

Word Classes and Word Order

August Fenk (august.fenk@uni-klu.ac.at)

Department of Media and Communication Studies and Department of Psychology, Alps-Adriatic University of Klagenfurt
Universitaetsstrasse 65-67, 9020 Klagenfurt, Austria

Gertraud Fenk-Oczlon (gertraud.fenk@uni-klu.ac.at)

Department of Linguistics and Computational Linguistics, Alps-Adriatic University of Klagenfurt
Universitaetsstrasse 65-67, 9020 Klagenfurt, Austria

Keywords: Function words; content words; word frequency.

Wordclass effects in the serial position curve

In a previous study (Auer, Bacik, & Fenk, 2001) subjects listened a German scientific text by Glasersfeld. A tone at the end of some of the sentences ($n = 10$) signaled them to recall immediately as much as possible from this sentence. A weak primacy effect and a significant recency effect appeared despite the differences between the test sentences (in length, syntactical structure, etc.). These position effects were operationalized as the gradients from the middle part to the first quarter (primacy part) and to the last quarter (recency part) of the sentences. In a statistical reanalysis (Fenk & Fenk-Oczlon, 2006) using the very same operationalization we found (Wilcoxon test) significantly higher relative recall scores in content words than in function words in all three parts of the sentences. This reanalysis was inspired by an unexpected but for linguists even more interesting wordclass-effect, i.e. a significantly higher proportion of function words in the primacy part and of content words in the recency part of the sentences.

A general tendency of function words to occupy initial positions?

To find out whether this tendency was a characteristic only of the author Glasersfeld we inspected texts – each third sentence, if at least 4 words long - from 9 further German authors (4 scientific and 5 literary texts). Table 1 and Figure 1 show the differences between the first and the last quarter of 10 sentences from each of the 10 authors (Fenk-Oczlon & Fenk, 2002). Müller (2005) proved these effects in 30 texts (300 sentences) from 3 different Romanic languages.

In a running text, function words will play a dominant role in referring to what has been said in the preceding sentences, and this reference is usually done in the first part of the sentence (cf. the rule “old before new”). But this tendency is in line with similar tendencies requiring or allowing a more general cognitive explanation.

The more frequent before the less frequent

The word order in frozen conjuncts is often explained by the rule “short before long” (A) which is also applicable to our tendency “function words before content words”. But shortness is, as we know since Zipf, favoured by high token frequency. Can we, therefore, replace rule (A) by the rule

(B) “the more frequent before the less frequent”? In such a competition (Fenk-Oczlon, 1989) rule (B) was the clear winner, and it might be the winner in our wordclass-specific effects, too. This would mean an increased generality of a covering law that states a rather constant flow of linguistic information: Initial positions of a string being *per se* of higher informational content should therefore be occupied by items of high frequency and low informational content.

Table 1: Differences in the distribution of word classes

	1. quarter	4. quarter	diff.
function	3.36	2.67	sign. $p < .01$
content	2.74	3.46	sign. $p < .01$
diff.	sign. $p < .05$	sign. $p < .01$	

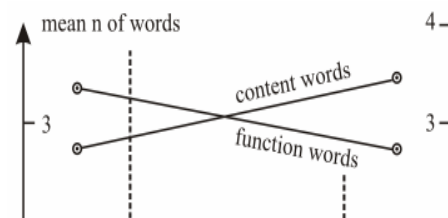


Figure 1: Function words decrease as content words increase

References

- Auer, L., Bacik, I., & Fenk, A. (2001). Die serielle Positionskurve beim Behalten echter Sätze. Paper presented at the 29. Österr. Linguistiktagung, Klagenfurt
- Fenk, A., & Fenk-Oczlon, G. (2006). Within-sentence distribution and retention of content words and function words. In P. Grzybek (Ed.), *Contributions to the Science of Text and Language. Word Length Studies and Related Issues*. Dordrecht: Springer
- Fenk-Oczlon, G., & Fenk, A. (2002). Zipf's tool analogy and word order. *Glottometrics*, 5, 22-28.
- Fenk-Oczlon, G. (1989). Word frequency and word order in freezes. *Linguistics*, 27, 517-556.
- Müller, B. (2005). *Die statistische Verteilung von Wortklassen und Wortlängen in lateinischen, italienischen und französischen Sätzen*. Doctoral dissertation, University of Klagenfurt, Austria.