

Corpus-based analysis of change in a dead language: the case of assimilation in Latin *con-*

The study of dead languages presents general problems even in the case of well-documented languages with large surviving written corpora (inscriptions and manuscripts). In the case of (Classical) Latin, the bulk of the remaining corpus is not contemporary but is often several times removed from the period when the language was spoken. When studying variation phenomena that point at language change, such factors are to be taken into account. For the phonologist it is a crucial question to what extent the patterns found in the corpus reflect phonological reality and thus contribute to a realistic picture of phonological change.

Consonantal assimilations at prefix–stem boundaries have been objects of detailed scrutiny in the past (Buck 1899, Prinz 1949 and 1953, García González 1996), and while the major patterns have been made clear, these studies did not involve proper quantitative analyses. The present paper attempts to take a close look at the assimilation between the nasal of *con-* and stem-initial *l* (e.g. *conlatus/collatus* ‘carried’; for the nature of the nasal consonant cf. de Vaan *s.v.* and Cser 2011:80–81). The actual corpus we use is vol 1 of CLCLT-5 (release 2002, cca 50 million words, we do not include medieval texts in the study). Frequency of attestation and the proportion of assimilated vs. non-assimilated forms is compared and a correlation is found, although the set of relevant forms attested at least about 300 times in the corpus is small (six words altogether; the rarest word we included is *conlabi* ‘collapse’ with 290 occurrences).

Prefixation led in many cases to lexicalisation, which in turn resulted in drastic phonological modifications at the prefix–stem boundary as well as within the stem. The pace and the extent of lexicalisation, however, was highly variable. Furthermore, prefixation also involved recomposition in all periods of the documented history of Latin. The varying pace of lexicalisation and the varying degree of transparency coupled with the phonological processes that took place at prefix–stem boundaries resulted in a not particularly straightforward relation between written forms and probable phonological variants. It is also clear that beginning with the 1st century AD etymologically oriented habits of spelling began to gain ground, but did not affect all words of a similar composition to the same extent.

Given these difficulties it is all the more notable that if one looks at these *conl-/coll-* words that are attested in the corpus with sufficient frequency to allow for any meaningful generalisation, there seems to be a correlation within this small set between frequency of attestation and frequency of assimilation. This may possibly point to a relatively faithful representation of phonological reality in the extant manuscript corpus.

References

- Buck, Carl Darling (1899) Notes on Latin orthography. *The Classical Review* 13:116–9, 156–167.
- Cser, András (2011) The floating C-Place node in Latin. *Journal of Linguistics* 47:65–85.
- de Vaan, Michiel (2008) *Etymological dictionary of Latin and the other Italic languages*. Leiden: Brill.
- García González, Juan José (1996) Asimilación de prefijos en inscripciones latinas. In: A. Bammesberger, F. Heberlein (eds) *Akten des VIII. Internationalen Kolloquiums zur Lateinischen Linguistik*. Heidelberg: Winter.
- Prinz, Otto (1949) Zur Präfixassimilation im antiken und im frühmittelalterlichen latein. *Archivum Latinitatis Mediae Aetatis* 21.87–115.
- Prinz, Otto (1953) Zur Präfixassimilation im antiken und im frühmittelalterlichen latein. *Archivum Latinitatis Mediae Aetatis* 23.35–60.