Discourse reference by topics in Old and Middle English
Ans van Kemenade University of Nijmegen, CLS/Dept. Of English Language and Culture

This paper explores some aspects of the interrelation between syntax and discourse in Old and Middle English.

In van Kemenade and Milicev (2005) and van Kemenade and Los (2005), evidence is presented that a set of Old English short adverbs, including most prominently *þa* and *þonne*, both literally meaning ‘then’, have the morphosyntactic function of separating the clausal subdomain hosting topic elements from the rest of the clause. For concreteness’s sake, the following partitioning of the clause is assumed:

\[
\begin{array}{ccc}
\text{CP} & \text{Topic domain} & // \text{þa/þonne} // \ldots. \\
\text{Clause typing} & \text{discourse anaphoric material} & \text{discourse particle} \ldots.
\end{array}
\]

‘Topic’ here is to be understood in the sense of representing material that is presupposed or ‘given’ in the previous discourse. Elements typically occurring in this subdomain in Old English include:

1. pronouns of various kinds, such as personal pronouns (subject or object), indefinite pronouns and the impersonal subject pronoun *man*.
2. Definite NP’s, i.e. subject NP’s premodified by a demonstrative pronoun of the *se* series, or NP’s (subject or object) consisting solely of a demonstrative pronoun of the *se* series used independently.

The special ‘high’ position of personal pronouns is well-known in the literature. The behaviour of 2 is more puzzling, especially in view of the fact that definite NP subjects in Old English may appear in a lower position as well. The hypothesis pursued in this paper is that demonstrative pronouns, beside being definite markers, can be anaphoric as well. This is especially clear from the fact that *se* demonstratives can be used independently in Old English. What distinguishes high definite NP’s from lower definite NP’s is that high definite NP’s contain a demonstrative pronoun that refers concretely to a discourse referent.

In the transition to Middle English, the demonstrative *se* series is lost and replaced by the definite determiner. This reduces the scope for discourse anaphoric marking and has a profound influence on the marking of discourse cohesion in interrelation with the syntactic position of subjects. Put briefly (and cutting several corners), the topic domain as assumed here comes to be reanalysed as the (single) position for the syntactic subject. This happens in subclauses earlier than in main clauses, which makes sense given the discourse functions of main and subclauses.

If this scenario is correct, we expect profound changes in the discourse linking of topic/subject material between Old English and Middle English. This will be explored in this paper, using quantitative methodology inspired by Bresnan, Cueni, Nikitina and Baayen (2005), in which the properties of the element in the topic/subject position (pronominality, animacy, definiteness, anaphoricity, discourse accessibility, length) are measured, making exhaustive use of the York Corpus of Old English and the Penn/Helsinki Parsed Corpus of Middle English.