwestern Brazil served as models of replication for Tariana, an Arawak language (Aikhenvald 2002: 98–9). For example, East Tucanoan languages have a morpheme meaning ‘also’ (e.g. Tucano ke’ra) which also marks the plural of some kinship terms and serves as an associative plural marker with other kinship terms and personal names. In Tariana there is a structurally similar morpheme (-sinĩ) which also means ‘also’; at the same time it marks the plural of a few kinship terms and has uses as an associative plural marker on proper nouns and kinship terms. There is reason to assume that this isomorphism is due to a replication process from East Tucanoan languages to Tariana, and there is also reason to maintain that this is an instance of replica grammaticalization (see section 3.1.3) whereby Tariana speakers created a new number category by grammaticalizing a nominal modifier ‘also’ to a plural marker on the model of East Tucanoan languages. Assuming that this reconstruction is correct, we would be dealing with a restructuring process in Tariana resulting from grammaticalization. This hypothesis is supported by the fact that, first, no Arawak language other than Tariana has an associative plural and, second, that the associative plural in the replica language Tariana does not appear to have been grammaticalized to the extent that it has been in East Tucanoan languages (Aikhenvald 2002: 98–9); on the basis of many other cases of contact-induced change, replica categories are likely to be less grammaticalized than their model categories (see section 3.4.5 below).

Another example adduced by Aikhenvald to illustrate restructuring concerns plural agreement with numerals. In the Arawak language Baniwa, nominal plural marking is confined to human nouns. Speakers of the fellow North Arawak language Tariana have a more complex structure of number marking, but one that exactly corresponds to that of their East Tucanoan-speaking neighbors: while Tariana basically retains the distinction between human and inanimate number marking, there nevertheless appears to be a gradual extension from human to non-human referents: The use of the plural form is optional when there is a numeral attribute for ‘2’ or ‘3’ but obligatory for nouns above ‘3’. Assuming that this is another case where Tariana speakers have replicated use patterns of East Tucanoan languages, this suggests that Tariana speakers have extended plural marking from human to non-human referents. As is to be expected in this process, extension was gradual, affecting first higher numerals, before spreading to the lower numerals ‘2’ and ‘3’, where plural marking is optional, that is, where it has not been conventionalized. Such a development is in line with an overall extension pattern of grammaticalization according to which the use of markers for human referents tends to be extended to denote inanimate referents; it is suggestive of context extension (see section 2.2.2, 3.1.2).

Not seldom, contact-induced grammaticalization entails restructuring. Grammaticalization involves first and above form–meaning units that are put to new, i.e. to more grammatical uses. But these units are part and parcel of phrasal, clausal, or even larger use patterns which are also affected by grammaticalization. Accordingly, contact-induced grammaticalization is not confined to morphology but tends to include larger discourse structures too. For example, Tariana speakers of northwestern Brazil use their interrogative pronouns, such as iwana ‘who?’, to mark relative and complement clauses on the model of Portuguese (Aikhenvald 2002: 182). But replica grammaticalization from interrogative to subordinating marker is not restricted to these markers. Complement clauses precede the predicate of the main clause in Tariana, but in the case of this grammaticalization, Tariana speakers replicated the Portuguese use pattern, where the complement clause follows the main clause predicate. Accordingly, (29a) reflects the Portuguese structure of (29b). Thus, the extension from interrogative to relative/complement marking appears to have triggered restructuring, that is, the introduction of a new arrangement of meaningful units.

(29) Tariana (North Arawak) and Portuguese (Aikhenvald 2002: 182–3)

a. Tariana
di- sata dhima- pidanã kwana- sika
INFRR.INTER rest
'He asked who remained.'
b. Portuguese
   ele perguntou quem restou.
   he asked who remained
   'He asked who remained.'

In a number of other cases, what appears to be suggestive of restructuring can equally well be interpreted as being suggestive of grammaticalization. In one type of restructuring, the replica language exhibits two (or more) structural options (say, A and B) to express one and the same grammatical function. When the model language has only one structure (A) for that function, restructuring may have the effect that speakers of the replica language narrow down these options to A, thereby establishing a one-to-one equivalence relation with the model language; we proposed the term narrowing for this process in section 2.2.4. Now, a common grammaticalization process involves context generalization (see 1.3), whereby the use of some form or construction (A) is extended, and this process can be at the expense of another form or construction (B), which subsequently may be lost. In such cases, restructuring can be induced by grammaticalization.

The following example illustrates this process. Not uncommonly in the Arawak languages of northwest Amazonia, word order is pragmatically based, where, for example, the order of words in adpositional or possessive constructions depends on whether the possese or the adposition is focused or not. Thus, in the Arawak language Baniwa, (30a) and (30b) have the same meaning, but in (30b), João is in focus. East Tucanoan languages on the other hand have a fixed word order [dependent – head]; that is, the possese invariably follows the possessor and the adposition its complement. Now, the North Arawak language Tariana, which has been heavily influenced by East Tucanoan languages, appears to have replicated the structure found in these languages by generalizing (30a), which is identical in Baniwa and Tariana – with the result that (30b) is an ungrammatical structure in Tariana: there is no variation in word-order, that is, there is now only one word-order arrangement that is isomorphic with that of East Tucanoan languages (Aikhenvald 2002: 167).

(30) Baniwa (North Arawak; Aikhenvald 2002: 167)

   a. João i- siu
      John INDF- to
   b. ri- siu João
      3.SG.NF to John
      'to John'

Overall, the relationship between restructuring and grammaticalization remains unclear. While there are examples to show that the former is an outcome of the latter, there are also examples to suggest that restructuring constitutes a mechanism that may trigger grammaticalization. Once again, language contact between East Tucanoan languages and the North Arawak language Tariana in northwest Amazonia provides an example (see also section 4.1.4). Most Arawak languages of the Upper Rio Negro area have a number of locative case markers, distinguishing between locative/directional, directional/allative, ablative, and perlocutive ('along') case marking. East Tucanoan languages on the other hand have only one multifunctional locative case, expressing location, direction to and from, as well as temporal functions. As the description by Aikhenvald (2003: 8–10) suggests, Tariana has given up the locative case distinctions characterizing its fellow Arawak languages by replicating the East Tucanoan model: being strongly influenced by East Tucanoan languages, Tariana speakers have generalized one of their case suffixes, the allative case suffix -se, to a catch-all locative case used for static location, ablative, and allative functions, thereby matching the locative case polysemy of East Tucanoan languages.31

On the surface, this is a straightforward instance of restructuring not involving grammaticalization,32 as a matter of fact, however, even this case has been shaped by grammaticalization. What appears to have happened in Tariana is that the allative case suffix -se was extended to contexts previously reserved for other case functions, such as locative, directional, and ablative functions, thereby assuming a wider range of functions. This process involved two main parameters of grammaticalization (see section 1.3, (5)), extension and desemanticization, in that – with the extension to more contexts – the erstwhile case marker -se lost its specific function as an allative marker in favor of a more general locative function. On the basis of the description provided by Aikhenvald (2003: 7–10) it would seem, however, that in this case grammaticalization was an epiphenomenal product of restructing in that, rather than saying at changing the functional and contextual structure of a case marker, Tariana speakers are likely to have been primarily concerned with establishing an equivalence relation (or structural isomorphism; see section 6.1) with the corresponding locative case marker of East Tucanoan languages.

To conclude, restructuring is an important mechanism of contact-induced replication, but not infrequently it appears to be shaped or influenced by grammaticalization; but, alternatively, it may also be a force that determines what can or cannot be grammaticalized.

3.4.4 On spontaneous replication

As the preceding discussion may have shown, contact-induced grammaticalization is a ubiquitous process. Still, not all processes of grammatical replication do
correspond to the definition proposed in section 1.3, and in the present section we will look at some of these processes.

One kind of process concerns cases where speakers simply replicate, or copy, a grammatical category of the model language in the replica language without invoking principles of grammaticalization. This means typically that they adopt a grammatical category of the replica language to the structure of a corresponding category in the model language; we will discuss this process in section 4.1.4. Another kind of process, to be discussed in section 4.4.3, concerns renewal, whereby an existing mode of grammatical expression is replaced by a new, that is, less grammaticalized mode, thereby initiating a new cycle of grammaticalization.

Finally, there is what may be loosely called spontaneous replication. What this term is meant to refer to is that speakers of an L1 having an imperfect knowledge of another language, L2, notice that there is a grammatical category in an L2 for which there is no equivalent category in the L1. In such a situation they may apply principles of replication, including grammaticalization, to guide them in their use of the L2. Having only an imperfect knowledge of the L2, they may apply principles of grammaticalization either less extensively or more extensively than is done by L1 speakers of that language. The result is either an undergeneralization or an overgeneralization of these principles (Deborah Ziegeler, p.c.).

The following example may illustrate the notion spontaneous replication. Perhaps the majority of the world’s languages do not have definite or indefinite articles. Now, when speakers of such languages (= L1) acquire Western European languages as L2s, which are characterized by the presence of strongly grammaticalized articles, it is likely to happen that such speakers do not use articles where L1 speakers of the relevant Western European language would use them (undergeneralization), or else use them also in contexts where they are not used by L1 speakers (overgeneralization). The literature on language contact between Slavic languages on the one hand and Germanic and Romance languages on the other offers a wealth of examples of this case of spontaneous replication, as already pointed out by Schuchardt (1884: 112) with reference to Slavic L1 speakers in contact with German and Italian.

While spontaneous replication is no doubt relevant for understanding the mechanism giving rise to new use patterns and grammatical categories in contact situations, we have little to say about it since our concern is not with idiosyncratic or temporally restricted language use but primarily with use patterns and categories that have acquired some stability of expression across space, time, and social interaction. Accordingly, when speakers of languages having no articles are in contact with languages such as English, they tend either to undergeneralize or overgeneralize articles in their use of English in spontaneous speech, and vice versa. But once such idiosyncratic usage is conventionalized we predict that there will be directionality, in that speakers of replica languages having no articles are likely to create articles on the model of a language having articles, while speakers of languages having articles are unlikely to lose their articles when in contact with a language lacking articles, since such a development would contradict principles of grammaticalization.

3.4.5 Space, time, and degree of grammaticalization

One observation that surfaces in studies of contact-induced grammatical replication is that, as a rule, replicated grammatical categories are less elaborated, i.e. less grammaticalized, than the corresponding model categories. In fact, what characterizes areas of intense language contact is that there tend to be continua of grammaticalization that correlate with space and time. Such continua may range from use patterns that are optional, i.e. determined by discourse-pragmatic factors at one end of the area to obligatory, fully grammaticalized use patterns at the other end.

As the findings made by Friedman (1976; 1994) show, this is the situation not uncommonly encountered in the Balkan sprachbund. He demonstrates that there is a synchronic continuum extending from discourse-based variation in Bulgarian to fully grammaticalized structures in Macedonian and Albanian (Friedman 1994). The following example illustrates this situation.

One of the properties characterizing Balkan languages is what is technically known as object reduplication, which means that the direct or indirect object noun or pronoun is cross-referenced in the verb phrase by a clitic pronoun agreeing with the object in gender, number, and case or case function (see section 5.2.3.1 (b)). Example (31) from Macedonian illustrates this structure.

(31) Macedonian (Friedman 1994: 102)

    Na mamočeta mu ja davan knigata.
    Na boy.DEF him.DAT it.ACC.F give.1.SG.PRES book.DEF.F
    'I give the boy the book.'

In Macedonian and Albanian, object reduplication is obligatory with indirect objects as well as in some other contexts; in other contexts again it correlates positively with definiteness, specificity, or discourse prominence (see Friedman 1994: 102–4). In Bulgarian on the other hand, object reduplication is entirely
facultative, being subject to discourse-bound variation. Thus, (32) would be a normal equivalent of the Macedonian sentence in (31), even though (31) could also occur in Bulgarian in specific pragmatically marked contexts. Object reduplication in Bulgarian does not correlate significantly with such parameters as definiteness, specificity, word order, or disambiguation of case functions but is associated with topicality, and indirect objects are reduplicated 2.5 times more than direct objects.

(32) Bulgarian (Friedman 1994: 104)

Na момчето дам книгата.
to boy.DEF give.1.SG.PRES book.DEF.F
'I give the boy the book.'

To summarize, object reduplication has been conventionalized to categorical status in Albanian and Macedonian with indirect objects and to some extent with definite direct objects, whereas in Bulgarian it has remained a use pattern whose occurrence is determined not by morphosyntactic but entirely by discourse-pragmatic variables. This difference correlates with the history of object reduplication:

The dialects reflect in synchronic spatial terms the diachronic development. Just as earlier Slavic documents that show reduplication began first with pronouns and earliest in southwest Macedonia and latest in northeast Bulgaria, so too in the modern northern and eastern Macedonian dialects that are transitional to Serbo-Croatian and Bulgarian, e.g. in Kumanovo and Kukus/Kilkis, object reduplication occurs with less consistency than in the west-central dialects [...]. (Friedman 1994: 105)

What these findings suggest is that there is a significant correlation between space, time, and degree of grammaticalization: Southwest Macedonia represents the earliest stage, where object reduplication has developed characteristics of an obligatory, full-fledged category, while northeast Bulgaria at the other end of the continuum represents the youngest stage, where grammaticalization is still in its earliest stages, in that object reduplication is still a 'discourse-bound pragmatic device characteristic of colloquial style' (Friedman 1994: 105).

The label “Macedonian” refers to a number of varieties differing from one another in substantial typological properties. These differences clearly exhibit an areally defined pattern (Friedman 1976). The Macedonian possessive perfect (‘have’-perfect) uses ima `have` as an auxiliary and the main verb in the past passive participle (PPP). The structure of the possessive perfect in Macedonian can be described most appropriately in terms of a correlation of linguistic and areal parameters: There is a continuum extending from northeastern Macedonia, where there is essentially no possessive perfect, to southwestern Macedonia, where there is a fully grammaticalized possessive perfect, while intermediate varieties of Macedonian exhibit intermediate forms of grammaticalization. This areal patterning also correlates with diachronic development: As the data presented by Friedman (1976; see also Drinka forthcoming) suggest, there was a wave of innovation giving rise to a fully grammaticalized possessive perfect in southwest Macedonia, gradually diffusing northeast, and the farther it diffused the fewer properties of a perfect category it acquired. The extreme northeast of Macedonia has not been affected by this wave, hence there is (as yet) no genuine possessive perfect.

What these data suggest is that the hypothesis proposed at the beginning of the present section is well-founded: replica categories differ in a principled way from their corresponding models in that they tend to exhibit properties such as the ones listed in (33).

(33) Properties of replica categories, distinguishing them from model categories

a. They are used less frequently.

b. They are associated with a smaller range of contexts.

c. They are less clearly associated with the grammatical meaning.

d. Their use is more likely to be determined by discourse-pragmatic than by morphosyntactic parameters.

e. Their use is more likely to be optional than obligatory (see chapter 2 for more details).

A number of authors (see, for example, Soper 1996; 287; Aikhenvald 2002: 7) have pointed out in some way or other that replica categories frequently differ from model categories on the lines sketched in (33); we will discuss a number of examples illustrating this point in section 4.4. Within the framework of code copying, which also includes replication, Johanson (e.g. 1992; 2002a) observes that a copy can always be distinguished from its original; note, however, that the evidence adduced by Johanson is of a different nature from the one used here: rather than grammaticalization, Johanson invokes adaptation (or accommodation) to the structure of the replica language (= the base code in his terminology) that accounts for the fact that the copy (or imitation) differs from the model (= the model code).

The following is a canonical example illustrating the rise and structure of a replica category. Most North Arawak languages of northeastern Brazil do not have any copula verb, while all Tucanoan languages have one. A notable exception is Tariana, which has a locative-existential copula alia; note that Tariana has been deeply influenced by East Tucanoan languages, which have copulas corresponding to Tariana alia. But unlike East Tucanoan languages, Tariana has no conventionalized copula for identity and equation clauses. But
younger people may extend the use of the locative-existential copula *alia* to express such clauses. In doing so, they establish structural isomorphism with East Tucanoan model languages such as Tucano. The model and the replica categories differ, however, in the following way: While the former is an obligatory category, i.e. one which must be used, the replica category is not—it is strongly pragmatically marked, being used only when contrastive focus is expressed.

There is a cross-linguistic grammaticalization path according to which the use of a locative-existential marker (e.g., *X is at Y*) may be extended to identifying (*The X is Y*) and/or classifying copula uses (*Y is an X*). Initially, this extension is optional and tends to be confined to pragmatically marked contexts (stage I). It is only at a more advanced stage that the marker loses its pragmatic function and becomes an obligatory copula (stage II). This suggests that young Tariana speakers have replicated the East Tucanoan stage II copulas by grammaticalizing their locative-existential copula *alia*, but that grammaticalization has not proceeded beyond stage I.

To conclude, Tariana *alia* as an identity/equation copula of younger speakers has most of the properties of a replica category listed in (33), differing from the East Tucanoan model categories in the following way: Being restricted to one pragmatic function, it is used less frequently, is associated with a smaller range of contexts, its use is optional, and since its primary function appears to be pragmatic, expressing a copular relation is not its primary function.

On the basis of (33) it seems possible to determine in a situation where no diachronic information is available which is the model and which is the replica category. Accordingly, if we find a case which does not support this hypothesis then such a case is in need of explanation. One factor that may be considered to constitute a problem to the hypothesis is time. In the early stages of language contact, replica categories tend to exhibit a larger range of the properties listed in (33), but the longer and the more intense the contact is, the more of the properties tend to be lost, and in the end, the two categories may become structurally indistinguishable. Obviously, in such cases, (33) will no longer apply.

### 3.5 Conclusions

Our main concern in the present chapter was to relate grammaticalization theory to contact-induced language change; in doing so, we were building on the foundation laid by Haase (1992) and Nau (1995). While there are examples that appear to be at variance with the unidirectionality hypothesis (section 3.4.1), overall grammatical replication was found to be in line with principles of grammaticalization. We were dealing with cases where in connection with language contact verbs meaning 'want' or 'go to' developed into future tense markers, relational nouns into adpositions, or nouns for 'thing' into possessive/genitive markers; but we are not aware of any language where language contact was responsible for a development in the opposite direction, i.e. from future marker to verbs for 'want' or 'go to', etc.

These observations furthermore suggest that contact-induced change as studied here is gradual rather than abrupt, involving a process roughly of the following kind: 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 a high frequency of use, and eventually may emerge as a full-fledged grammatical category.

Still, the chapter raises a number of questions that we are unable to answer, such as the following: To what extent are the developments sketched here the result of universal principles of grammaticalization and to what extent are these due to the specific factors obtaining in situations of language contact, including the structures of the model and the replica languages? What the observations made seem to suggest is that in addition to the kinds of motivation that have been defined so far (see, for example, Heine, Claudi & Hünnemeyer 1991; Hopper & Traugott 1993) there are others that have to be taken into account in the study of grammaticalization. They relate to communicative goals, such as making the categories existing in the languages in contact mutually compatible and more readily intertranslatable, or to social goals, inducing speakers, for example, to talk like their neighbors, or in a way that is socioeconomically or otherwise profitable.

Another question is how the various processes of contact-induced grammaticalization to be observed in a given language are related to one another: do they interact with or presuppose one another? For example, how are the different processes to be observed in one and the same replica language, e.g. in Pipil (Campbell 1987), Basque (Haase 1992), or Tariana (Aikhenvald 2002), interrelated? There is no conclusive answer to this question—nor has this question ever been addressed in any systematic way in grammaticalization studies.

There is one question where we are able to propose at least a partial answer, namely the question of whether there are certain functional categories that are more likely to arise in situations of language contact than others. As we saw in the course of this chapter, certain kinds of categories are fairly likely to evolve, such as new tense and aspect markers, adpositions, case markers, conjunctions, discourse markers, definite and indefinite articles, etc., while there is little evidence for some other categories arising as a result of language contact,
such categories relating to personal deixis or negation. We have not been able to detect any more general principle underlying this differential behavior — one that would be able to account, for example, for the fact that future tense markers are highly likely to grammaticalize (see section 3.3) while past tense markers are considerably less so.

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 chapter may have shown is that these are in no way mutually exclusive alternatives; on the basis of our data on contact-induced grammaticalization both can be expected to be potentially present in virtually any given case of language change.