The relationships between grammaticalization, agglutination, lexicalization and analogy in Latin and other languages[1]

Michèle FRUYT
(Université de Paris IV-Sorbonne)
michele.fruyt@gmail.com

The concept of 'grammaticalization' was created by Antoine Meillet, who coined two French words: the noun Fr. *grammaticalisation* and the reflexive verb Fr. *se grammaticaliser*. He developed the concept in two articles (1912 and 1916) that became, later on, two chapters of his 1921 book. Two of the main topics developed by Meillet were the formation of subordination conjunctions and verbal periphrases in the Romance languages.

Meillet's definition of grammaticalization is « *le passage d'un mot autonome au rôle d'élément grammatical* » (Meillet 1921: 131), which we could update as «a new linguistic element coming into the grammatical area of a given language». Meillet gave to the concept of grammaticalization a smaller extension than contemporary linguists do, or we could say that contemporary linguistics have over-extended the concept created by Meillet.

Therefore it seems useful to define and delimit the extension of the grammaticalization phenomenon, by contrast with other connected phenomena that occur with it or without it. We would like to study here the relationships between grammaticalization, agglutination, lexicalization and analogy. In some cases, they occur at the same time, and in some other cases, they occur separately or in various combinations.

1. AGGLUTINATION

Agglutination is a productive process of word formation that exists in its own right. We will use the word here according to F. de Saussure's terminology. It occurs when several «words» are joined together in their syntagmatic sequence. They then become one single new «word», a single new lexeme or lexical item, which has its own features on the phonetic, morphological, syntactic, semantic and referential levels. We then have a

---

new sign with a new signifier (in the sense of Fr. *signifiant*) and a new signified (in the sense of Fr. *signifié* according to F. de Saussure's terminology).

We will consider here only the agglutination processes that belong to the common base of a given speech community, those that have been institutionalized and conventionalized in the language, those that Bernard Pottier calls a *lexie* and that he defines as any memorized lexical unit (Fr. *toute unité lexicale mémorisée*) or any memorized sequence of words (Fr. *toute séquence mémorisée*), and that we may call a *complexeme*[2].

1.1. The formation of adverbs or prepositions from frozen sequences

Agglutination is usual in the formation of adverbs and prepositions[3], for example temporal adverbs, coordinators, and connectors such as: Engl. *far away, in the distance*, *near by*, *to-gether, there-fore, never-the-less*; Fr. *en même temps*; *de ce fait*; *néanmoins*; Lat. *praeter-eā* « in addition to that », *eō-modo* « in this way », *nihilō-minus* « nevertheless ».

Many subordinators also have their origins in agglutinated sequences: Engl. *as far as.., when-ever ..., who-ever*; Fr. *après que ..., avant que ..., dès lors que..., alors que..., tandis que ..., lors-que..., bien que...*; Lat*. quō-modo, quā-rē, quam-ob-rem, quam-uīs,* as well as many prepositions: Engl. *in front of*, Fr. *en face de*; Engl. *because of,* Fr. *en raison de, à cause de*.

A most interesting category is that of the temporal deictic adverbs meaning « today ». In many Indo-European languages, they result from the agglutination of a nominal syntagm (a noun phrase) containing a deictic adjective and a noun meaning « day »: literally « this day ». This is so common that it is often followed, in the diachrony of the same language, by a cyclic renewal, where the noun « day » is repeated a second time, for example:

- Ancient Greek (Attic) τήμερον « today » from a deictic element and the noun ἡμέρα « day ». A cyclic renewal has occurred in ἡ τήμερον ἡμέρα (Dem.) lit. Engl. « the day which is today, the day of today », lit. Fr. « le jour d'aujourd'hui ».

- Latin *ho-diē* « today » > Old-French *hui* « today », followed then by a cyclic renewal: Fr. *au-jour-d'hui* « today » from *au jour d'(hui).*

---

[2] For the coinage of this English word by Ph. BALDI and P. CUZZOLIN in order to translate Fr. *lexie* or *lexie complexe* (according to Bernard Pottier's terminology), see FRUYT (forthcoming 2011, Berlin, note 11).

[3] Fruyt (1990 : passim).

- German *heute* results from the agglutination and shortening of the sequence of two words, containing the deictic stem *hi-* and the noun « day » in the instrumental case as shown by Old-High-German: O.-High-Germ. **hiu tagu*; Old-High-Germ. *hiutu*, M.-High-Germ. *hiute.*

- The same origin is to be found in Modern English *today* (as shown by Old-English *to daeg* literally « on (this) day »).

This agglutination process can also be illustrated in many other temporal adverbs, such as:

- Engl. *nowadays* « at the present day » from Middle-Engl. *now* (adverb), *a* (preposition) « on, to », and *day* (noun);

- Germ. *heutzutage* « nowadays » from *heute* (adverb « today »), *zu* (preposition "to" »), *tage* (noun « day »),

as well as in many nouns denoting parts of the day. The name for « afternoon » results from the agglutination of a prepositional syntagm or phrase with a shift and trans-categorization into a noun in Engl. *an after-noon* or Fr. *une après-midi.*

Latin shows an agglutinated sequence for the nouns meaning « midday » (literally « the middle of the day » in the locative and ablative case) and « tomorrow » (literally « the next day » also in the locative and ablative case): Lat. *merī-diē* « midday, noon », *postrī-diē* « tomorrow ».

The very long temporal adverbs usually have this kind of origin and they are the result of the agglutination of more than two elements: Fr. *dorénavant* comes from *d'ore en avant*, based on four or five words joined together at various periods of the language during the centuries from Latin to Modern French.

1.2. Agglutination in various morphological categories

The agglutination process is to be found in the formation of all morphological categories.

1.2.1. Verbs

The agglutination process may be responsible for the formation of verbs. The complementizer (in the accusative as a direct object) is incorporated into the initial verb in Lat. *animum aduertere* (*ad* + accusative) « to draw attention to » > *anim-aduertere* (+ acc.). The incorporation of another type of complementizer (in the ablative case with a separative meaning) is documented in *manū mittere* « to free (a slave) > *manū-mittere*. In *uēnum īre* « to go to the sale, to be put on sale » > *uēn-īre* « to be on sale », the first element *uēnum* is probably the directive accusative of a process noun meaning « sale », which is not

documented elsewhere in Latin. Nevertheless, this kind of formation is not usual in Latin for verbs. It is much more usual for other grammatical categories.

It is well documented in several Indo-European languages, e.g. Engl. *turn upside down*. Its French semantic equivalent also belongs to this kind of structure : Fr. *bouleverser* "turn upside down, change completely" is an agglutination of two verbs: *bouler* "turn upside down" and *verser* (same meaning), (see *DHLF* p. 278). Fr. *saupoudrer* "spray salt or sugar on (something)" also belongs to this category, from the noun *sau* "salt" (an old and dialectal variant (XIVth c.) of Mod.-Fr. *sel*) and the verb *poudrer* (a denominativ verb from the noun Fr. *poudre* "powder").

1.2.2. Adverbs

Many enunciative adverbs are due to agglutination, e.g. an infinitive meaning "know" or "see" and a modal verb meaning "it is possible, it is permitted" in Lat. *scīre licet* lit. « it is permitted to know » > *scī-licet* « of course » ; *uidēre licet* lit. « it is permitted to see » > *uide-licet* « of course ».

An interesting group is constituted by the modal adverbs expressing possibility (epistemic modality) and meaning « perhaps »: a usual formation in the I.-E. languages is a freezing of a verbal syntagm with a modal verb expressing possibility and the infinitive of the verb « be », e.g.:
- Engl. *may be* (cut off from the clause *it may be* (*that*)…).

- Fr. *peut-être* (issued from the verbal syntagm *il peut être que …*)*.

Lat. *fors sit an* "may be" followed by the subjunctive (the most usual form being *fortasse*) displays a similar formation, since the first word *fors* means "chance", while the second is the verb "be" in the present subjunctive with a possibility value and *an*, the final word, was probably originally a subordinator that was cut off from the subsequent subordinate clause it originally introduced. The original literal meaning would be "the chance may be that" and therefore "it could happen that".

1.2.3. Free-choice quantifiers

The formation of free-choice quantifiers in the Indo-European languages often involves the freezing of a personal inflected modal verb referring to volition, either in the second person or the third one, e.g. Lat. *quī-uīs, quī-libet* lit. « who-ever you want, who-ever you please» and therefore « anyone» with *uīs* "you want" (from *uolō* "want", 2nd person sg.) and *libet* impersonal verb "it pleases (someone), it is pleasing (for someone)".

1.2.4. Nouns and adjectives

And, of course, usual nouns and adjectives with a lexical meaning are issued from the rigidification of an agglutinated nominal syntagm, where the noun which is the head of the syntagm is modified by an adjective or another noun in the genitive case. In this last situation, the genitive case may still display the productive inflectional genitive ending (e.g. long -*i* in Lat. *agrī cultūra* or *agrīcultūra)* or an archaic and fossilized genitive ending (e.g. *–ās* in Lat. *pater familiās*)*.

This kind of coherent sequence functions as one and the same syntactic and semantic unit with one and the same referential value. This lexical structure is often used in Latin and the Romance languages in order to create denominations of precise entities in the various technical and scientific vocabularies. This lexical organisation is often underlined by some classification concepts based on cognition, reflecting the way the speech community perceives the denoted entities.

Recently, in the geological and industrial domain, was coined the noun Fr. *terre rare* lit. "rare earth" in order to designate a kind of earth, recently discovered, which contains several very precious metals used in electronics. The two words, the noun Fr. *terre* and the adjective *rare* are both very usual – *terre* being a generic and polysemic term -, but their agglutinated sequence forms a new lexical unit with a very specific new denotation.

A similar structure could be mentioned in Fr. *carte orange* (from Fr. *carte* "card" and *orange* "which displays the orange colour"), which denotes a pass for public transportation in Paris. In this recent coinage, two very usual terms combine in a new specific designation.

The Latin noun *crēta* must have denoted a kind of earth whose main property was originally felt as being its whitish colour (fine clay, chalk, etc.). But this word may be associated with several other words with which it constitutes a complexeme (see note 2) denoting various kinds of earth : *crēta figulāris* or *crēta figlīna* "potter's clay", *crēta fullōnia* "fuller's earth", *crēta argentāria* "silversmith's whiting", *crēta sūtōria* "shoemaker's earth", etc. Each of these complexemes denotes a different product (with different chemical properties) and therefore each of them functions semantically as a whole.

1.3. Negation

A good and well-known example of agglutination which finally ends with the complete de-semantization of a constituent is provided by the Latin and French negative lexemes. It is also a grammaticalization case, since negations are grammatical lexemes.

In logical terms, a well-documented pattern is the negation of the scalar minimum, when the speaker denotes the total absence of the smallest element possible. The Latin standard negation *nōn* comes from

the agglutination of the inherited I.-E. negation *ne and a neuter form of the numeral « one ": Lat. *ne oinom "not one" > nōn « "not". The specific Latin negation nihil "nothing" is also the result of the inherited negation *ne plus a noun hīlum meaning "a very small thing" and originally, probably, the curve of a bean: Lat. *ne hīlum "not (even) a tiny thing"> nihil "nothing". Lat. nēmō "nobody" from *ne hemō or *ne homō meant literally "not a single human being", "not one human being". Here, the constituent homo has not been completely de-semanticized, since the word may only be applied to human beings.

Similarly, we have a cyclic renewal in the standard negation in Old-French, which is the association of ne (phonetically derived from the unstressed Latin negation non) and a concrete noun denoting the scalar minimum, a very small step for distance, a very small drop used with the verbs meaning "see": Fr. ne …pas, ne …point, ne …goutte, etc. In reality, we should rather talk of helix[4] renewals (Fr. renouvellement en spirale) as A. Meillet did (1921: 140), since the renewal shows a lexical or morphological re-encoding, while the structure stays the same.

Other languages also provide negative lexemes formed from the agglutination of the numeral "one" behind a previous negation: Engl. no one < *no one "not one"; Ancient Greek οὐδείς "nobody", lit. "not one" (numeral "one" in the masculine), οὐδέν "nothing" lit. "not one" (numeral "one" in the neuter).

In Engl. nobody, we have the same kind of structure with a semantic change for the noun body, which became a quantifier. It is not totally de-semanticized, since it still contains the semantic feature "human being", but it has lost its specific semantic features.

The reinforced negation in contemporary English "No way!" "absolutely not!" also shows a semantic change for the noun way and probably a strong de-semanticization. It could, one day, become a standard negation in English.

1.4. Verbal periphrasis

The constitution of verbal periphrases involves an agglutination and also a certain kind of grammaticalization called "morphologization", e.g.:

- in the Latin future passive infinitive : amātum īrī, lectum īrī (supine plus passive infinitive of īre "go");

- in the Latin sequence "infinitive + habeō", that gave the future tense in most Romance languages;

---

[4] Fr. spirale may be translated as Engl. helix : Fruyt (forthcoming).

- in "*habeō* + past passive participle" as a past periphrasis in the Romance languages. It had at first a perfective, resultative meaning[5], then - as the degree of grammaticalization became complete - it took on an aoristic past meaning.

This last evolution is documented in colloquial French in the "*passé composé*", which has become the temporal equivalent of the "*passé simple*" (Fr. *j'ai vu* instead of *je vis*).

There are various speeds for the complete grammaticalization into an aoristic value. While it is already completed in Modern spoken French, it is not yet in Spanish. In Medieval-Spanish (M.-F. Delport 2004), we do have the periphrasis " 'have' + past passive participle", but it still denotes the present consequences of an accomplished action (the present result of a passed action). We have the same perfective value in Late Latin, but no periphrasis yet.

There are various stages before a periphrasis arrives at a full and complete degree of grammaticalization. This precise case confirms the tendency according to which French would be the Romance language that showed the most rapid development, while Italian was slower and Spanish the slowest.

2. LEXICALIZATION

Lexicalization is a special case of freezing: a morphological freezing. It is often followed by the entry into the lexicon of a new element, either a lexeme or a lexeme constituent.

A common situation of lexicalization in the inflectional languages is shown when a given word, inflected in a certain case, gender and number, looses the possibility of morpho-syntactic variation and becomes morphologically frozen. The frozen form itself may become the signifier of a new lexeme, generally an adverb, since we generally categorise as adverbs most of the lexemes that are not liable to inflection.

This is well known in the I.-E. languages: e.g. the process must have occurred in the Latin adverbial type *bene, longē*, since this adverbial morpheme's long *–e* is probably an old instrumental singular ending. The Latin adverb *temper-ī* "at the right time" (cf. the same formation for Fr. *à temps*) is probably the freezing and lexicalization of a locative case in the singular (with long *–i*) of *tempus, -oris* Nt "time".

Thus we see that lexicalization can preserve linguistic forms that have otherwise become obsolete, as in Lat. *pater familiās*, where there is still documented the no longer productive genitive sg. inflectional ending -*as*. The long *–e* in *longē* was maintained in this adverb, whereas it no longer functioned as a productive ending in the Latin thematic declension.

---

[5] We use *perfective* and *aoristic* here in order to refer to the values documented, respectively, by the Ancient Greek perfect and aorist.

In the same way, Lat. *uel* "or", a usual coordinator, comes from the morphological freezing of a form of the verb "want": *uolō* "want" in the 2nd pers. sg. of the athematic[6] present indicative: *uel-si* "you want" > *uelli* > *uell* > *uel* or in the imperative 2nd pers. sg. (*uel*).

The various lexicalization examples we have just mentioned do not have the same relationship with grammaticalization. The coordinator *uel* is a grammaticalization case (mentioned by Meillet himself), since the new item is a grammatical lexeme. The lexical item *uel* as a coordinator is a more grammatical lexeme than *uel* as a form of the verb *uolō* "want" (as is shown by the frequency list of the LASLA, Delatte *et alii* 1981). The coordinator shows a decrease in its degree of autonomy and a decrease in its semantic value. But A. Meillet probably exaggerates when he sees a total de-semantization of this word, since the coordinator *uel* still keeps some kind of meaning (being used for inclusive alternative), while opposed to Lat. *aut* "or" (being used for exclusive alternative) or to the additive coordinators: Lat. *et, -que, atque* "and".

On the other hand, the first forms we mentioned, the adverbs *bene, longē, temperī,* illustrate instances of lexicalization without being cases of grammaticalization, since these lexemes have a lexical meaning and they are not grammatical lexemes.

These various relationships between lexicalization and grammaticalization show that there are many combinations possible between grammaticalization and the connected phenomena we are studying here.

3. VARIOUS COMBINATIONS

3.1. The combination of lexicalization and agglutination

The kind of lexicalization we have just described may often be found combined, in the same occurrence, with an agglutination process. It often occurs in the formation of adverbs in an inflectional language.

As we have seen, there is both agglutination and morphological freezing in all the words we have mentioned meaning "today". It is also the case for all the subordinators we have quoted in the agglutination process (Lat. *quā-rē, quō-modo, quam-ob-rem.)*

This situation is also to be seen outside of the adverbial category. A good example in Latin is the series of temporal lexemes: in the Latin adjectives *nu-diūs-tertius*, *nu-diūs-quartus* literally "it is today the third day", "it is today the fourth day", that is to say "two days ago", "three days ago", we have three elements linked together. In these words, the last element –*tertius, -quartus* is an ordinal number and it is the variable part of the word. But *nu-* is frozen and even demotivated, and *diūs* is an ancient form of the word for "day" which was already no longer in use**.**

---

[6] For the athematic present stem: cf. 3rd sg. *uul-t,* 2nd pl. *uul-tis,* present infinitive *uel-le.*

3.2. The combination of agglutination and grammaticalization

In fact, any case of agglutination could be considered as a freezing, and more precisely as a kind of freezing applied to a syntagmatic unit. It is not a morphological freezing, but a syntagmatic freezing.

There are some cases of grammaticalization that do not involve an agglutination, but only a lexicalization (as we will see below with Lat. *licet, uel*). Conversely, any kind of agglutination is not necessarily a grammaticalization case. Grammaticalization needs a grammatical element to be involved.

Terms like Engl. *social security*[7]*,* Fr*. carte bleue, sécurité sociale* show an agglutination of several words into a new lexeme, so that the first element Engl. *social* or second element Fr. *bleue, sociale* are downgraded into constituents. But they are not cases of grammaticalization, since neither the whole lexeme nor its constituents are grammatical elements.

Some of the other examples we have looked at above are not so obviously cases of grammaticalization. In some cases, they could also be considered as simple agglutinations, e.g. in the Engl. *once upon a time*, Fr. *dorénavant* (< *d'or-en-avant*), depending on whether these adverbs or their constituents may or may not be considered as grammatical elements.

Subordinators and prepositions are usually considered as grammatical lexemes, since they are relators. Therefore, the agglutinated sequences (subordinators) Lat. *quā-rē, quō-modo, quam-ob-rem*, (prepositions) Engl. *in front of*, Fr. *en face de* may also be considered as grammaticalization cases. Finally, the Latin subordinators *quārē, quō-modo, quam-ob-rem* display an accumulation of our three phenomena: agglutination, lexicalization, grammaticalization. The prepositions Engl. *in front of*, Fr. *en face de* are grammaticalization cases based on an agglutination and a semantic weakening for the nouns Engl. *front* and Fr. *face,* which originally denote body parts. These nouns have undergone a loss of some of their semantic features, but not of all of them: actually, this kind of preposition is a lexico-grammatical lexeme. There are grammatical lexemes that have a grammatical function as relators, but

---

[7] The lexical item *social security* is easily recognized as the agglutinated sequence of a two word syntagmatic sequence (since the first element is an adjective), while some linguists might consider *oyster card* and *visa card* as compounds (rather than agglutinations) since the first element is a noun. But English is a difficult language in which to distinguish between nominal compounds and nominal agglutinations. It seems that the language user does not usually perceive a difference between the two formations (except in some special cases like *hay fever* and *hay-fever*). The complex nouns *visa card, oyster card* were recently coined by analogy on a pre-existing, highly productive pattern. In this pattern, synchronically nominal compounding and agglutination have merged.

that also have concrete semantic features and express a specific relationship between two entities.

3.4. The combination of lexicalization and grammaticalization

Not all grammaticalization cases involve lexicalization and, conversely, not all lexicalization cases are cases of grammaticalization.

3.4.1. Some lexicalizations also involve grammaticalization cases. Lat. *aduersus* or *aduersum* (+ accusative) functions as a spatial preposition meaning "towards, in the direction of": it derives from the morphological freezing in the nominative or accusative singular (masculine) of the past passive participle in **-to-* (Lat. *–tus, -a, -um*) of the motion verb *aduertere* "to turn towards" or medio-passive *aduertī* "to turn oneself towards". The original meaning of *aduersus* (nominative) or *aduersum* (accusative) as a participle was "turned in the direction of" with an agreement with the subject or the object of the verb. Since a preposition is a grammatical lexeme, while a participle of a motion verb is not, this lexicalization is also a case of grammaticalization.

Latin *licet* underwent a trans-categorization from a modal verb "it is permitted, allowed" (expressing possibility and permission) to a usual concessive subordinating conjunction *licet* + subj. "although" in Late Latin. Loss of autonomy here is a good criterion for this change. The modal verb had more autonomy than the subordinator as far as its morphology, syntax and semantics were concerned, whereas the subordinator is used only in very precise syntactic and semantic environments and it is no longer subject to inflection.

In the same way, as we have seen above, the Latin coordinator *uel* "or" mentioned by Meillet (1921: 169 and 1912) is the result of a lexicalization and grammaticalization of a form of the verb "want", without any agglutination.

3.4.2. But the situation is more ambiguous for some other examples of lexicalization, since it is difficult to draw a clear-cut borderline or boundary between grammatical and non grammatical items. Let us take the example of the temporal adverb Lat. *mānī* or *māne* "early in the day, in the morning". It comes from the lexicalization in the ablative singular of the archaic adjective *mānis, -e* "good" and follows, therefore, a semantic evolution well documented in other I.-E. languages such as French : Fr. *de bonne heure, de bon matin* "early in the morning" with the adjective Fr. *bon* "good". Here is a semantic evolution from an evaluative adjective (which is already a specific kind of adjective) to a special kind of adverb, since "in the morning" could be considered as a temporal quantifier: so is this a grammaticalization case or not?

3.4.3. Finally, some other examples may be considered as simple lexicalizations without grammaticalization, since neither the final whole

term nor its constituents are grammatical elements, and since they involve only lexical terms. The Latin long -o adverbs are derived from the morphological freezing of the ablative singular of thematic adjectives: the lexicalization of *continuō* "in a continuous way" (adverb) compared to *continuus, -a, -um* "continuous" (adjective) may not be considered as a grammaticalization case, since neither the adjective nor the adverb are grammatical elements.

3.4.4. Thus grammaticalization may or may not co-occur with agglutination and lexicalization. The cases may be more or less adequate to fulfil the usual criteria and parameters of grammaticalization that linguists usually work with. Some cases are marginal, some others are almost perfect.

The prototypical type of grammaticalization is represented by Old-Fr. *beau-coup* (analyzed by Marchello-Nizia 2006, p.142-148). In Old-French, the starting point was a sentence where *beau coup* was a true nominal syntagm, a free noun phrase containing two words with a lexical meaning, in sentences such as: *le chevalier frappa un beau coup* "the knight made a good stroke". Fr. *coup* is a concrete noun meaning "stroke" and Fr. *beau* an adjective meaning "beautiful, good". But the whole agglutinated sequence then becomes a quantifier: Fr. *Il a beaucoup marché*; *il y a beaucoup de livres.*

In this word, we find almost all the parameters of grammaticalization:

a) the agglutination of two words into one single lexeme;
b) the downgrading of the <u>two</u> original words (and not only of one of them);
c) the trans-categorization of the sequence into a new category: from a nominal syntagm (a free syntactic noun phrase) to an adverb;
d) the entrance into the grammatical area of the language: from a lexical meaning to a grammatical meaning and function. From two lexical words, the whole resulting lexeme has become a grammatical (or lexico-grammatical) word and more precisely a quantifier;
e) the loss of semantic features for the two original words: from a specific context to a non specific context;
f) the loss of the morpheme boundary and, later on, de-motivation.
Only the phonetic melting of the sequence has not yet occurred.

Latin and Old-French negations (Lat. *nōn* and Old-Fr. *ne ... pas, point, mie, goutte*) are usually considered as the best illustrations of grammaticalization, since they were mentioned by Meillet. But they are not as prototypical as Fr. *beaucoup*. Their starting point is a standard negation and their arrival point is also a standard negation. Therefore the whole operation is only the incorporation of an element ("one" or a concrete noun) into a pre-existing negation. It is a formal reinforcement on the signifier level of a pre-existing lexeme, which keeps the same syntactic function. On the other hand, in Fr. *beau-coup* we see the

coalescence of two elements into a new one, which does not have the same syntactic function as either of the two original elements.

But Latin and French negations are still good examples of grammaticalization, since :

a) both the starting point and the arrival point are grammatical lexemes;

b) the core of the grammaticalization process is the incorporated element "one" or the noun meaning "step", "drop", etc. This element has indeed been grammaticalized, since it has lost its status as a free lexical lexeme and has become a constituent part. It has also lost all its semantic features in Modern French (although it had not yet lost them in O.-Fr.).

4. ANALOGY

Another concept we should look at in the context of grammaticalization is analogy.

4.1. Analogy according to F. de Saussure

According to F. de Saussure, analogy is a fundamental linguistic concept and is responsible for the systemic organisation of any given language. As we can see in the following texts, Saussure considers as fundamental the linguistic consciousness and perception of the speech community and of the individual language user:

La conscience du sujet parlant selon F. de Saussure :
- *ELG* 192 : "<Grand principe> : Ce qui est réel dans un état donné du langage, c'est ce dont les sujets parlants ont conscience à un degré quelconque ; tout ce dont ils ont conscience, et rien que ce dont ils peuvent avoir conscience".
- *CLG/E* 2759 et suivantes : I R 2.65-66 : "Entre l'analyse subjective des sujets parlants eux-mêmes (qui seule importe!) et l'analyse objective des grammairiens, il n'y a donc aucune correspondance, quoiqu'elles soient fondées toutes deux en définitive sur la même méthode (confrontation de séries)".
- *Le troisième cours* 275 : "Pour la masse parlante, la perspective où se présentent les termes, c'est la réalité. Ce n'est pas un fantôme, une ombre. D'un autre côté, le linguiste doit, s'il veut comprendre un état de langue, se mettre lui-même dans cette perspective et abandonner la perspective diachronique ou historique qui sera pour lui une gêne, un empêchement. La perspective verticale ou diachronique ne concerne que le linguiste".

In these texts, Saussure speaks of the basic process of analogy as the combination of the consciousness of the language user and the

systemic organisation of a given language. A partial translation of the following texts could be: analogy is the manifestation "of an intelligent transformation"; "the association occurred in the brain of one and the same individual and only one quarter of a second was necessary to draw" the analogical form from the original one; "when new forms appear, everything happens ... by de-combining existing forms and re-combining other forms with the material provided by the first ones".

So, as shown in the following texts, analysis and re-analysis are intelligent operations made by the human brain and based on the perception of certain structural relationships (using certain abstractions made from the spoken language in order to reach an underlying structural level):

- F. de Saussure 2005 : ELG 161 : "et une langue quelconque à un moment quelconque n'est pas autre chose qu'un **vaste enchevêtrement de formations analogiques**, les unes absolument récentes, les autres remontant si haut qu'on ne peut que les deviner. <**Demander à un linguiste de> citer des formations analogiques, c'est donc comme si l'on demandait à un minéralogiste de citer des minéraux**, ou à un astronome de citer quelques étoiles, je commence par le dire pour qu'il n'y ait aucune méprise sur la valeur que nous attribuons à ces faits : ce ne sont pas des faits exceptionnels <et anecdotiques>, ce ne sont pas des *curiosités* <ou des anomalies>, mais c'est la substance la plus claire du langage partout et à toute époque, c'est son histoire de tous les jours et de tous les temps".
- L'analogie (p. 160) comme "phénomène de **transformation intelligente**".
- *ELG* 189 : "le fait s'est passé entre formes on ne peut plus contemporaines, puisque **l'association s'est faite dans le cerveau du même individu**, et qu'il n'a pas fallu même qu'**un quart de seconde** pour conclure de *bél-essi* à *thêr-essi*". *(extension de la finale grecque –essi au datif pluriel issue d'une mécoupure)*
- *ELG* 189 : "nous retrouvons la condition primordiale de toute opération morphologique. Elle porte sur la diversité ou sur le rapport des formes simultanées".
- *ELG* 191 : "quand les formes nouvelles surgissent, tout se passe...par décomposition des formes existantes et recomposition d'autres formes au moyen de matériaux fournis par les premières"[8].

---

[8] Saussure finally proposes the methodological tool of the "linguistic square" (a basic figure with four terms), a certain kind of relationship of similarity: "Le carré linguistique": *ELG* 228: "Toutes considérations possibles sur un fait linguistique sont immédiatement enfermées en une figure simple et partout la même, comprenant quatre termes ».

As we can see, Saussure already had a cognitive approach to analogy and systemic perception of the structure of a given language. He mentions the human brain of a given speaker as the *locus* of these innovations, and he mentions the mental operation as being very rapid (a quarter of a second), which probably means that he conceived these operations as being situated quite "deep" in the areas of the memory, perhaps in a similar way to our automatic reactions.

4.2. Analogy as a standardisation factor

Analogy is also a factor of standardisation of the paradigms of a given language

We all know of analogical forms created by young children or non-native speakers. Instead of the irregular form in use in the speech community, they create a standardized form at the "discourse" level, but one which does not exist in the institutional language. In so doing, they show their need for regular paradigms; they make the paradigms more regular.

In one of his chapters on grammaticalization, Meillet mentions this creative analogical process by children in the creation of Fr. *vous disez* instead of *vous dîtes* (by analogy with *nous rendons, vous rendez* and *nous disons, ils disent*):

> Meillet 1921: 130: chapter "L'évolution des formes grammaticales" : "Les procédés par lesquels se constituent les formes grammaticales sont au nombre de deux.... L'un de ces procédés est l'analogie ; il consiste à faire une forme sur le modèle d'une autre .... soit par exemple les types français : *nous finissons, vous finissez, ils finissent ; nous rendons, vous rendez, ils rendent ; nous lisons, vous lisez, ils lisent* ; sur *nous disons, ils disent,* l'enfant qui apprend à parler est conduit à former *vous disez* sans avoir jamais entendu pareille forme : c'est une forme dite *analogique*. Toutes les formes régulières de la langue peuvent être qualifiées d'analogiques ; car elles sont faites sur des modèles existants, et c'est en vertu du système grammatical de la langue qu'elles sont recréées, chaque fois qu'on en a besoin. .... la forme obtenue par le fonctionnement du système grammatical reproduit le plus souvent une forme déjà entendue et enregistrée dans la mémoire.....
>
> (p. 131) L'autre procédé consiste dans le passage d'un mot autonome au rôle d'élément grammatical. Par exemple *suis* est un mot autonome dans la phrase ... *je suis celui qui suis*, et a encore une certaine autonomie dans une phrase telle que : *je suis chez moi* ; mais il n'est presque plus qu'un élément grammatical dans : *je suis malade, je suis maudit*, et il n'est tout à fait qu'un élément grammatical dans : *je suis*

parti, je suis allé, je me suis promené ... où ce qu'on appelle
... l'auxiliaire n'est qu'une partie d'une forme grammaticale
complexe exprimant le passé..... *suis* est devenu une partie
constituante d'une forme grammaticale. Ces deux procédés,
l'innovation analogique et l'attribution du caractère
grammatical à un mot jadis autonome, sont les seuls par
lesquels se constituent des formes grammaticales nouvelles.

    (p. 133) Tandis que l'analogie peut renouveler le détail
des formes, mais laisse le plus souvent intact le plan
d'ensemble du système existant, la "grammaticalisation" de
certains mots crée des formes neuves, introduit des
catégories qui n'avaient pas d'expression linguistique,
transforme l'ensemble du système".

As we know, the effect of analogy is usually an increase in
grammatical regularity and a decrease or the elimination of grammatical
irregularity. It often results in the alignment of "abnormal" forms on
"normal" ones, i.e. on those which are perceived as "regular" by the
speaker.

Latin morphology provides a great number of instances of such
analogical alignments, since Latin has a strong tendency to develop
regular paradigms. In the third declension of nouns, there is a tendency to
have the same number of syllables (isosyllabism) for all the inflected
forms of the same lexeme: for example, the nominative sg. *bōs* "cow",
with only one syllable, was felt as irregular compared to the rest of the
forms, which had two syllables (*bouis* genitive sg.). Therefore there was
an attempt at regularization with the creation of a new analogical and
regular nominative with two syllables: *bouis*. The same thing happened to
the nominative *Iūpiter* (the god Jupiter), which seemed irregular
compared to the rest of the forms: accusative *Iouem,* genitive *Iouis,*
dative *Iouī,* ablative *Ioue.* Therefore we find an occurrence of a new
"regular", analogical nominative: *Iouis* (Enn. *Ann*. 64).

Lat. *iter, itineris* Nt "path, way, travel" is irregular for the number of
syllables and the alternation between *r* and *n*. Therefore two kinds of
attempts were made in order to create regular forms: either the creation
of a new regular genitive analogical to the nominative or, conversely, the
creation of a new nominative analogical to the genitive. From the
nominative *iter* was created the genitive *iter-is* (Naeu. *Tr*. 38; Acc. *Tr*.
627) and from the genitive *itiner-is* was created the nominative *itiner* (Pl.
Merc. 913, Lucr. 6,339).

The Latin intensifier *ipse* comes from the agglutination of the
endophoric *is* and the reinforcing particle *pse*. In Archaic Latin, we already
find the inflectional morpheme at the end of the word: *ips-e, ips-a, ips-
um*, etc. But we have traces of the previous situation, where the inflection
was applied to the first word: in the feminine *ea-pse*. The fact that *ea-pse*
was transformed into *ips-a* with a shift of the inflectional ending towards
the end of the word is a regularization at a time when the internal
morpheme boundary was lost and the formation of the sequence was de-

motivated. This re-analysis into *ips-a* is based on analogy and on the pre-existing knowledge of the speech community, due to the influence of the memorized prototypical situation where the inflectional ending was always at the end of the word.

4.3. Various relationships between analogy and grammaticalization

The above example of *ipse* could be considered as a grammaticalization case, since *ipse,* as an intensifier, is a grammatical lexeme and since its two constituents are also grammatical lexemes. Therefore, we have here both grammaticalization and analogy. But in the previous examples, the words for "cow", "Jupiter", "path" display analogical changes without grammaticalization, since they have a lexical meaning and not a grammatical function.

When A. Meillet "discovered" the concept of grammaticalization, he opposed it to 'analogy' as a concept with quite different consequences on the evolution of languages. He thought that grammaticalization was responsible for major changes in languages, while analogy was only responsible for small changes (such as the one-off replacement of Fr. *vous dîtes* by *vous disez)*. We prefer to follow Saussure in this, and we consider analogy as a fundamental process in language development, a strong creative and innovative factor for change.

4.4. Pre-existing patterns and analogy

Analogy is the transfer of a pre-existing productive pattern to a new situation. It is the extension of that pattern and it is based on the perception and analysis of the linguistic architecture by the speech community. The pattern according to which the new analogical data is created is already productive in the speech community. The transfer is based on some kind of similarity between two elements or two situations. It is the result of a certain kind of "reasoning" and intellectual activity by the speech community: analogy is the result of the infrastructure of the language as it is perceived by the language user. Analogy is, therefore, under the control of the speech community. This is an important fact that distinguishes analogy from some grammaticalization cases.

5. TWO KINDS OF GRAMMATICALIZATION: WITH AND WITHOUT ANALOGY

This conception of analogy allows us to distinguish between two kinds of grammaticalization: that which involves an analogical process and that which does not. We must therefore distinguish between the

innovations that are made according to a pre-existing pattern and those that are made without such a pattern.

5.1. Grammaticalization may interact with analogy

A "morphologization" is a certain type of grammaticalization that happens when a new element enters a pre-existing morphological category and takes on its morphological features.

For example, in Latin, the inflectional paradigm of the third person singular reflexive pronoun *se* was created according to that of the personal pronouns of the first and second person singular: *mē, tē.* In Latin, the reflexive pronoun *sē (sibi)* has the same kind of inflection as the personal pronouns *ego (mē, mihi)* and *tū (tē, tibi)* in the accusative and dative[9]. Therefore we have a system with three parallel paradigms, and the paradigm of *sē* is analogical to that of *mē* and *tē*. The reflexive paradigm *sē* was thus created in Latin itself according to the pattern displayed by the 1st and 2nd person. This is based on analogy and it is a case of paradigmatization, which is a sub-group of grammaticalization, since *sē* here has entered a pre-existing grammatical system. This is not the entry into the lexicon of the language (lexicalization), but the entry into its morphology (morphologization).

In this case, the innovation is not random. On the contrary, it is a motivated reproduction of a pre-existing pattern. This kind of innovation pre-supposes the existence of a linguistic consciousness in the speech community in a certain area of the language's structure: the consciousness of a systemic organisation and how to extend it. The speech community may not be aware of it, but the innovation is due to an analysis or re-analysis, and there is an element of human thinking, intellectual activity and reasoning in this innovation.

We would like here to mention incidentally that the most conscious creations are voluntary ones such as Engl. *social security*, Fr. *sécurité sociale*, Fr. *carte orange*. They have been created in various circumstances for various needs by one or several speakers, just as Engl. *rail-way*, Fr. *chemin de fer* were created in the 19th century by engineers who needed a new name for the denotation of a new extralinguistic referent (*designatum*). The creation of these lexemes definitively involved active control by the speech community. But these words are not grammatical elements.

5.2. Grammaticalization without analogy

We also speak of "grammaticalization" for processes that do not involve any reasoning or intellectual activity for the speaker and the

---

[9] It is only lacking a nominative.

members of the speech community, and which are not under the control of the speech community. Many grammaticalization cases are only the agglutination of a sequence of words *in situ*, such as the formation of Fr. *au-jour-d'hui,* Engl. *in front of*, Fr. *en face de*. It is also the case for Latin and Old-French standard or specific negations (Lat. *nōn, nēmō, nihil,* Fr. *ne ... pas*): their formation involves a phenomenon over which the speech community did not have any control. The new periphrases that entered the verbal paradigm of a given language at a given time did not result from a decision of the speech community, but were induced slowly by common use all along the evolution of the language over several centuries. These cases are only the rigidification of a certain structure, linked to a certain frequency in the spoken language of the whole community[10].

Nevertheless, when a form agglutinated *in situ* arises, it may be immediately submitted to human intellectual activity and integrated into the pre-existing system, as shown by the Latin modal verb *nequeō* "I can not". We first have an agglutination *in situ* without any human intellectual intervention, since we start from the impersonal passive **neque ītur* "it does not go" (with the verb *īre* "go"), which becomes one single word: *nequītur* "it is not possible". Then the human linguistic intervention leads to the re-interpretation of this word as a form of a paradigm. Since an isolated form deprived of a paradigm is an abnormal and irregular situation, a normal and regular situation was created with a completely new paradigm: *nequ-eō* "I cannot" (infinitive *nequ-īre*).

Another modal verb also comes from the agglutination *in situ* of a sequence containing a negation which underwent a re-analysis: *necesse est* "it is necessary to" from a sequence meaning "there is no escape": Lat. **necessis est* "there is no escape to" > **necessest* ==> a re-interpretation as: *necesse est* "it is necessary".

The Latin verb *rē-fert* has the same kind of origin, with successively both a freezing *in situ* and a re-analysis: Lat. *meā rē fert* lit. "it goes according to my interest" == > *meā rē-fert* "it is important for me". In *meā rē-fert*, we have the re-interpretation of *rē* ablative sg. of the noun *rēs* "thing" as a kind of preverb *rē-* with a long *e*, associated with the productive preverb *re-* with a short *e*.

6. CONCLUSION

There are thus several kinds of processes to which we could apply Meillet's definition of grammaticalization as: "a new linguistic element coming into the grammatical area of a given language", a proto-typical example of which is Fr. *beau-coup,* which shows all the parameters of grammaticalization. There are also several kinds of units that could be considered as grammatical elements and as the *locus* of a grammaticalization process: a lexeme or lexical unit (as in Fr. *beau-coup*),

---

[10] In a comparison with geology, this would be a kind of sedimentation.

a lexeme form in a paradigm (in the verbal periphrases), a lexeme constituent as "one" in the Latin negation.

In any event, a contrastive study with analogy, based on pre-existing patterns, allows us to distinguish two major kinds of grammaticalization:

a) those that are under the control of the speech community and which involve a perception and analysis of linguistic data by the language users;

b) and those that happened independently *in situ* in the syntagmatic chain and without the speech community noticing them or acting upon them. Therefore, we may be wrong when we see grammaticalization as one and the same phenomenon.

This shows us the necessity of relativizing our use of the word and concept of *grammaticalization* and of seeing it in the context of the series of other linguistic processes which may or may not be instrumental in its manifestations. Agglutination, lexicalization and semantic change (and weakening) are more general than grammaticalization, since they occur in all areas, grammatical and non-grammatical ones. They are the primary phenomena: they are mechanisms, while grammaticalization is a selection of data inside a certain area. And there are many ways of forming linguistic elements inside this area. Analogy is one of them. Since analogy is the mirror of the infrastructure of the language, it is much more general than grammaticalization and it displays an intrinsic qualitative difference from all the other phenomena we have studied here. Therefore the grammatical elements formed by analogy should be set apart from the other types of formations.

REFERENCES

BÉGUELIN, Marie-José, to be published, « La problématique des identités diachroniques selon Ferdinand de Saussure », in : Proceedings of the colloquium *Le changement linguistique et ses théorie*s (18-21.2.2008, Programme doctoral CUSO, Univ. of Freiburg, Switzerland).

BERENDONNER, Alain, to be published: « L'opposition lexical / grammatical est-elle opératoire? », in : Proceedings of the colloquium *Le changement linguistique et ses théories* (18-21.2.2008, Programme doctoral CUSO, Univ. of Freiburg, Switzerland).

DELATTE Louis & EVRARD Etienne & GOVAERTS Simone & DENOOZ, Joseph, 1981, *Dictionnaire fréquentiel et Index inverse de la langue latine*, Liège, LASLA.

DELPORT, Marie-France, 2004, *Deux verbes espagnols :* haber *et* tener*, Paris, Editions hispaniques.

DHLF : *Dictionnaire historique de la langue française, Le Robert*, A. Rey (ed.), Paris, Dictionnaires Le Robert, 2000.

FRUYT, Michèle, 1990, « La formation des mots par agglutination en latin »*, Bulletin de la Société de Linguistique de Paris* 85, 1990, 173-209.

FRUYT, Michèle, 1994: « Lexique et conscience linguistique: la motivation», in : D. Conso & N. Fick & B. Poulle (eds.), *Mélanges François Kerlouegan,* Besançon, Annales littéraires de l'Université de Besançon (515), 1994, 255-267.

FRUYT, Michèle, 1998: « La grammaticalisation en latin », in : B. Garcia-Hernandez (ed.), *Estudios de Lingüistica Latina, Actas del IX Coloquio Internacional de Lingüistica Latina,* Madrid, Ediciones clasicas, 877-890 (9th International Colloquium on Latin Linguistics, Madrid, Autonoma, 14-18 April 1997).

FRUYT, Michèle, 1998: « Le renouvellement dans l'évolution linguistique: quelques faits latins », in : B. Bureau & Ch. Nicolas (eds.),*Moussyllanea, Mélanges de linguistique et de littérature anciennes offerts à Claude Moussy*, Louvain-Paris, Peeters, 77-87.

FRUYT, Michèle, 2004: « Somes Cases of Grammaticalisation in Latin: Negation, Correlation and Modal Lexemes »*,* in: G. Haverling (ed.), *Classica et Mediaevalia* 55, Dec. 2004, Aarhus (Danemark),.

FRUYT, Michèle, 2005: « Grammaticalization and Latin » (ICHL, 16th International Conference on Historical Linguistics, Copenhagen, 11-15 Aug. 2003); in : M. Fortescue & E. Skafte Jensen & J. Mogensen & L. Schoesler (eds.), *Historical Linguistics* 2003, Amsterdam, Benjamins.

FRUYT, Michèle, 2008-a (in collaboration with A. Orlandini), « Some Cases of Grammaticalization in the Verb in Late Latin », in : R. Wright (ed.), *Latin vulgaire, latin tardif VIII*, *Actes du VIIIème colloque international sur le latin vulgaire et tardif*, Oxford, 6-9 Sept. 2006, 230-237.

FRUYT, Michèle, 2008-b: « Adverbes latins, grammaticalisation et lexicalisation »*,* in : M. Fruyt & S. Van Laer (eds.), *Adverbes et évolution linguistique en latin*, (Colloquium *Les adverbes latins*, Centre A.Ernout, Paris IV-Sorbonne, 6-8 June 2004), Paris, L'Harmattan, 49-66.

FRUYT, Michèle, forthcoming : « Grammaticalization in Latin », in : Ph. Baldi & P. Cuzzolin (eds.), *New Perspectives on Historical Syntax*, vol. 4 (2011), Berlin, Mouton de Gruyter.

Giacalone-Ramat, Anna, 1998, « Testing the Boundaries of Grammaticalization », in : A. Giacalone-Ramat & P. Hopper (eds.)*The Limits of Grammaticalisation*, , Amsterdam, Benjamins, 107-127.

Lehmann, Christian, 1985, « Grammaticalization: Synchronic Variation and Diachronic Change », *Lingua e stile* 20, 303-318.

Lehmann, Christian, 1991, « Grammaticalization an Related Changes in Contemporary German », in : Traugott E. & Heine B. (eds.), *Approaches to Grammaticalization,* Amsterdam, Benjamins, 493-535.

Marchello-Nizia, Christiane, 2006, *Grammaticalisation et évolution linguistique*, De Boeck, Bruxelles.

Meillet, Antoine, 1912, « L'évolution des formes grammaticales », *Rivista di scienza*, vol. XII, n° XXVI, 6 (=*Linguistique historique et linguistique gén*érale 1921, 130-148).

Meillet, Antoine, 1916, « Le renouvellement des conjunctions ». *Annuaire de l'Ecole Pratique des Hautes Etudes* 1915-1916*,* section historique et philologique, Paris (= *Linguistique historique et linguistique générale* 1921, 159-174).

Meillet, Antoine, 1921, *Linguistique historique et linguistique générale*, Paris, Champion.

Petit, Daniel, 1999, *Swe- en grec ancien : la famille du pronom réfléchi. Linguistique grecque et comparaison indo-européenne*, Leuven, Peeters, coll. linguistique, Société de linguistique de Paris (LXXIX).

Saussure (de), Ferdinand, *Cours de linguistique générale*, Ch. Bally & A. Séchehaye (eds.), 1st ed. 1916; 3rd ed. 1931 (= *CLG*); *Cours de linguistique générale,* Tullio de Mauro (ed.), Paris, Payot, 1972 (= *CLG*); *Cours de linguistique générale*, R. Engler (ed.), Wiesbaden, Harrassowitz, 1967-1974 (= *CLG/E*); *Ecrits de linguistique générale*, S. Bouquet & R. Engler (eds.), Paris, Gallimard, 2002; *Le troisième cours = « Notes préparatoires pour le cours de linguistique générale 1910-1911 »* and E. Constantin, *Linguistique générale. Cours de M. le Professeur de Saussure 1910-1911*, D. Gambarara & C. Mejia Quijano (eds.), *Cahiers Ferdinand de Saussure* 58, 2005, 71-290.

Tara, George Bogdan, unpublished, *Les périphrases verbales avec le verbe habeo en latin tardif*, Thesis defended in February 2008, University of Paris IV.