

Examining stylistic influences on the evolution of *do*-support

Aaron Ecay

Introduction Warner (2005) presents an account of the evolution of *do*-support which attributes the decline observed in *do*-support rates around 1575 to a change in the social meaning of the incoming variant. Specifically, he proposes that “a new evaluative principle for register variation is introduced in the late 16th century,” and that this principle, militating against the usage of “do not” sequences and/or the contraction of auxiliary verbs (including “do”) and “not,” can explain the decline in the use of *do*-support in negatives, without need for grammatical explanations.

Corpora Warner’s study was based on Ellegård (1953)’s corpus, which was specially collected to study *do*-support. We present data drawn from the Penn Parsed Corpus of Early Modern English and the Parsed Corpus of Early English Correspondence. This combined corpus provides a dataset which, though slightly smaller, is roughly comparable in size to Ellegård’s, but which was not a targeted (and thus potentially biased) sample. We will see that this data replicates Warner’s basic conclusions, while also filling in finer details.

Preliminaries The first observation of the present study is a methodological one. Warner used both type-token ratio and word length as measures of a text’s style, combining both measurements into a single index of style. He reports that the two separate variables each correlate with the usage of *do*-support and that they correlate well with each other. PC data do not bear out these observations. Type-token ratio and word length are somewhat correlated ($R^2 = 0.12$). However, the two style indices do not associate with *do*-support usage in the same way across the pre- and post-1575 periods (Table 1). Word length displays the positive-then-negative association with *do*-support described by Warner. Type-token ratio however is always negatively associated with *do*-support. In the spirit of seeking to replicate Warner’s findings, we will thus take only word length into account as a measurement of style in the present study.

Replication Moving on to the data, we follow Warner in dividing the texts into high- and low-style halves. We find as Warner did that the dip in *do*-support is a property entirely of the higher register (Fig. 1). However, our results differ from his in the area of affirmative questions. Warner does not report on affirmative questions, because Ellegård’s data do not show an overall decline in *do* usage in that context. The PC data do show a decline however, and it relates to style just as the other contexts examined by both us and Warner do (the blue lines in Fig. 1). This evidence stands opposed to Warner’s conclusion that it is *not* which mediates the decline in *do*-support usage, by triggering a novel negative social evaluation after 1575 and temporarily depressing the surface rates of *do*-support.¹

Extension Han and Kroch (2000) have proposed that the trajectory of negative imperatives is impacted by a grammatical effect centered around 1575. Specifically, for them the loss of verb-raising consists of two changes. The first, which takes place before 1575, involves the loss of raising to T, while verbs continue to raise to a lower Asp projection. Imperatives, which for them do not contain T, are not affected by this change. On the other hand, after 1575 raising to Asp begins to be lost, meaning previously-inert imperatives are suddenly affected. Examining the data with Warner’s proposed register split, we see that we must include style in this description as well (purple lines in Fig. 1). In higher-style texts, negative imperatives behave just as Han and Kroch describe; however in low-style texts the trajectory is monotone increasing, just as in all other contexts. This could indicate that Han and Kroch were mistaken in their account, or that the high and low styles may be tracing different grammatical paths to the same outcome, and only in the higher register is Han and Kroch’s account fully realized.

Ecay (2012) points out the significance of affirmative declaratives for the rise of *do*-support, arguing that they betoken an intermediate, never homogeneously realized, grammatical step between Middle and Modern English. Examining them through Warner’s style lens shows that there is an appreciable gap between the two register groups at the beginning of the change, which collapses after 1575 (Fig. 2). This is the opposite of what happens in modern *do*-support contexts, where early style effects are mild, if they exist at all, and become pronounced only after 1575. This result strengthens Ecay’s conclusion that the affirmative declarative context has a different status from the others. On the other hand, the fact that *do* usage in low-style affirmative declaratives begins to decline only in 1625 or slightly later calls into question the widespread hypothesis² that there was a single reanalysis event responsible for both the eventual disappearance of affirmative declarative *do* and the dip in *do*-usage in other contexts.

Conclusion Examining a new source of data shows that Warner’s basic intuition that a stylistic effect was at work in the history of *do*-support was on the right track. However, we propose, given the PC data, that the negative evaluation attached to *do*, not to *not* as Warner proposed. This explains the suppression of *do* in high style texts across all sentence types. It also explains why *do* in affirmative declaratives begins to decline 50 years earlier in high-style texts than in

¹We cannot attempt to replicate the other prong of Warner’s support for his conclusion, involving contraction-facilitating vs. -forbidding word orders in negative questions, because Ellegård’s data unlike the PCs vastly oversamples negative questions; the random sample the PCs provide is too small to extract useful data on negative questions.

²To enumerate non-exhaustively a few examples, this hypothesis is adopted by Kroch (1989), Warner (1993), and Postma (2010).

low-style ones: in the former case, it is the stylistic effect that is driving the decline, and not the loss of affirmative declarative *do* from speakers' grammar. This proposal has implications for purely grammatical accounts of *do*-support, as we have discussed. The data presented also provide fertile ground for building quantitative models of the interaction between linguistic change social forces. Finally, this study serves as an illustration of the value of parsed corpora to replicate and extend – and thus strengthen – familiar results.

References Ellegård, A. (1953) *The auxiliary do: The establishment and regulation of its use in English*. Han, C-H. and Kroch, A. (2000) "The rise of *do*-support in English: implications for clause structure." *Proceedings of NELS* 30. Postma, G. (2010) "The impact of failed changes." In: *Continuity and Change in Grammar*, pp. 269-302. John Benjamins. Warner, A. (1993) *English auxiliaries: Structure and history*. Cambridge University Press. Warner, A. (2005) "Why *do* dove: evidence for register variation in Early Modern English negatives." *Language Variation and Change* 17, pp. 257–280.

	Pre-1575	Post-1575
Type-token ratio	-	-
Mean word length	+	-

Table 1: The association (sign of the coefficient) between style measures and *do*-support usage as measured by a logistic regression. Other predictors in the model were clause type (\in {negative declarative, affirmative question, negative question}), year of composition (standardized as z-score) and a random intercept per author.

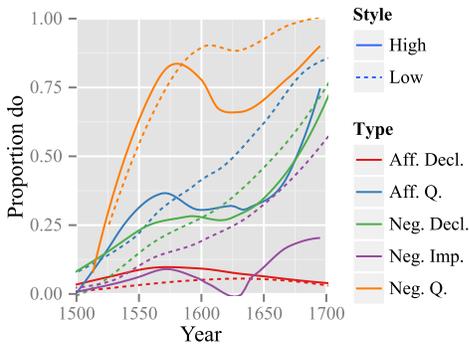


Figure 1: Trajectories of *do*-support in high- and low-word-length texts. The high-word-length all show noticeable dips or flat regions around 1575–1625, whereas the dashed (low) all increase without pause (with the exception of negative questions, which are the least frequent category and which have already gone virtually to completion).

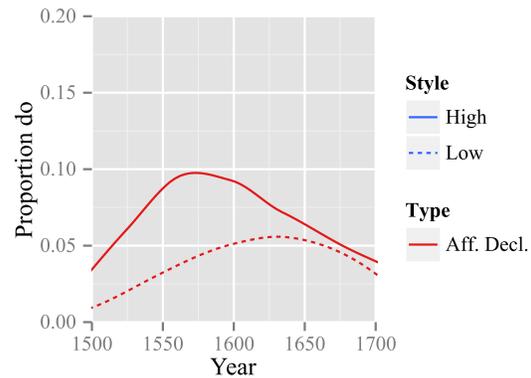


Figure 2: The effect of word length on *do* in affirmative declaratives (excerpted and zoomed in from Fig. 1)