

Historical Developments in the Marking of English Relative Clauses

C.E.A. Diertani
University of Pennsylvania

This paper focuses on the historical developments in the marking of relative clauses in English. Relative clause marking in English displays a great deal of variation between overt and null *wh*-elements and overt and null complementizers; there are thus four possible patterns of marking, all of which are attested in English, albeit with very different synchronic and diachronic frequencies:

- (1) Overt *wh*-element, null complementizer;
- (2) Null *wh*-element, overt complementizer;
- (3) Overt *wh*-element, overt complementizer;
- (4) Null *wh*-element, null complementizer

The data used in this paper are drawn from two syntactically annotated corpora of historical English¹, covering about six hundred years of texts, the latest texts from ca. 1710. The antecedents of each clause were coded for various syntactic and textual properties. Some relative clauses could not be captured using the technology available; these included conjoined relative clauses; multiple relative clauses belonging to a single antecedent; and extraposed relatives. Thus the corpus contains only in-situ relatives, none of which share antecedents. This amounts to some 37,416 relative clauses.

In the earliest Middle English texts, by far the most frequent clausal marking is that in (2) above. However, as Figure 1 shows², this type decreases dramatically, while its opposite, the type (1), increases via the classic S-shaped curve to become the most common type. Contact relatives, with no overt markers, also increase in frequency.

Restrictive and non-restrictive relative clauses are not distinguished in the conventions of annotation used in the corpus; in order to approximate the distinction, I examined a subset of the corpus, using bare quantifiers for restrictive clauses and proper names as stand-ins for non-restrictive relative clauses. The results were strikingly different: in the “restrictive” type (cf. Figure 2), pattern (1) never really takes off, and patterns (2) and (4) are almost equally common by the late Early Modern English period, whereas among “non-restrictive” clauses (cf. Figure 3), pattern (1) becomes virtually the only type in the latest texts.

Given the low frequency of the double-marked type in (3), one might argue that this is a single development: because double-marking is strongly dispreferred, the increased use of overt *wh*-elements led to the decrease in overt complementizers. In this paper, however, I argue the opposite, that there were in fact two separate developments: an increase in overt *wh*-elements in specifically non-restrictive relative clauses and an overall decrease in overt complementizers in the language generally. This is due to the existence of the two minority patterns (if the two constructions (1) and (2) were truly in complementary distribution, one would not expect to get cases of both, nor of neither), and to the fact that the incidence of doubly-marked clauses appears to peak around the same time that the frequencies of the singly-marked clause types are most nearly equal. In addition, if the decrease in overt complementizers were caused by the increase in overt *wh*-elements, it would be expected that the change in frequency for overt *wh*-elements should be the same as that of *null* complementizers, and the decrease in overt complementizers should be parallel to that of null *wh*-elements. This parallelism does appear in the earliest Middle English texts, but becomes increasingly less evident as time progresses. This is not what one would expect if the presence of one marker type directly impacted the other. Ongoing work on complementizers in dependent clauses of all types so far does appear to indicate that overt complementizers have indeed been decreasing in all environments.

¹ I.e. the Penn-Helsinki Parsed Corpus of Middle English (Kroch and Taylor 2000) and the Penn-Helsinki Parsed Corpus of Early Modern English (Kroch, Santorini, and Delfs 2004).

² Note that there are no data from periods 2, 5, and 7 on the chart.

Figure 1

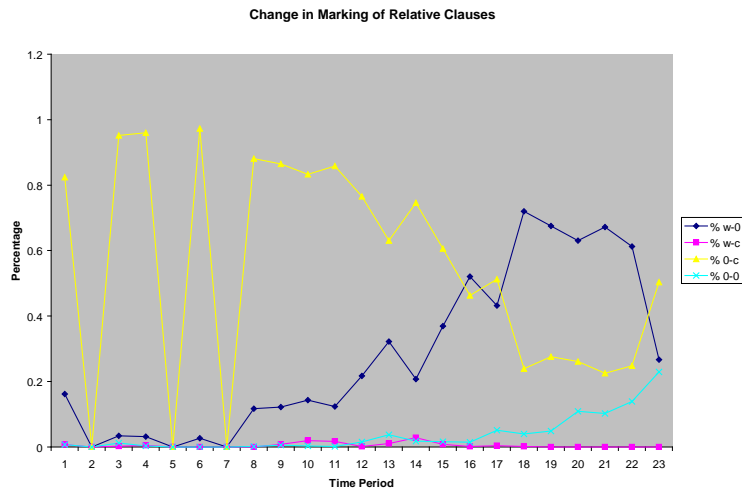


Figure 2

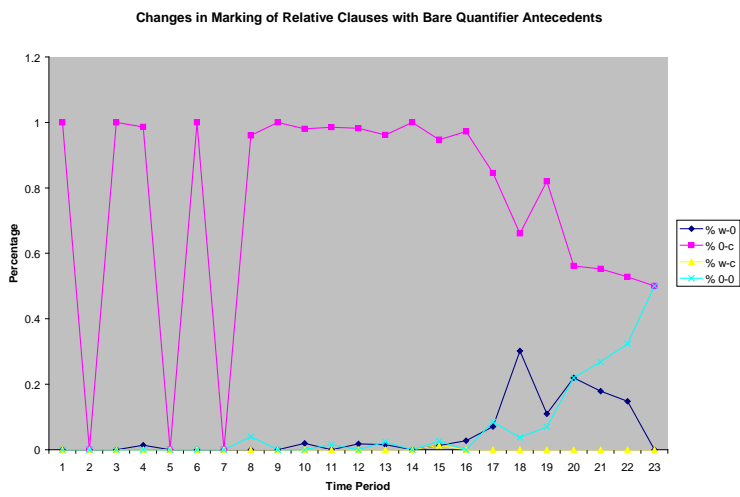


Figure 3

