

What are etymological (and etymographical) units made of: vocables or lexemes?

Éva Buchi

1. Introduction

I would like to share with you some thoughts on a metalexical and metalexicalographical topic which is theoretical in its essence, but poses very practical problems to practitioners of historical lexicology and lexicography. I myself have to deal with them mainly in two kinds of contexts. First as a lexicographer at the ATILF lab in Nancy, France, where we are working on various etymological dictionaries, amongst them the *Dictionnaire Étymologique Roman* (DÉRom), a research project which brings together more than fifty scholars from fourteen (mostly European) countries. Secondly, I have to tackle this problem in my capacity as a teacher of the European Master in Lexicography (EMLex), which my university, Université de Lorraine, offers in collaboration with eight other universities.

Etymological dictionaries are typically defined as “*DICTIONAR[IES] in which words are traced back to their earliest appropriate forms and meanings*” (Hartmann & James 1998 s.v.). The element *word* in such definitions, although intuitively comprehensible, lacks technical rigour, and is therefore ambiguous. Consequently I will use instead the threefold terminology (as well as the typographical conventions attached to it) established within the theoretical framework of Meaning-text theory (see Mel’čuk 2012: 1: 21-44): *wordform* (defined as ‘segmental linguistic sign that is autonomous and minimal, i.e., that is not made up of other wordforms’), *lexeme* (‘set of wordforms, and phrases, that are all inflectional variants’), and *vocable* (‘set of lexical units –lexemes or idioms– whose signifiers are identical, whose signifieds display a significant intersection, and whose syntactics are sufficiently similar’).

Thus, in order to quote a simple (and simplified) example, *table1* and *tables1* are wordforms of the lexeme TABLE1 ‘article of furniture consisting of a flat top and legs’, and together with the lexemes TABLE2 ‘arrangement of items in a compact form’, and TABLE3 ‘upper flat surface of a cut precious stone’, TABLE1 makes up the vocable TABLE. This diagram shows the two wordforms which form the lexeme TABLE1, the two wordforms which constitute the lexeme TABLE2, and the two wordforms which make up the lexeme TABLE3, and how these three lexemes, because of their clear semantic link, compose the vocable TABLE.

Let’s go back now to Hartmann and James, who state that an etymological dictionary is a “*DICTIONARY in which words are traced back to their earliest appropriate forms and meanings*”. The question one might ask is: when they use the term *word*, should we understand ‘wordform’, ‘lexeme’, or ‘vocable’? You will probably agree that ‘wordform’ can be excluded as a valid interpretation: tracing back each wordform –say each inflectional variant of a verb– to its origin hardly forms part of the objective of etymology. But the decision between the possible interpretations ‘lexeme’ and ‘vocable’ cannot be made that easily. I did not find any indication in theoretical work: whilst every etymology manual defines in great detail what is an etymon, nobody seems to be interested in defining the unit whose origin the etymon constitutes. In order to decide between the two options, the lexeme or the vocable as etymological unit, rather than speculating theoretically, I propose we have a look at the current practice in etymological dictionaries.

2. Examples

2.1. Spanish *escaparate*

Of course the problem presents itself only with polysemous vocables, for instance with the Spanish masculine noun *ESCAPARATE*, which is made up of two lexemes: *ESCAPARATE1* ‘glass-door cabinet used for displaying delicate things’ and *ESCAPARATE2* ‘shop window used for displaying samples of what is sold in the shop’. As indicated by Corominas’s *Breve diccionario etimológico de la lengua castellana*, Spanish *ESCAPARATE* is borrowed from Old Dutch *SCHAPRADE* n. ‘closet (in particular kitchen cupboard)’. We can deduce from this entry that Corominas considers that the vocable and not the lexeme constitutes the etymological unit. Otherwise, as Old Dutch *SCHAPRADE* only means ‘closet’ and not ‘shop window’, he would have stated two etymologies: he would have said that *ESCAPARATE1* ‘glass-door cabinet’ is borrowed from Old Dutch *schaprade* ‘closet’, whereas *ESCAPARATE2* ‘shop window’ has been coined in Spanish, as a metaphorical semantic evolution from *ESCAPARATE1* ‘glass-door cabinet’.

Corominas’s practice is in accordance with Untermann’s approach to etymology, which I will quote here through Thomas Krisch’s translation: “For me, etymology is defined as: establishing and describing the process which produces a new sequence of phonemes and assigns a meaning to it, using given vocabulary and given grammatical means, in order to meet a requirement which emerges” (Krisch 2010: 317, quoting Untermann 1975: 105). If producing a new sequence of phonemes is constituent to the etymological unit, derivatives and compounds, for instance, merit an etymology, but internal creations which constitute semantic evolutions do not.

This implicit choice in favour of the whole vocable and not the individual lexeme as the etymological unit seems to be quite common in etymological dictionaries. If you know of counterexamples in the domains you are working in, I would be very interested to hear about them in the discussion.

2.2. German *Stichwort*

One apparent counterexample I found in Kluge’s etymological dictionary for German (Kluge & Seebold 2002), in the entry *Stichwort*. First, I was under the impression that for Kluge, the vocable *STICHWORT*, a neuter noun, contained four lexemes: *STICHWORT1* ‘signal (for instance a phrase) given to a performer to begin a specific speech or action’, *STICHWORT2* ‘word placed at the beginning of a dictionary entry’, *STICHWORT3* ‘key word used for structuring texts’, and *STICHWORT4* ‘hurtful remark’, this last lexeme being obsolete in contemporary German. Kluge indicates actually two etymologies for these four lexemes: *STICHWORT1-3* are explained as an internal creation of German, a compound with the verb *STECHEN* ‘to point out’, whereas *STICHWORT4* is analysed as another German compound, which contains the noun *STICH* ‘hurtful action’. A superficial reading of this entry made me assume this was an example for the lexematic approach to etymology, that Kluge attributed a specific etymology to at least one lexeme, *STICHWORT4*. But I was wrong: there is actually no clear semantic link between *STICHWORT1-3* on the one hand and *STICHWORT4* on the other: *STICHWORT1-3* have to be modeled as one vocable, and *STICHWORT4* as another. So we are in presence of two homonyms, both of which are of course etymologized independently, and the underlying understanding of the etymological unit is again that of a vocable.

2.3. French *agio*

The *Trésor de la langue française* (TLF) contains an entry *agio* which displays three lexemes: AGIO1, which is termed archaic, meaning ‘excess value of one currency over another’, AGIO2 ‘bank fees including interest, commission, and exchange’, and AGIO3 ‘dishonest speculation at the stock exchange’. In the etymology section of this entry, the TLF states: “Empr[unt] à l’ital[ien] *aggio*” [“borrowing from Italian *aggio*”]. As the TLF indicates only one etymology for these three lexemes, its etymological unit seems to be the vocable AGIO. But a closer look shows that this situation allows in reality for two possibilities: either only part of the three lexemes (one or two) are borrowed from Italian: then the etymological unit is indeed the vocable AGIO as the set of the synchronically linked lexemes AGIO1, AGIO2, and AGIO3; the individual lexemes are not etymologized. Or all three lexemes are borrowed from Italian –at the same moment or at different moments of the history of French–, in which case the etymological unit would rather be the sum of the three lexemes.

The TLF-entry, which was published in 1973, does not allow us to establish which one of these interpretations is correct. But the problem had to be addressed in the context of the TLF-Étym project, which proposes a selective revision of the etymologies contained in the *Trésor de la langue française informatisé* (TLFi). The entry *agio* was compiled by Franz Rainer, a professor at Vienna University of Economics and Business and a renowned specialist in Romance banking terms. Franz Rainer began by establishing first attestations for all three lexemes: for AGIO1 ‘excess value of one currency over another’, he could not go back further than 1679, the dating already proposed by TLF. But he dated AGIO2 ‘bank fees including interest, commission, and exchange’ back to 1723, and AGIO3 ‘dishonest speculation at the stock exchange’ to 1727. Then he looked for precise information on the etymon, Italian *aggio*. According to Cortelazzo’s and Zolli’s *Dizionario etimologico della lingua italiana* (DELI₂), the Italian vocable AGGIO contains two lexemes: AGGIO1 ‘excess value of one currency over another’, documented since 1498 (in the form of the irregular plural *agie*) and AGGIO2 ‘discount on the amount of a tax granted to state employees’, documented since 1892.

A comparison of the semantics of French AGIO and Italian AGGIO shows that only one of the three French lexemes was borrowed without doubt from Italian:

(1) French AGIO1 ‘excess value of one currency over another’ (documented since 1679) goes indeed back to Italian AGGIO1, which presents the same meaning and has been documented since 1498.

(2) Italian AGGIO2 ‘discount on the amount of a tax granted to state employees’ was not borrowed by French: it is of no relevance for French etymology.

(3) There does not seem to exist a counterpart within Italian AGGIO to French AGIO2 ‘bank fees including interest, commission, and exchange’: in contemporary Italian, we would use the phrase *spese di commissione*. Thus AGIO2 most probably represents an internal creation of the French language, more specifically a semantic evolution from AGIO1 ‘excess value of one currency over another’, the semantic link between the two lexemes lying in the sememe ‘commission’¹.

(4) As for French AGIO3 ‘dishonest speculation at the stock exchange’, there is no way it could have been borrowed from Italian, because Italian AGGIO does not present such a meaning (AGIO3 would translate into Italian *speculazione disonesta*). AGIO3 actually analyses as another

¹ As there is some doubt about the meaning of *aggio* in the early Italian documentation, Franz Rainer does not completely exclude the possibility of AGIO2 being as well a borrowing from Italian.

French innovation, a semantic evolution from AGIO₂ ‘bank fees including interest, commission, and exchange’, the common denominator being the sememe ‘profit’. Franz Rainer explains convincingly the appearance of this new lexeme in the aftermath of what was called the Law affair. John Law was a Scottish economist who became a millionaire by issuing huge amounts of share certificates of his Mississippi Company to the French. These shares were ultimately rendered worthless, and initially inflated speculation about their worth led to a chaotic economic collapse in France in the 1720. This semantic innovation is thus strongly linked to a French economic context and could only have been made within the French language.

In short, the French vocable AGIO represents a borrowing from Italian, because its first lexeme, AGIO₁, was borrowed from this language. But AGIO₂ and AGIO₃ are to be considered internal creations of the French language.

3. Conclusion

So what are etymological (and etymographical) units made of: vocables or lexemes? In general, dictionary entries are made up of vocables like Spanish ESCAPARATE, German STICHWORT¹ and STICHWORT², and French AGIO. For this reason, the easy answer to my question is that etymological units are made of vocables. Indeed, most etymological dictionaries –as we saw, that is the case for Corominas, Kluge, and TLF– implicitly proceed as if this were the case, without, however, discussing this option, for instance in their prefaces. But in my opinion, Spanish ESCAPARATE₂ ‘shop window used for displaying samples of what is sold in the shop’ cannot be considered a borrowing from Dutch, and French AGIO₃ ‘dishonest speculation at the stock exchange’ cannot be considered a borrowing from Italian. I think these lexemes deserve to be properly etymologized as internal creations of the Spanish and the French language.

For this reason, I would like to advocate *a contrario* that individual lexemes and not whole vocables are best hypostatized as etymological and etymographical units. Once you look at it properly, this ruling seems quite obvious. I would actually be surprised if you disagreed with me when I state that etymologies which put the different lexemes of a vocable to the centre of their attention are better, more precise, more complete etymologies. But if this is so evident –and I really think it is–, why did the discipline of etymology have to wait for 2016 and a humble class of the European Master in Lexicography for this finding to be put forward?

I think the answer lies in the terminology we use, and ultimately in our conceptualization of the units of the lexicon of a language. If we call ESCAPARATE₂ ‘shop window used for displaying samples of what is sold in the shop’ a simple “meaning” of the “word” ESCAPARATE, then there is of course no need for etymologizing it: only “words”, i.e. linguistic signs, i.e. sets of signifiers, signifieds, and syntactic properties, may and should be etymologized. If the “word” *table*, a noun, presents three “meanings”, then of course *table* is the etymological and etymographical unit. But if we analyse TABLE as a vocable made up of three lexemes, each of them presenting a signifier, a signified and syntactic properties, the stage is definitely set for the lexeme becoming the etymological and etymographical unit.

4. Bibliography

- Corominas, Joan (1973³ [1961¹]): *Breve diccionario etimológico de la lengua castellana*. Madrid: Gredos.
- DELI₂ = Cortelazzo, Manlio & Zolli, Paolo (1999² [1979–1988¹]): *Dizionario etimologico della lingua italiana*. Bologna: Zanichelli.
- Hartmann, R. R. K. & James, Gregory (1998): *Dictionary of lexicography*. London/New York: Routledge.
- Kluge, Friedrich & Seebold, Elmar (2002²⁴ [1883¹]): *Etymologisches Wörterbuch der deutschen Sprache*. Berlin/New York : De Gruyter.
- Krisch, Thomas (2010): "Etymology". In: Luraghi, Silvia & Bubenik, Vit (edd.): *Bloomsbury Companion to Historical Linguistics*. London/New Delhi/New York/Sydney: Bloomsbury: 311-322.
- Mel'čuk, Igor' A. (2012/2013): *Semantics: From meaning to text* (2 volumes). Amsterdam/Philadelphia: Benjamins.
- TLF = Imbs, Paul & Quemada, Bernard (edd.) (1971–1994): *Trésor de la langue française. Dictionnaire de la langue du XIX^e et du XX^e siècle (1789–1960)* (16 vol.). Paris: Éditions du CNRS/Gallimard.
- TLFi = Imbs, Paul & Quemada, Bernard (edd.) (1971–1994): *Trésor de la langue française informatisé. Dictionnaire de la langue du XIX^e et du XX^e siècle (1789–1960)*. Nancy: ATILF: <http://atilf.atilf.fr/tlf.htm>.
- TLF-Étym = Steinfeld, Nadine (ed.) (2005–): *Trésor de la Langue Française Étymologique (TLF-Étym)*. Nancy: ATILF: <http://www.atilf.fr/tlf-etym>.
- Untermann, Jürgen (1975): "Etymologie und Wortgeschichte". In: Seiler, Hansjakob (ed.): *Linguistic Workshop III. Arbeiten des Kölner Universalienprojekts*. Munich: Fink: 93-117.