

CLIPP

Christiani Lehmanni inedita, publicanda, publicata

titulus	New reflections on grammaticalization and lexicalization
huius textus situs retis mundialis	https://www.christianlehmann.eu/publ/lehmann_new_reflections.pdf
dies manuscripti postremum modificati	28.04.2005
ocasio orationis habitae	Symposium "New Reflections on Grammaticalization"; 17.-20.6.1999, Potsdam
volumen publicationem continens	Wischer, Ilse & Diewald, Gabriele (eds.), <i>New reflections on grammaticalization</i> . Amsterdam & Philadelphia: J. Benjamins (TSL, 49).
annus publicationis	2002
paginae	1-18

New reflections on grammaticalization and lexicalization

Christian Lehmann

Abstract

The essential difference between grammar and lexicon is the following: The grammar is concerned with those signs which are formed regularly and which are handled analytically, while the lexicon is concerned with those signs which are formed irregularly and which are handled holistically. A sign is lexicalized if it is withdrawn from analytic access and inventorized. On the other hand, for a sign to be grammaticalized means for it to acquire functions in the analytic formation of more comprehensive signs. Both processes regularly, but not necessarily involve a reductive component. Consequently, grammaticalization is not the mirror image of lexicalization.

The genesis of members of minor word classes, in particular adpositions and conjunctions, has often been treated as an instance of grammaticalization. However, minor word classes are not necessarily classes of grammatical formatives. In particular, there are more lexical and more grammatical adpositions. For instance, before *auf Grund (von)* ‘on the basis of’ can ever get grammaticalized to a grammatical preposition, it must first be lexicalized to the lexical preposition *aufgrund (von)*. In this sense, grammaticalization presupposes lexicalization.

Thus, lexicalization and grammaticalization are processes that have much in common and are, to a certain extent, parallel. The mirror image of grammaticalization is degammaticalization, and the mirror image of lexicalization is folk etymology.¹

¹ This paper was first presented at the Philosophical Faculty of the University of Tübingen in December 1998. I thank the participants of that guest lecture and the audience of the Potsdam symposium, in particular Carmen Pensado and Martin Haspelmath, for helpful criticism and suggestions.

1. Theoretical bases

The purpose of this contribution is to clarify the concepts of 'grammaticalization' and 'lexicalization' in their mutual relationship (cf. Moreno Cabrera 1998). Such an explication cannot possibly justify all previous uses of these concepts, in particular not all of those reported or endorsed in Lehmann 1989. Also, to the extent that it tries to be consistent, the explication necessarily leads to unwonted results.

1.1. Analytic and holistic approaches

Given an object of cognition of some complexity, the human mind has two ways of accessing it. The analytic approach consists in considering each part of the object and the contribution that it makes to the assemblage by its nature and function, and thus to arrive at a mental representation of the whole by applying rules of composition to its parts. The holistic approach is to directly grasp the whole without consideration of the parts. This can be done if the object itself is already familiar or if, by its contours or its contextual setting and function, it bears an essential analogy to some familiar object.

The two approaches complement each other in various ways.

1. If confronted with a familiar object, we tend to take the holistic approach; if confronted with an unfamiliar object, we take the analytic approach.
2. For a given specific object, we can often switch between the two approaches by making a fresh analysis of what used to be familiar or by disregarding compositional parts in favour of the function of the whole.
3. A given complex object may only be analyzed in certain parts or aspects, while the internal structure of other parts remains out of consideration.

To illustrate:

- E1. a. X chooses the correct approach to Y.
 b. X takes the correct approach to Y.

In E1.a the combination of the relational noun *approach* with its prepositional dependent, and the combination of the transitive verb *choose* with its direct object, are interpreted by general rules of semanto-syntax.

In E1.b the combination *X [takes (Z) approach] to Y* constitutes a proper part of the sentence. Its contour and function are analogous to the simpler construction *X approaches Y (in a Z way)*.

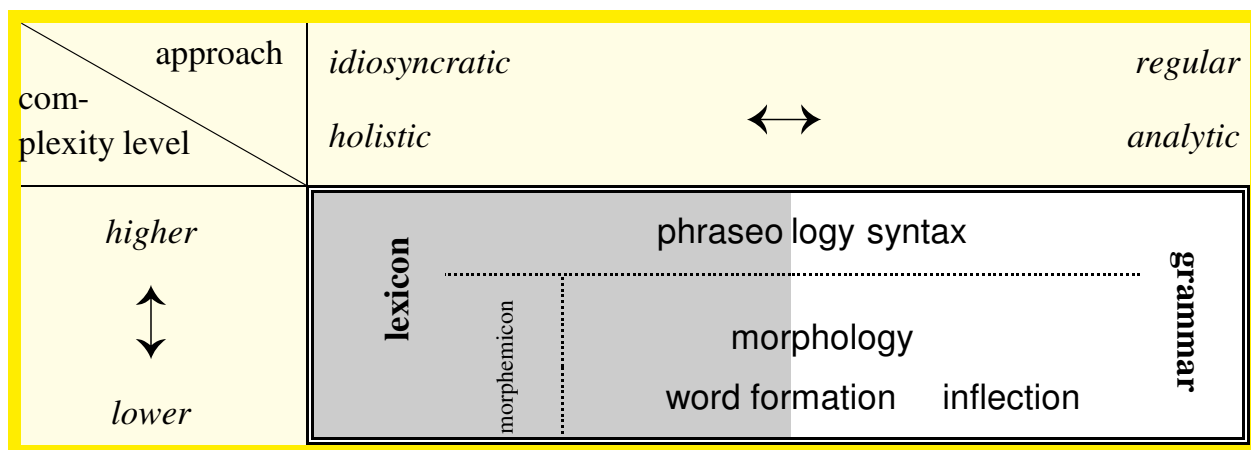
E1 thus illustrates the above generalizations:

1. The relatively unfamiliar collocation *choose approach* is construed analytically, while the familiar collocation *take ... approach* is construed holistically.
2. The collocation *choose ... approach* could instead be accessed holistically, whereby the specific contribution of *choose* would essentially be foregone, and the whole would be largely synonymous with *take ... approach*; and again, the collocation *take ... approach* could instead be accessed analytically, whereby *take* would regain a more literal sense (contrasting, e.g., with *abandon*), and the resulting constructional meaning would be slightly different.
3. The holistic approach treats *take ... approach* as a proper part of the construction, which it is not in the analytic approach. However, this does not mean that the construction of E1.b is an unanalyzed whole, since we can still integrate the contributions of each of the elements in the slots X, Y and Z with the help of general compositional rules.

1.2. Lexicon and grammar

The system of linguistic signs is subdivided into lexicon and grammar. The relationship between the two components and their organisation in terms of subcomponents is represented in S1.

S1. *Lexicon and grammar*



On the horizontal axis of S1, the lexicon differs from the grammar. The vertical axis is associated with the hierarchy of levels of grammatical structure. The latter is, of course, only partially represented in the lexicon. The most idiosyncratic part of the lexicon is the morphemicon, which contains all the lexical and grammatical morphemes of the language. Accessing a collocation XY holistically means treating it as an entry of the inventory, as a lexical item. If this mode of access to XY gets more prominent in language activity, it is the initial step of the lexicalization of this sequence.

Accessing a collocation XY analytically means treating it as a grammatical construction in which the structural properties of either X or Y or both matter and make a regular contribution to the pattern. If this mode of access gets more prominent in language activity, it is the initial step of the grammaticalization of XY.

In the following two sections, we will see that lexicalization and grammaticalization apply alternatively to a construction, while they apply successively to an item.

2. Lexicalization and grammaticalization as alternatives

2.1. Verb and coverb in Jaminjung

The initial step in the processes of grammaticalization and lexicalization does not yet involve any noticeable changes in the collocation. So far, those are but alternative modes of treating the collocation XY. However, they lay the ground for the further fate of XY. To see this, let us take an example from Jaminjung, a Non-Pama-Nyungan language of Northern Australia.² The language has two word classes which are at stake here. One is the class of verbs which is closed and comprises about 30 members. Understandably, these verbs have a very general meaning, are highly polysemous and in this resemble the function verbs or even auxiliaries of more familiar languages. The other class is called coverbs. It is an open class which comprises such concepts as are covered by verbs and adverbs in more familiar languages. The coverbs have valence just like the verbs, but they do not take a subject and instead combine with a verb much like an adverb does. E2 is an example, combining the verb *-angga* with the coverb *warlŋginy*.

E2.jirramabuny-anggawarlŋginy
 two 3.DU-GO.PRS on.foot³
 ‘two are walking’ (Schultze-Berndt 2000, DB, D14105)

Now the collocation of verb plus coverb can be accessed either analytically or holistically. In the former case the verb functions like a grammatical verb. E3 is an example.

E3.jiwayurru buru-mayan ga-gba=biya
 bower.bird return-CONT 3.SG-BE.PST=NOW
 ‘the bower bird was going back and forth then’ (Schultze-Berndt 2000, 2-79)

The meaning of the sentence is construed in a bottom-up fashion by the following compositional operations. First, the coverb is combined with its (nominal) dependents – none

² All the data and most of the analyses of this language are taken from Schulze-Berndt 2000.

³ CONT continuous, DU dual, PRS present, PST past, SG singular.

in E3. Then the verb is first combined with its nominal dependents – here, the subject – and next with the coverb phrase. If they have nominal dependents in common, these and their roles are unified. In this way, the meaning of the whole is a regular function of the meaning of the parts and their relations.

In this approach, the collocation of verb and coverb works as a pattern, with two slots to be occupied by members of two clearly distinct categories, one of which – the verb – constitutes a structured paradigm. The two slots can be filled essentially in mutual independence. To the extent that the bulk of the concrete meaning of the whole is contributed by the coverb, the verb only functions as an aspectual operator which converts the coverb into a finite clause. In this construction, it is a grammatical verb.

If, instead, the collocation of verb plus coverb is accessed holistically, the verb retains a concrete meaning, as in E4.

E4. *jirrib ga-rdba-ny*
 married 3.SG-FALL-PST
 ‘he/she got married’ (Schultze-Berndt 2000, JAM 013)

Here, the verb and the coverb do not each take their dependents. Instead, the complex formed by the two essentially functions like a derived verb in whose meaning the meaning of the parts in isolation cannot necessarily be re-identified and which takes dependents as a whole. Although the collocation bears an outer resemblance to the pattern observed in the former case, no filling of the slots is possible which could bear a semantically regular paradigmatic relationship to the one of E3. In this approach, the collocation of verb and coverb functions as a simple verb, enriching, as it were, the inventory of the verbs. The complex is, thus, lexicalized.

2.2. The German preposition *zu*

Similar examples could, of course, be adduced from serial verb constructions all over the world. Often it is the same verb which gets both grammaticalized to a function verb and, finally, to an aspectual operator, and in other collocations gets lexicalized by merging with a contextual component. This will be illustrated with the German preposition *zu*. Synchronically, this preposition has a number of uses which vary in the extent to which they form regular patterns. Originally, this was a local preposition with allative and locative functions, similar to French *à*. These two functions appear in E5.

E5. a. Der Prinz begab sich zur Königin. ‘The prince betook himself to the queen.’
 b. Der Prinz residierte zu Potsdam. ‘The prince resided at Potsdam.’

The allative use of *zu* evolves into a purposive one, which in the end gives us the subordinator of the infinitive. The development can be envisaged as proceeding along a gradience whose steps are illustrated by E6.

- E6. a. Der Prinz begab sich zur Königin. ‘The prince betook himself to the queen.’
 b. Der Prinz begab sich zur Jagd. ‘The prince betook himself to the hunt.’
 c. Der Prinz begab sich zum Jagen. ‘The prince betook himself to hunting.’
 d. Der Prinz entschied sich zum Jagen. ‘The prince decided in favor of hunting.’
 e. Der Prinz entschied sich zu jagen. ‘The prince decided to hunt.’

At each stage of this evolution, the preposition is sensitive to the syntactic categories that constitute its context.⁴ The dependent is a concrete and an abstract NP, respectively, in E6.a and b. The noun in the abstract NP is an infinitive in E6. c and d. Finally, the complement of the preposition is a bare infinitive in e. At the same time, the superordinate verb is one of locomotion as long as the preposition has an allative or purposive sense. Once the latter does no more than subordinate an infinitive, the superordinate verb can be a complement-taking verb. In the end, such verbs may even require *zu* as a marker introducing the dependent infinitive. At this endpoint of the grammaticalization process, *zu* is but an obligatory slot filler in a construction which is formed by compositional rules of syntax.

The uses of *zu* which form the chain leading from E5.a = E6.a to E6.e instantiate productive patterns. This is not so with the locative use appearing in E5.b. The combination of locative *zu* with town names is obsolete. If a toponym which is not a town name, such as *Capri*, *Hessen*, *Dänemark*, is substituted for *Potsdam*, the sentence becomes outright ungrammatical. The combination of locative *zu* with common nouns is illustrated in E7.

- E7. a. Der Prinz war zu Hause. ‘The prince was at home.’
 b. Der Prinz kam zu Pferde. ‘The prince came on horseback.’
 c. Der Prinz siegte zu Wasser und zu Lande.
 ‘The prince triumphed on land and sea.’

Each of the collocations of *zu* with its dependent in E7 is lexicalized. In E7.a, we cannot substitute *Hause* by *Hotel*; *Esel* ‘donkey’ instead of *Pferde* in E7.b, and *Fluß/Ufer* ‘river/bank’ instead of *Wasser/Lande* in E7.c are impossible. The collocations do instantiate a pattern, viz. the one illustrated by E5.b; but the pattern is obsolete. Consequently, the phraseologisms of E7 are but remnants of an earlier pattern. In regular locative prepositional phrases, *zu* is replaced by other prepositions.

2.3. Grammaticalization of a construction

⁴ Cf. Diewald (in this vol.) for a detailed analysis of the role of the context in grammaticalization.

The examples show that one cannot properly say that a given element as such is either grammaticalized or lexicalized. Instead, it is the construction of which the element is a constituent which may embark on either course.⁵ If this is so, then the grammaticalization of a construction does not entail the grammaticalization of any of its component elements. Consider, for example, the construction consisting of a *verbum dicendi* as superordinate verb and a subordinate clause, as in E8.b.

- E8. a. Irvin apologized, he didn't hit me on purpose.
 b. Irvin said he didn't hit me on purpose.

It seems appropriate to say – as it has been said for at least a century – that the collocation of a *verbum dicendi* and a sentence specifying the content of the communication, as it appears in E8.a, has been grammaticalized into a complex sentence in E8.b. No formative is visible in E8.b which would have specifically undergone grammaticalization in this process.

E9 translates E8 into German.

- E9. a. Erwin entschuldigte sich; er habe mich nicht absichtlich getroffen.
 b. Erwin sagte, er habe mich nicht absichtlich getroffen.

Here, the sentence rendering the content of the speech act is in the subjunctive present in both cases. In E9.a, the subjunctive expresses that the speaker does not vouch for what he is saying. In E9.b, it is triggered by the governing verb. Consequently, the subjunctive becomes more grammaticalized in this development.

The analogy between the English and the German case warrants the generalization that the grammaticalization of a particular formative is but a by-product of the grammaticalization of a construction. If there is an element that mediates the relation between the constituents of a construction, then grammaticalization of the construction will involve grammaticalization of this element. But if there is no such element present, grammaticalization may proceed, anyway.

Those who are familiar with my earlier work on grammaticalization will notice that this implies a slight extension of the concept. The traditional conception, which centers around the grammaticalization of a linguistic sign, sees this in the intersection of a set of paradigmatic and syntagmatic relations. This, of course, entails the existence of a construction that the sign in question is a part of and that is grammaticalized together with the latter (cf. Lehmann 1995[S], §2 and 1995[T]:175-178). What I propose here is to apply

⁵ As Meillet (1915:170) says: “c'est le rôle dans la phrase qui décide de tout”. Cf. also Bybee et al. 1994: 11.

the criteria of paradigmatic and syntagmatic autonomy to a construction regardless of whether it contains a constituent in which the symptoms of grammaticalization crystallize. This is a small and controlled extension which avoids the undesirable consequence that anything which enriches the grammar must be called grammaticalization. The extension is of relevance for the analysis of isolating structures, on whose grammaticalization much more empirical work is necessary.

3. Lexicalization and grammaticalization in succession

3.1. Lexical and grammatical members of word classes

Up to now, we have viewed grammaticalization and lexicalization as two alternatives which may apply to a given construction. However, there is yet another sense in which the two processes complement each other. Before we can turn to it, the theory of word classes needs some clarification.

Word classes are sometimes divided into lexical and grammatical word classes. It is, for instance, assumed that nouns, adjectives and verbs form lexical word classes, while prepositions and conjunctions form grammatical word classes. As a matter of fact, there are lexical and grammatical words in each of the word classes, as illustrated in T1 from Spanish (the English translations illustrate the same point).⁶

T1. *Lexical and grammatical members of word classes in Spanish*

lexical			grammatical		
category	example	meaning	category	example	meaning
noun	posesión	possession	pronoun	suyo	his
adjective	rojo	red	pro-adjective	tal	such
verb	existir	exist	auxiliary	estar	be
adverb	atrás	behind	deictic adverb	ahí	there
preposition	tras	behind	gramm. prep.	de	of
conjunction	mientras	while	subordinator	que	that

⁶ Kortmann & König 1992:684 arrange some deverbal English prepositions on a continuum from least to most grammatical.

The subdivision in T1 shows that word classes do not differ in that some are lexical and others are grammatical. Instead, the criterion of lexical vs. grammatical is independent of the word classes and yields two subclasses of each of them. Needless to say, no sharp boundary between these two subclasses is intended.

By consequence, it is not the case that the so-called minor parts of speech have something particularly grammatical about them.⁷ Therefore the transition of, e.g., a relational noun into a preposition or a conjunction does not amount to the grammaticalization of the former, as is so often assumed. If, for instance, a relational noun such as Span. *base* appears in a preposition like *a base (de)* ‘on the basis (of)’, this is often called grammaticalization of the noun *base* to the preposition *a base de*. In reality, however, the appurtenance of any linguistic unit to a word class – preposition in the case at hand – implies first and foremost its appurtenance to the inventory, i.e. to the lexicon.⁸ The genesis of a preposition like *a base (de)* is therefore, first of all, a process of lexicalization of this sequence. Once such new lexical items have been created, they can undergo grammaticalization.

If grammaticalization changes a lexical element of a given category into a grammatical one, then it follows that grammaticalization by itself does not touch the syntactic category or may, at any rate, leave it untouched.⁹ These are cases of grammaticalization which cannot be construed as reanalysis.¹⁰

3.2. Lexicalization and grammaticalization in Spanish

In what follows, we consider the relationship between lexicalization and grammaticalization in three different areas of Spanish, viz. prepositions, conjunctions and verbs with prepositional government.

3.2.1. Complex prepositions

⁷ The category of the particle is especially often associated with the notion of grammatical word or morpheme.

⁸ This insight is already foreshadowed in the *Cours de linguistique générale*. Saussure (1916[1985]:186) writes: ‘on attribue généralement les prépositions à la grammaire; pourtant la locution prépositionnelle *en considération de* est essentiellement lexicologique, puisque le mot *considération* y figure avec son sens propre.’

⁹ ‘Decategorialization’ is presented in Hopper & Traugott 1993, ch. 5.3 as something essential in grammaticalization. However, this should not be interpreted as ‘shift from major to minor category’, but rather as ‘shift from content words to function words’ (Haspelmath 1998:329).

¹⁰ Cf. Haspelmath 1998 for critical discussion.

Just like other languages, Spanish possesses a number of structural types of complex prepositions.¹¹ Of these, only the type illustrated in E10 is of present interest.

E10. Tenemos que venderlo por debajo del precio. ‘We have to sell it under price.’

Prepositions of this structure have been formed since the Old Castilian stage. T2 shows a selection of Spanish prepositions which were complex at the stage of Old Castilian and are mere secondary prepositions nowadays.¹²

T2. *Complex prepositions in Spanish*

form		meaning
Old Castilian	Mod. Castilian	
des de	desde	since
baxo de	bajo	below
a cabo de	cabe	beside
ante de	ante	before
face a	hacia	towards

All the prepositions of the first column of T2 have it in common that their last element is a primary preposition, namely either *de* or *a* (cf. Meyer-Lübke 1899:295-298). Only when the complex changes into a (secondary) simple preposition, the primary preposition may disappear.

We see that the syntactic pattern of the formation of complex prepositions is based on the combination of two components, as schematized in S2.a. The first is a semantically specific expression, to be called semantically specific relator (SSR) in what follows. The second is a PrepP which, in turn, is formed with a primary, thus, semantically empty, preposition. The latter only serves the structural function of forming a PrepP. Once this is guaranteed, the syntactic nature of the SSR and its syntagmatic relation to the PrepP are of secondary importance. In the SSR slot we find prepositions (*de*, *ante*), adverbs (*bajo*) and relational

¹¹ Cf. Raible 1992, ch. I.1 and, in particular, p. 11f on the strength of the subtype *a N de*.

¹² This is, of course, not to say that all Spanish prepositions originate in this way. There are also deadverbial prepositions (on which s. Meyer-Lübke 1899:159-164) such as *dentro* ‘inside’ (< Vulgar Latin *de intro*), which initially govern a direct complement (*dentro la casa* ‘inside the house’), but from the second half of the the 13th century on govern their complement by means of *de* (*dentro de la casa*).

nouns (*face*); and the syntactic relation between the two syntagms varies accordingly. If the SSR is a preposition or a relational noun, we have government; in the other cases, we have modification or apposition. As the example of *des de* (< Vulgar Latin *de ex de*) in T2 shows, this pattern is of old age. Comparison with French *dès* reveals that it goes back to Proto-Romance.

It is probably not too important that the internal syntax be entirely correct here, because the SSR is not meant to function as a compositional constituent of a syntactic construction, but to create a semantically specific form of prepositional subordination. Since language is a goal-directed activity, we may assume that the reanalysis shown in S2.c is already targetted with the formation of those syntactically complex expressions.

S2. Genesis of complex prepositions by reanalysis

a. <i>input</i>	[[SSR [primary prep [NP]] _{PrepP}] _{PrepP}
b. <i>transition</i>	[[SSR primary prep] [NP]] _{PrepP}
c. <i>output</i>	[secondary prep [NP]] _{PrepP}

The reanalysis goes hand in hand with the lexicalization of the complex preposition, since step S2.b subtracts the formation from the rules of syntax.¹³

In the initial phase, the primary preposition was needed for its structural function. After lexicalization, this function is integrated into the complex consisting of the SSR plus preposition. The internal structure of the latter is no longer relevant. It may either be blurred by phonological attrition, as in *cabe* and *hacia*;¹⁴ or the primary preposition may be dropped, as in *bajo*.¹⁵

¹³ Cf. Vincent 1997. He derives Ital. *dopo cena* ‘after dinner’, by the reanalysis formalized in S2 (with Prep instead of SSR in the first position), from Vulgar Latin [*de* [*post cenam*]]. – It should be obvious in general that the univerbation of *bc* in *abc* presupposes a bracketing *a[bc]*, and consequently, in particular, that lexicalizations of the kind observable in T2 presuppose the kind of rebracketing shown in S2. It is not clear how Haspelmath (1998:330-333) avoids this conclusion.

¹⁴ Intervocalic *d* may get lost more generally from the 15th century on; thus *puede* ‘can’ may appear as *pue*.

¹⁵ This development would go in the opposite direction of the one mentioned in fn. 12. In the case at hand, an alternative analysis is possible, as pointed out by M. Haspelmath (p.c.). Namely, the phrase *bajo de NP* never passes through S2. Instead, *bajo* at first governs a prepositional complement; later, government becomes direct, so that *de* is dropped. This analysis, while plausible in the cases at hand, is obviously not available for the cases of T2.

3.2.2. Complex conjunctions

Just like other languages, Spanish possesses a number of structural types of complex conjunctions.¹⁶ Of these, only the type illustrated in E11 is of present interest.

E11. No parece mal que los españoles tengan sus patatas, con tal que nosotros tengamos nuestras papas. ‘It seems o.k. for the Spanish to have their pommes de terre, provided we can have our potatoes.’

T3 shows a mixed selection – this time, at the synchronic level – of complex and other secondary conjunctions of Spanish.

T3. *Secondary conjunctions in Spanish*

form	meaning
en la medida en que (c. ind.)	to the extent that
a pesar de que (c. ind.)	despite the fact that
a no ser que (c. subj.)	lest
después de que (c. ind./subj.)	after
pese a que (c. ind.)	although
con tal que (c. subj.)	provided
siempre que (c. subj.)	as long as
para que (c. subj.)	in order that
aunque (c. ind./subj.)	though
porque (c. ind.)	because

All these conjunctions have it in common that their last element is a primary conjunction, viz. *que*.¹⁷ Only when the complex changes into a (secondary) simple conjunction, such as *como*, ‘as’ and *mientras* ‘while’, *que* disappears. For several of these complex conjunctions, the subordination is syntactically completely regular. For instance, in *a pesar de [que S]*, *después de [que S]*, *pese a [que S]*, we can substitute a concrete NP such as *mis esfuerzos*

¹⁶ See Raible 1992 for a cross-linguistic survey and a universal theory of clause linkage.

¹⁷ Cf. Meyer-Lübke 1899:611f. Meillet 1915 is among the first to analyze the genesis of Romance conjunctions in terms of grammaticalization.

‘my efforts’ for the constituent [*que* S]. This state is represented in S3.a (where SR stands for ‘[generic] subordinator’).

S3. *Genesis of complex conjunctions by reanalysis*

a. <i>input</i>	[[SSR [SR [S]] _{NP}] _{PrepP}] _{AdvP}
b. <i>output</i>	[[sec. conj] [S]] _{AdvP}

We see that the syntactic pattern of the formation of complex conjunctions is based on the combination of a SSR with a subordinate clause which, in turn, is formed with a semantically empty conjunction. The latter only serves the structural function of subordinating the clause. Once this is guaranteed, the syntactic nature of the SSR and its syntagmatic relation to the subordinate clause are of secondary importance. In the SSR slot we find prepositions, adverbs and combinations thereof; and the syntactic relation between the two syntagms varies accordingly.¹⁸ If the SSR ends in a preposition, it governs the subordinate clause, otherwise it bears a modifying or appositive relation to the latter.¹⁹

The conjunction thus developed functionally contains a subordinator and by virtue of this directly takes a clause as its complement. This state is symbolized in S3.b. In this last phase, the subordinator *que*, which remains visible in *porque* and *aunque*, may be suppressed, as in *como* and *mientras*.

At the last stage of the evolution of complex prepositions and conjunctions, the structural element which subordinates the complement and which differs between PrepPs and subordinate clauses, disappears. As a result of this, there are several particles such as *como*, *mientras* which function both as preposition and as conjunction.

Incidentally, the mood to be observed with complex conjunctions cannot be derived from their constitution, but is simply the same which is used in the respective subordinate clause types if they are introduced by a simple conjunction. This is further evidence for the conception that the passage through S3 is a goal-directed process, where new collocations are fitted into a given schema.

¹⁸ Cf. Kortmann 1996 for the various combinations found in European languages and Herlin 1999 for complex temporal conjunctions in Finnish.

¹⁹ The same schema applies, of course, to the formation of conjunctions in other Romance and, mutatis mutandis, numerous further languages. Cf. Harris & Campbell 1995:288 for Romance.

3.2.3. Prepositions governed by verbs

Verbs govern their complements in different syntactic functions. Among these is the PrepP as complement. Just as a verb governs the case of its complement, it may govern the specific preposition of this PrepP. T4 contains some Spanish examples.

T4. *Prepositional government in Spanish*

form	meaning
acabar/terminar por	end with
creer en	believe in
parecerse a	resemble
asombrarse de	wonder at
acabar de	finish
abusar de	abuse

Just as in complex prepositions, those prepositions which are used to govern the complement are exclusively primary prepositions. No Spanish verb governs any of the prepositions of T2. This is, thus, completely parallel to the complex prepositions. We might therefore feel tempted to speak of a lexicalization of verb-preposition collocations in T4. However, the combination here remains discontinuous. Neither is there any evidence for a reanalysis, analogously to the previous cases, whereby the verb would form a constituent together with the governing preposition. Therefore the traditional description is to be preferred, whereby the verb is a lexical unit in itself and determines the occurrence of a particular preposition in its complement. Anyway, this property of the verb is a lexical-grammatical property, just as if the preposition were part of its lexeme.

4. Reanalysis, grammaticalization and lexicalization

The reanalyses shown in S2 and S3 destroy a regular syntactic construction with no compensation at the syntactic level. The reanalysis therefore entails a loss in compositionality. It is the essential step in the lexicalization of complex prepositions and conjunctions. It is true that, in the cases reviewed, a new preposition or conjunction evolves by reanalysis. However, as was said in the beginning, this is not a case of grammaticalization, because the particle thus developed is not a grammatical element. It could be further grammaticalized, such as Latin *de*, *ad*, *in* were grammaticalized to Spanish *de*, *a*, *en*. This process, however, does not involve further reanalysis.

4.1. Grammaticalization vs. lexicalization

Every monomorphemic unit is, by definition, already in the lexicon and therefore cannot be lexicalized. Only complex units may be lexicalized. Again, relatively few morphemes and even fewer complex units are contained in the grammar. Morphemes, and also complex units, may therefore be grammaticalized.

Complex units may be grammaticalized without having been lexicalized. For instance, the combination of a preposition with its governed case, or the combination of a conjunction with a mood, may be grammaticalized. These constellations are usually not analyzed as discontinuous linguistic units (cf., however, Touratier 1979). This would, in fact, presuppose a reanalysis and ensuing lexicalization of the combination.

Grammaticalization involves an analytic access to a unit, lexicalization involves a holistic access to a unit, a renunciation of its internal analysis. Both processes do not concern signs in isolation, but signs in their paradigmatic and syntagmatic relations. However, this is just where the essential differences between the two processes are.

Let $[XY]_Z$ be a complex construction which undergoes grammaticalization or lexicalization. Then the differences between the two processes consist in two aspects. First, in grammaticalization there may be a constituent of Z, e.g. X, which is the focus of the process and which is changed into a grammatical formative by it. In lexicalization, there is no such constituent; the lexicalization affects Z as a whole. From this it follows that lexicalization necessarily concerns an internally complex unit, whereas we may reasonably speak of grammaticalization even with respect to simple units.

Second, in grammaticalization the internal relations of Z become more strict and constrained. This regards, in particular, the relation between X and Y or between X and Z. Again, in lexicalization the internal relations of Z become irregular and get lost.

A consequence of this explication of the notion of lexicalization is that the coalescence of two grammatical morphemes must be called lexicalization. Here are a couple of examples. In Spanish *desde*, appearing in T2, the grammatical prepositions *de*, *ex* and *de* are combined to a new preposition. In English *himself* (with the other inflected forms), the accusative of the personal pronoun is combined with the semigrammatical morpheme *self* to yield the reflexive pronoun. From this perspective, the evolution of German *möchte* into a new lexeme (infinitive *möchten*)²⁰ is a merger of the semigrammatical lexeme *mögen* ‘may’ with the inflectional category of the subjunctive II into a new attenuated volitive meaning.

²⁰ studied in G. Diewald's contribution to a workshop of the 21. Annual Meeting of the DGfS, 1999

4.2. Lexicalization and its converse

What was said on lexicalization may be summarized in T5.

T5. *Lexicalization*

component \ complex unit	lexicon	grammar
access	holistic	analytic
structure	opaque	transparent
function vs. structure	irregular	compositional
process	← lexicalization → folk etymology	

The horizontal axis of T5 is the horizontal axis of S1. Lexicalization is a process constantly at work in ordinary language activity. T5 shows that the inversion of lexicalization is not grammaticalization. Bestowing structure onto a hitherto opaque expression is not an automatic ingredient of language activity, but demands an enhanced measure of creativity. The operation is called folk etymology (cf. Untermann 1975) and is by magnitudes rarer than lexicalization.

A final terminological remark is necessary. The adjective *lexical* has two meanings in linguistics, 1) belonging to the inventory, 2) having a specific, concrete meaning. In the latter sense, *lexical* is opposed to *grammatical*, as displayed in S1. In the former sense, however, both words with a concrete meaning and grammatical formatives belong to the inventory. In particular, the morphemicon (core of the lexicon) contains both the lexical and the grammatical morphemes of a language. Lexicalization is a process in which something becomes lexical in the first of the two senses. The term *idiomaticization* has essentially the same meaning. Lexicalization as a process in which something becomes lexical in the second sense would be the same as degrammaticalization, to which we turn in the following section.

4.3. Grammaticalization and its converse

What was said on grammaticalization may be summarized in T6.

T6. *Grammaticalization*

level	higher		lower
complex unit	sentence	word	stem
function vs. structure	iconic		arbitrary
unit concerned	word	morpheme	feature
manipulation	free		obligatory
meaning	specific		abstract
process	→ grammaticalization ← degrammaticalization		

The horizontal axis of T6 is the vertical axis of S1. Grammaticalization is a process constantly at work in ordinary language activity. T6 shows that the inversion of grammaticalization is not lexicalization. Giving autonomy to a hitherto dependent expression is not an automatic ingredient of language activity, but demands an enhanced measure of creativity. The operation is called degrammaticalization (cf. Ramat 1992) and is by magnitudes rarer than grammaticalization.

5. Summary

Both lexicalization and grammaticalization are reductive processes which constrain the freedom of the speaker in selecting and combining the constituents of a complex expression. Insofar, both processes can be regarded as a transition of an expression from *parole* into *langue*. This is in consonance with the conception of *langue* as the language system whose semantic subsystem consists of the lexicon and the grammar. Lexicalization and grammaticalization are the two janus-faces of the creation of the language system in *parole*, of the *Versprachlichung* of the world.

Grammaticalization and lexicalization are not mirror images, but orthogonal to each other. Both are reduction processes (cf. Lehmann 1989), but in a different sense. Grammaticalization reduces the autonomy of a unit, shifting it to a lower, more strictly regulated grammatical level, more precisely, into the right lower corner of S1. Lexicalization reduces

the inner structure of a unit, shifting it into the inventory, more precisely, into the left lower corner of S1.

While we may reasonably speak of lexicalization only with respect to complex units, grammaticalization concerns a complex unit and may simultaneously affect in particular one of its constituents. The latter then evolves into a (more) grammatical formative. Such a unitary constituent is created by lexicalization to begin with. Insofar, lexicalization plays a role as the first phase, or perhaps rather preparatory phase, of grammaticalization. Again, it is not excluded that lexicalization and grammaticalization occur jointly in a given case. English *wanna* and *gonna* would be cases in point. In the former, the combination of a lexical and a grammatical morpheme lexicalizes to a modal, in the latter, the combination of semi-grammaticalized *going* with a grammatical morpheme is lexicalized and further grammaticalized.

We have seen that prepositions and conjunctions come about not by grammaticalization, but by lexicalization. Once they have come into existence, they may then be grammaticalized. Lexical change, however, is much more ephemeral than grammatical change. From among all the new prepositions and conjunctions, only a fraction is grammaticalized. All the others are abandoned and replaced by other neologisms. Those numerous complex prepositions and conjunctions which constantly come and go do not indicate incomplete grammaticalization processes, but are simply products of lexical change.

References

- Bybee, Joan L. & Perkins, Revere & Pagliuca, William 1994, *The evolution of grammar. Tense, aspect and modality in the languages of the world*. Chicago & London: University of Chicago Press.
- Harris, Alice C. & Campbell, Lyle 1995, *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Haspelmath, Martin 1998, "Does grammaticalization need reanalysis?" *Studies in Language* 22:315-351.
- Herlin, Ilona 1999, The lexicalization of Finnish conjunctions. Helsinki: University of Helsinki (unpublished paper).
- Hopper, Paul J. & Traugott, Elizabeth C. 1993, *Grammaticalization*. Cambridge: Cambridge University Press (Cambridge Textbooks in Linguistics).
- Kortmann, Bernd 1996, *Adverbial subordination. A typology of adverbial subordinators based on European languages*. Berlin & New York: W. de Gruyter.
- Kortmann, Bernd & König, Ekkehard 1992, "Categorial reanalysis: the case of deverbal prepositions." *Linguistics* 30:671-697.

- Lehmann, Christian 1989, "Grammatikalisierung und Lexikalisierung". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42:11-19.
- Lehmann, Christian 1995, "Synsemantika". Jacobs, Joachim et al. (eds.), *Syntax*. Ein internationales Handbuch. Berlin: W. de Gruyter (Handbücher der Sprach- und Kommunikationswissenschaft, 9/II); 1251-1266.
- Lehmann, Christian 1995, *Thoughts on grammaticalization*. Unterschleissheim: LINCOM Europa (LINCOM Studies in Theoretical Linguistics, 1).
- Meillet, Antoine 1915, "Le renouvellement des conjonctions." *Annuaire de l'École Pratique des Hautes Études*, section historique et philologique. Reprint: Meillet, Antoine 1921, *Linguistique historique et linguistique générale*. [Tome I.] Paris: Klincksieck; 159-174.
- Meyer-Lübke, Wilhelm 1899, *Romanische Syntax*. Leipzig: O.R. Reisland (Grammatik der Romanischen Sprachen, 3).
- Moreno Cabrera, Juan Carlos 1998, "On the relationships between grammaticalization and lexicalization." Giacalone Ramat, Anna & Hopper, Paul J. (eds.), *The limits of grammaticalization*. Amsterdam & Philadelphia: J. Benjamins (Typological Studies in Language, 37); 211-227.
- Raible, Wolfgang 1992, *Junktion. Eine Dimension der Sprache und ihre Realisierungsformen zwischen Aggregation und Integration*. Heidelberg: C. Winter (Sitzungsberichte der Heidelberger Akademie der Wissenschaften. Philosophisch-historische Klasse, 1992/2).
- Ramat, Paolo 1992, "Thoughts on degrammaticalization." *Linguistics* 30:549-560.
- Saussure, Ferdinand de 1916[1985], *Cours de linguistique générale*. Publié par Charles Bally et Albert Séchehaye avec la collaboration de Albert Riedlinger. Édition critique préparée par Tullio de Mauro. Paris: Payot.
- Schultze-Berndt, Eva 2000, *Simple and complex verbs in Jaminjung. A study of event categorization in an Australian language*. Nijmegen: Max-Planck-Institute for Psycholinguistics (MPI Series in Psycholinguistics).
- Touratier, Christian 1979, "Accusatif et analyse en morphèmes." *Bulletin de la Société de Linguistique de Paris* 74(1):43-92.
- Untermann, Jürgen 1975, "Etymologie und Wortgeschichte." Seiler, Hansjakob (ed.), *Linguistic workshop III. Arbeiten des Kölner Universalienprojekts 1974*. München: Fink (Structura, 9); 93-116.
- Vincent, Nigel 1997, "On the origin of prepositions". Paper delivered to the Philological Society, Manchester, 21 November 1997.