

Linguistic evidence and diachronic truth

Carola Trips

Institut für Linguistik/Germanistik, Universität Stuttgart

carola.trips@ling.uni-stuttgart.de

Recently, working with corpora has become more and more important in the field of linguistics. This is especially true for the field of diachronic linguistics (see e.g. Kroch and Taylor (1997), Pintzuk (1999), Pintzuk et al. (2000), Haeberli (2002), Trips (2002)) since there are no native speakers available that could be asked for linguistic evidence on one phenomenon or the other. And it is here that the word “evidence” seems to bear a special meaning.

If we have a look in the Oxford English Dictionary (OED) we find the following entry for evidence: “the quality or condition of being evident”, and in the Merriam Webster Dictionary we find “something that furnishes proof”. Now if we look at diachronic data it will instantly become clear that the evidence found here often furnishes proof only to a certain degree, and that this is inevitably the case due to the nature of diachronic data.

Morphological phenomenon: Development of suffixes in the history of English

- (1) *þæt he þær onfenge [ærc**e**bis**c**opes **hade**] ...*
that he there received archbishop's office
(cobede,Bede_3:21.248.11.2540)
- (2) *for þou iugeþ þe folk in **euenhede**, ...*
for thou judges the people in equality
(CMEARLPS,77.3378)

By investigating linguistic phenomena like, e.g., the development of suffixes in the history of English demonstrated with the examples in (1) and (2) we have to take a number of factors into account that are absent in other fields of linguistics like, e.g., neurolinguistics or psycholinguistics. Here, we always have to be aware of the fact that the diachronic corpora available consist of written records and that there always is a discrepancy between spoken and written language so that the data found in these corpora always reflect a restricted use of language. Being aware of these factors, we have to contemplate in how far they have an impact on our theory about this phenomenon, i.e., is it plausible to assume that the authors of the texts used *had* or *hed* respectively as free or bound morpheme only in written language and had a different status in their spoken language (in this case it is probably not). Further, sometimes we find a situation

where we have a written standard (as, e.g., in Old English times) that might obscure how people “really” wrote. Moreover, there are texts where we discover a discrepancy between the date of composition and the date of the manuscript:

Syntactic phenomenon: V2 order in Middle English

- (3) *his medlid liffe shewith oure Lorde in hym silfe ...*
this moderate life shows our lord in himself
(CMROLLTR,27.555)

Wrt the philological information for the *Prose Treatises from the Thornton Ms.* from Richard Rolle the PPCME2 notes that the date of the composition is between 1250 and 1350 whereas the date of the manuscript is between 1420 and 1500. If we then investigate the development of a syntactic phenomenon like, e.g., V2 orders (see example (3)) it is sometimes not clear whether the grammar of that text is the one of the original author or that of the author who wrote the manuscript.

All these factors are a special property of diachronic research work and lead to the understanding that here linguistic evidence is often more a linguistic hint than a proof and that this evidence has to be used cautiously. This is especially true if the problem of negative evidence is addressed: the probability of counter-examples cannot be reduced by enlarging the corpus, contrary to working with contemporary corpora.

In the talk I will present diachronic data for a number of morphological and syntactic phenomena and discuss the factors mentioned above. Further, I will contrast problems arising from working with diachronic data with problems that might arise from working with synchronic data. The data presented in the talk are from the largest annotated diachronic corpora (for English) available at present, The York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE), and The Penn-Helsinki Parsed Corpus of Middle English 2 (PPCME2).

References

- Haeberli, E. (2002). Features, categories and A-positions: cross-linguistic variation in the Germanic languages. Kluwer, Dordrecht.
- Kroch, A. and A. Taylor (1997). Verb movement in old and middle english: Dialect variation and language contact. In A. v. Kemenade and N. Vincent, eds., *Parameters of Morphosyntactic Change*, pp. 297–325. Cambridge University Press, Cambridge.
- Kroch, A. and A. Taylor, eds. (2000). *The Penn-Helsinki Parsed Corpus of Middle English, Second Edition (PPCME2)*. University of Pennsylvania, Philadelphia.
- Pintzuk, S. (1999). *Phrase Structures in Competition: Variation and Change in Old English*. Revision of Pintzuk 1991. Garland, New York.

- Pintzuk, S., G. Tsoulas, and A. Warner, eds. (2000). *Diachronic Syntax. Models and Mechanisms*. Oxford University Press, Oxford.
- Taylor, A., A. Warner, S. Pintzuk, and F. Beths, eds. (2003). *The York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE)*. University of York.
- Trips, C. (2002). *From OV to VO in Early Middle English*. John Benjamins, Amsterdam/Philadelphia.