

Explaining Case Syncretism: A Case Study on German Definite Articles

Remi van Trijp
Sony Computer Science Laboratory Paris
6 Rue Amyot
75005 Paris (France)
remi@csl.sony.fr

In his 1880 essay, Mark Twain famously complained that *The Awful German Language* is the 'most slipshod and systemless, and so elusive to grasp' language of all. Many linguists, even though of course not sharing Twain's harsh assessment, seem to agree that the German language is full of 'idiosyncrasies' or 'historical accidents' as the result of phonological and morphological processes (Barðdal 2009; Baerman 2009:229). One such system is the German paradigm of definite articles, which agree with their head noun in case, number and gender. These definite articles evolved from a system consisting of fairly transparent mappings between form and function in Old High German (OHG; 900-1100 AD) to a seemingly opaque and non-systematic set of case forms in New High German (NHG; see Figure 1). Numerous explanations have been offered for the decline of this system, such as phonological erosion or frequency effects, but all of them run into problems when compared to the empirical evidence (see Barðdal 2009 for a critical survey).

In this presentation, I will explore an alternative hypothesis, namely that the syncretism observed in German definite articles is **motivated** by the communicative needs and cognitive constraints of language users. This hypothesis is supported through a novel research method that 'brings back alive' the system of Old High German definite articles through a computational reconstruction in Fluid Construction Grammar (FCG; Steels 2011). The main advantage of a computational reconstruction is that it offers a fully operational and explicit way of testing all of the explanations put forward by linguists. More specifically, we can build a detailed 'performance profile' of the reconstructed language system and compare that profile to the performance of the 'synchronic' system of definite articles.

The surprising result of this comparison is that the NHG-system of definite articles outperforms its OHG predecessor in various linguistic assessment criteria, as shown in Figure 2. The NHG-system turns out to be easier to process, pronounce and acoustically distinguish than the OHG system. At the same time, the communicative power of the German language is unharmed: even though the NHG articles are less reliable as a cue for disambiguating utterances (shown in the measures 'informativity' and 'disambiguation power'), the language user does not experience this loss because the German grammar offers a sufficient amount of other cues (such as SV-agreement, number and gender) to keep communicative success at the same rate (shown in 'conditional informativity' and 'conditional disambiguation power'). A closer inspection further reveals that forms that are less efficient or that do not contribute to disambiguation in OHG correspond to the articles that have actually undergone significant changes over time.

The experiments thus show that the system of German definite articles strikes a balance between various linguistic pressures, and suggest that syncretic case forms were able to propagate in the German speech community because they were more advantageous for communication than the older forms. The case study also shows how computational reconstructions can offer valuable new contributions to the study of language.

References

- Baerman, M. 2009. Case syncretism. *The Oxford Handbook of Case*, ed. by A. Malchukov and A. Spencer, p. 219-230. Oxford: Oxford University Press.
- Barðdal, J. 2009. The development of case in Germanic. *The Role of Semantics and Pragmatics in the Development of Case*, ed. by J. Barðdal and S. Chelliah, p. 123-159. Amsterdam: John Benjamins.
- Steels, L. editor. 2011. *Design Patterns in Fluid Construction Grammar*. Amsterdam: John Benjamins.
- Wright, J. 1906. *An Old High German Primer* (2nd edition). Oxford: Clarendon Press.
- Wright, J. 1916. *A Middle High German Primer* (2nd edition). Oxford: Clarendon Press.

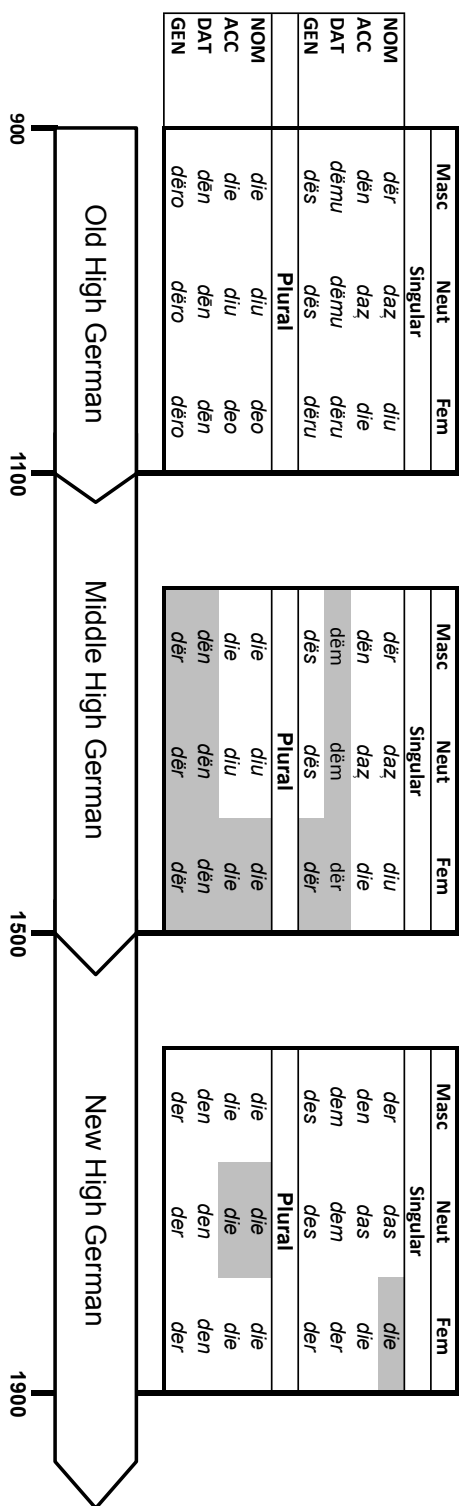


Figure 1: This Figure shows the paradigm of German definite articles in three different time periods: Old High German (900-1100; Wright 1906), Middle High German (1100-1500; Wright 1916), and New High German (from 1500 onwards). Over the course of time, the paradigm has collapsed multiple distinctive forms into syncretic articles. Grey cells indicate forms that have changed with respect to the previous time period.

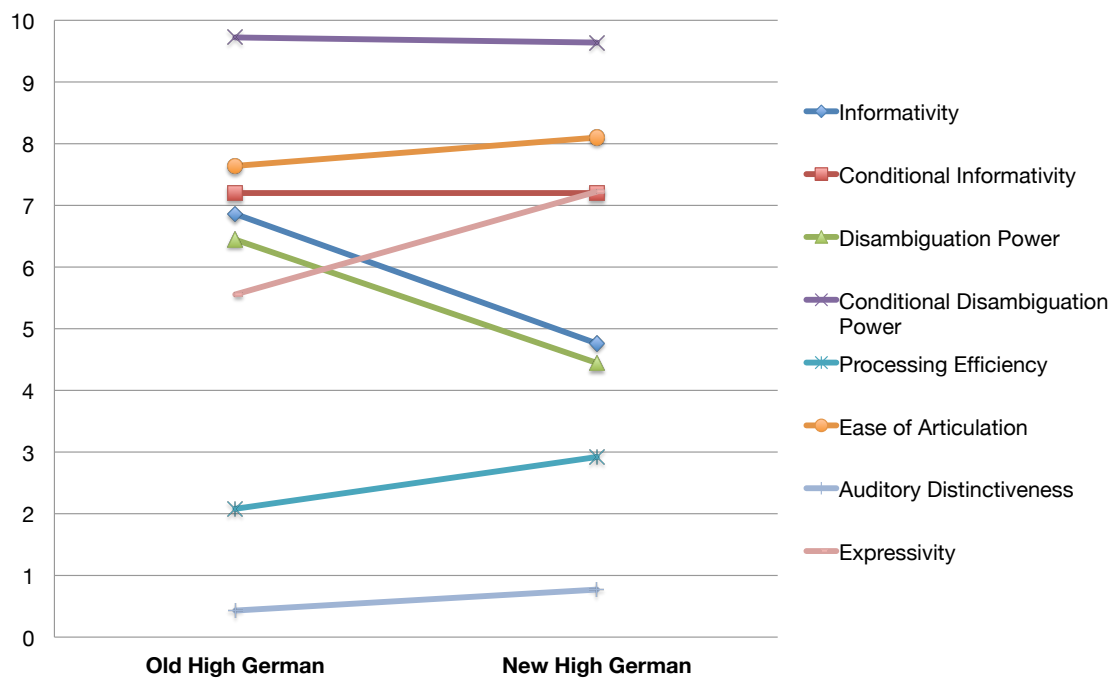


Figure 2: This Chart summarizes the main results of the experiments on a single scale. The linguistic criteria show that the overall communicative efficiency of the language (‘conditional informativity’ and ‘conditional disambiguation power’) has remained stable, while the system of definite articles has evolved to become more efficient in terms of processing, usage of memory, articulation and auditory distinctiveness.