

Topicalization is (Still) Disappearing: information uniformity as a dimension of specialization

Joel C. Wallenberg, *University of York*

This article presents new data on the decline of DP “topicalization” (i.e. object fronting) in Late Early Modern English and its implications for the hypothesis that speakers prefer informational uniformity in their utterances (Fenk and Fenk, 1980; Fenk-Oczlon, 1989, 2001; Levy, 2005; Levy and Jaeger, 2007, *inter alia*). This study applies the version of the hypothesis in Cuskley et al. (2021), that speakers distribute the information content of words as uniformly as possible across an entire utterance, to help explain the statistical patterns of syntactic change over long periods of time (in a similar way to Wallenberg et al., 2021). In particular, I will show that the topicalization of direct objects in English has declined since the Late Early Modern English period into the modern period, a pattern which is itself a continuation of the decline observed by Speyer (2008, 2010) in Middle English and Early Modern English. Furthermore, the decline persists even though the topicalization and non-topicalization orders are functionally differentiated (i.e. specialized) for different informational and prosodic contexts.

Speyer (2008, 2010) showed that the frequency of direct object topicalization, and fronting of other constituents, declined during much of the history of English. He also tied this to a particular prosodic effect: the desire of speakers to avoid a stress clash between the subject and the fronted constituent. He showed observationally and experimentally that topicalization sentences such as (1-a) below, in which a prosodically strong/heavy nominal object appears next to a similar nominal subject were dispreferred compared to topicalizations in which the subject was a pronoun, as in (1-b), and *in situ* sentences (2).

- (1) a. The cat Mary likes.
b. The cat she likes.
- (2) Mary/she likes the cat.

Speyer further convincingly argued that the avoidance of (1-a) is due to prosodic clash between object and subject, showing that experimental participants insert a pause between the two if they are required to produce sentences like (1-a). He then explained the decline in topicalization over the history of English as a consequence of the decline of verb-second (V2) orders, which allowed for a prosodically weaker verb to intervene between object and subject in sentences with fronting (as occurs in all the modern Germanic languages other than English). Note also that the same strategies which prevent the prosodic clash *also* prevent a peak of high information item (i.e. lower frequency words and phrases) at the beginning of a clause. However, even though topicalization and *in situ* object orders could be considered specialized for clash and non-clash contexts, this description does not explain why they still overlap in function Prince (1998).

I explore the hypothesis that, in addition to prosodic context, the use of the orders also depends on the degree of information uniformity each one allows a speaker to achieve in an utterance (as measured by the “Deviation of the Rolling Mean” (DORM) of a sentence’s ordered information content (surprisal) values, cf. Cuskley et al. 2021; Wallenberg et al. 2021). Because degree of uniformity is a continuous dimension, the two syntactic variants cannot specialize completely along it (Wallenberg, 2016, 2019); they therefore still overlap in their functions, and so compete in use Kroch (1994). The current study uses the *Penn Parsed Corpus of Early Modern English* and *Penn Parsed Corpus of Modern British English 2* (PPCEME, PPCMBE2; Kroch et al., 2004, 2016), and finds

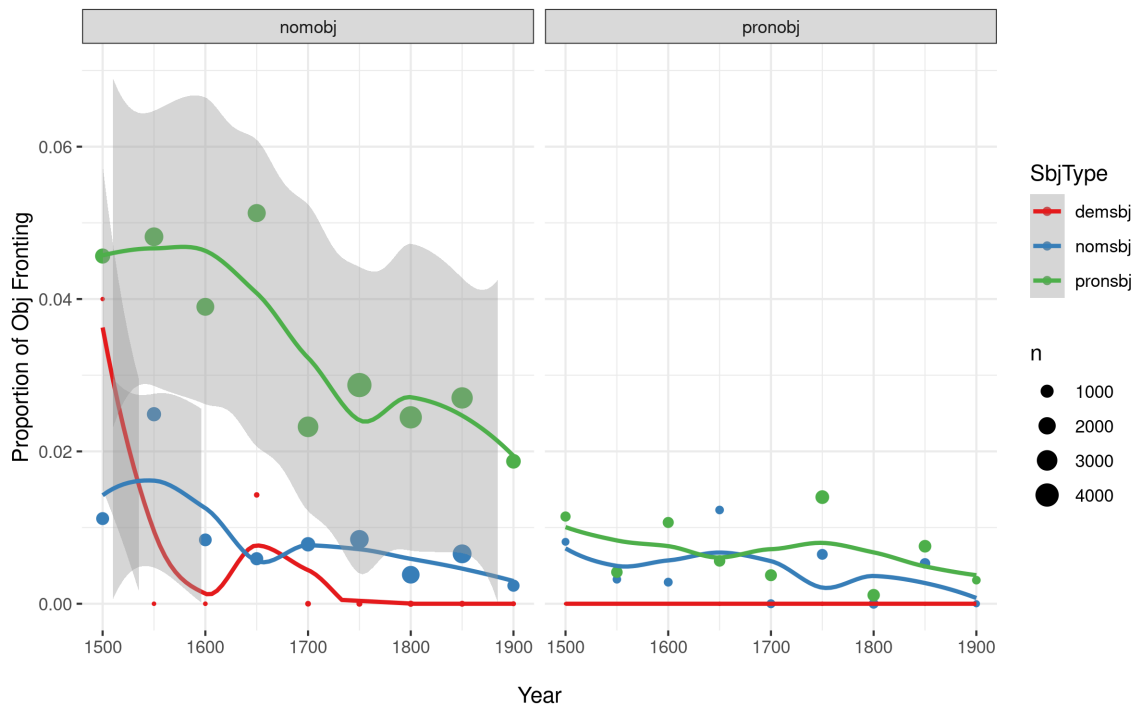


Figure 1: Decline of direct object DP fronting in Early Modern and Modern British English, as measured in the PPCEME and PPCMBE2 (4.5 million words from 1500 to 1920), faceted by object type (nominal vs pronominal and subject type (demonstrative, nominal, or pronominal).

two results confirming the incomplete specialization hypothesis: first, the decline in topicalization continued after the period in which V2 orders were lost, and secondly, the decline persists in clearly non-V2 sentences, even where the subject is a pronoun (see Figure 1).

This continued competition between the variants and consequent diachronic decline of one variant, even though it has slowed (as predicted in Wallenberg, 2016), is entirely what one would expect if the variants cannot completely differentiate in their contexts of use. The full article will go further in relating the speed of decline to the information uniformity DORM measures of the relevant populations of sentences.

References

- Cuskley, Christine, Rachael Bailes, and Joel Wallenberg. 2021. Noise resistance in communication: Quantifying uniformity and optimality. *Cognition* 214:104754.
- Fenk, August, and Gertraud Fenk. 1980. Konstanz im kurzzeitgedächtnis-konstanz im sprachlichen informationsfluß. *Zeitschrift für experimentelle und angewandte Psychologie* 27:402.
- Fenk-Oczlon, Gertraud. 1989. Geläufigkeit als determinante von phonologischen backgrounding-prozessen. *Papiere zur Linguistik* 40:91–103.
- Fenk-Oczlon, Gertraud. 2001. Familiarity, information flow, and linguistic form. *Typological Studies in Language* 45:431–448.
- Kroch, Anthony, Beatrice Santorini, and Lauren Delfs. 2004. Penn-Helsinki Parsed Corpus of Early Modern English, release 3. Size 1.8 Million Words.
- Kroch, Anthony S. 1994. Morphosyntactic variation. In *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, ed. K. Beals et al.
- Kroch, Anthony S., Beatrice Santorini, and Ariel Diertani. 2016. Penn Parsed Corpus of Modern

- British English 2nd edition, release 1. Size \sim 2.8 million words.
- Levy, Roger. 2005. *Probabilistic models of word order and syntactic discontinuity*. Stanford University.
- Levy, Roger P., and Florian T. Jaeger. 2007. Speakers optimize information density through syntactic reduction. In *Advances in neural information processing systems*, 849–856.
- Prince, Ellen. 1998. On the limits of syntax, with reference to left-dislocation and topicalization. *Syntax and semantics* 281–302.
- Speyer, Augustin. 2008. Topicalization and clash avoidance: On the interaction of prosody and syntax in the history of English with a few glimpses at German. Doctoral Dissertation, University of Pennsylvania.
- Speyer, Augustin. 2010. *Topicalization and stress clash avoidance in the history of english*. Topics in English Linguistics. Berlin/New York: Mouton de Gruyter.
- Wallenberg, Joel C. 2016. Extraposition is disappearing. *Language* 92:e237–e256.
- Wallenberg, Joel C. 2019. A variational theory of specialization in acquisition and diachrony. *The Determinants of Diachronic Stability* 254:245.
- Wallenberg, Joel C, Rachael Bailes, Christine Cuskley, and Anton Karl Ingason. 2021. Smooth signals and syntactic change. *Languages* 6:60.