

"Do so"

Beatriz Macía-Vega (Liverpool & Santiago de Compostela) and John Payne (Manchester)

In this paper, we present a corpus-based analysis of the *do so* construction. This analysis aims to establish the factors which drive speaker's choice of the *do so* construction in preference to alternative anaphoric devices such as auxiliary stranding and forms such as *do it*, *do this*, *do that*, *do likewise* and *do the same*. Examination of authentic data reveals the existence of counterexamples to each of the factors which have sometimes been considered categorical. For example in (1) we see a natural sounding counterexample to the well-known constraint that "orphans" must have adjunct status:

- (1) And nice to see the end of Tony Blair apologising for things which clearly weren't his fault, while failing to do so for any of the things which clearly are.

Here we believe it plausible to analyse *apologise for* as a prepositional verb, and the underlined "orphan" as a prepositional complement.

Such counterexamples suggest that the grammatical space occupied by the *do so* construction has been underestimated, and that a probabilistic model is appropriate. We therefore invoke the *logistic regression method*, as used for example by Bresnan, Cueni, Nikitina & Baayen (2005) in an analysis of the dative alternation, in order to quantify the importance of the various factors involved.

Reference:

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & Harald Baayen. (2005). "Predicting the dative alternation". In *Cognitive foundations of interpretation*, ed. by G. Boume, I. Kraemer & J. Zwarts. Amsterdam: Royal Netherlands Academy of Science.