

FITNESS LANDSCAPES IN CULTURAL LANGUAGE EVOLUTION: A CASE STUDY ON GERMAN DEFINITE ARTICLES

REMI VAN TRIJP

*Sony Computer Science Paris, 6 Rue Amyot,
Paris, 75005, France
remi@csl.sony.fr*

Computational experiments in cultural language evolution are important because they help to reveal the cognitive mechanisms and cultural processes that continuously shape and reshape the structure and knowledge of language. However, understanding the intricate relations between these mechanisms and processes can be a daunting challenge. This paper proposes to recruit the concept of *fitness landscapes* from evolutionary biology and computer science for visualizing the “linguistic fitness” of particular language systems. Through a case study on the German paradigm of definite articles, the paper shows how such landscapes can shed a new and unexpected light on non-trivial cases of language evolution. More specifically, the case study falsifies the widespread assumption that the paradigm is the accidental by-product of linguistic erosion. Instead, it has evolved to optimize the cognitive and perceptual resources that language users employ for achieving successful communication.

1. Introduction

There is a wide consensus among researchers working on cultural language evolution that language is a complex adaptive system (CAS; Steels, 2000; Beckner et al., 2009), which means that many linguistic phenomena are emergent properties of the complex interplay between cognitive, perceptual and social constraints on language users as they engage with each other in local interactions. Computational models are an important tool of the CAS approach because they demonstrate the consequences of these complex dynamics on a scale and time course that is unimaginable with human subjects (e.g. Baronchelli, Chater, Pastor-Satorras, & Christiansen, 2012; Beuls & Steels, 2013; Fay & Ellison, 2013).

As noted by Coupé, Shuai, and Gong (2013, p. 121), these models have recently shifted from investigating the evolution of simple one-to-one mappings to more realistic linguistic communication systems. However, this increase in sophistication comes with the need for new methods of deciphering what is going on in the models. This paper proposes to recruit the concept of *fitness landscapes* from evolutionary biology and computer science, and to apply it to cultural language evolution. Through a case study on German, it will illustrate how fitness landscapes can help to solve long-standing puzzles in language evolution.

1.1. *Linguistic Selectionism and Fitness Landscapes*

This paper subscribes to the theory of cultural language evolution as proposed by Steels (2011). One of the cornerstones of the theory is the biologically-inspired mechanism of *linguistic selectionism*, which involves two kinds of processes:

1. Processes that create linguistic variation in a population.
2. Processes that select variants to become dominant conventions.

The theory thus hypothesizes that linguistic variants may be selected if they increase the “linguistic fitness” of a language. In order to test this hypothesis, we need a good understanding of the relations between the space of possible languages (given the linguistic variants at hand) and their linguistic fitness. Fitness landscapes can offer a valuable tool for visualizing these relations.

Fitness landscapes were first introduced in evolutionary biology by Wright (1932), who visualized the space of possible gene combinations as a field, where the “height” of the field corresponds to the fitness of a genotype (in this case its replication rate). The problem of evolution can then be conceptualized as “a mechanism by which the species may continually find its way from lower to higher peaks in such a field” (p. 358–359). This idea soon became generalized to computer science, where it is used for evolutionary optimization (Richter, 2010).

Applying this approach to cultural language evolution, we can thus generate a space of possible languages through the processes that create variation in a population. Next, we can define a “fitness function” that evaluates how well each possible language is adapted for communication. As I will explain in section 2, this fitness function aggregates several linguistic selection criteria that language users experience when engaging in linguistic interactions, such as communicative success, disambiguation power and processing efficiency.

1.2. *A Linguistic Puzzle*

The best way to demonstrate the validity of the proposed method is to ground it into a concrete case study of language evolution that looks like an outright challenge to the theory. The experiments in this paper focus on the evolution of the German paradigm of definite articles, which is widely considered as one of the most intriguing puzzles in linguistics.

The big mystery of the paradigm is as follows. Each article is marked for three dimensions: case (nominative, accusative, dative and genitive), number (singular and plural) and gender (masculine, neuter and feminine). A fully transparent paradigm would thus consist of 24 distinct articles, one for each combination of these dimensions. As illustrated in Figure 1, however, the actual German paradigm has been evolving further and further away from such transparency. The Figure displays the paradigm as it appears in Old High German (OHG, 900–1100; Wright, 1906), Middle High German (MHG, 1100–1500; Wright, 1916) and New

High German (NHG, from 1500 onwards). Gray cells in the paradigm indicate when distinct forms collapsed into “syncretic forms” (i.e. where the same form covers multiple cells). One striking observation is that the OHG-paradigm counts twice as many distinct forms than the current system. For instance, where NHG has one syncretic form *die* for nominative and accusative plurals, OHG had a three-way gender distinction between masculine, neuter and feminine. So why did the speakers of German allow this more transparent system to crumble down to its current form? What can this case study tell us about language evolution?

	Masc	Neut	Fem
Singular			
NOM	<i>dër</i>	<i>daz</i>	<i>diu</i>
ACC	<i>dën</i>	<i>daz</i>	<i>die</i>
DAT	<i>dëmu</i>	<i>dëmu</i>	<i>dëru</i>
GEN	<i>dës</i>	<i>dës</i>	<i>dëru</i>
Plural			
NOM	<i>die</i>	<i>diu</i>	<i>deo</i>
ACC	<i>die</i>	<i>diu</i>	<i>deo</i>
DAT	<i>dën</i>	<i>dën</i>	<i>dën</i>
GEN	<i>dëro</i>	<i>dëro</i>	<i>dëro</i>

	Masc	Neut	Fem
Singular			
<i>dër</i>	<i>daz</i>	<i>diu</i>	
<i>dën</i>	<i>daz</i>	<i>die</i>	
<i>dëm</i>	<i>dëm</i>	<i>dër</i>	
<i>dës</i>	<i>dës</i>	<i>dër</i>	
Plural			
<i>die</i>	<i>diu</i>	<i>die</i>	
<i>die</i>	<i>diu</i>	<i>die</i>	
<i>dën</i>	<i>dën</i>	<i>dën</i>	
<i>dër</i>	<i>dër</i>	<i>dër</i>	

	Masc	Neut	Fem
Singular			
<i>der</i>	<i>das</i>	<i>die</i>	
<i>den</i>	<i>das</i>	<i>die</i>	
<i>dem</i>	<i>dem</i>	<i>der</i>	
<i>des</i>	<i>des</i>	<i>der</i>	
Plural			
<i>die</i>	<i>die</i>	<i>die</i>	
<i>die</i>	<i>die</i>	<i>die</i>	
<i>den</i>	<i>den</i>	<i>den</i>	
<i>der</i>	<i>der</i>	<i>der</i>	

900	1100	1500	1900
Old High German		Middle High German	New High German

Figure 1. This Figure shows the system of German definite articles in three different time periods: Old High German, Middle High German and New High German.

The answers suggested in the literature contradict each other. The most popular explanation is that non-systematic syncretism is simply a historical accident caused by phonological and morphological changes (Baerman, 2009). However, new syncretic forms do not randomly enter the paradigm. In fact, the data show such strong tendencies that Hawkins (2004, p. 63–86) argues that the increase in syncretism follows a universal hierarchy for case (nom > acc > dat) and gender (masc, fem > neut), whereby distinctions in lower dimensions are lost before distinctions in higher dimensions (e.g. dative before accusative). However, none of these answers can explain why the paradigm declined so rapidly from OHG to MHG, but then remained relatively stable for more than five centuries despite the availability of simpler variants in the Low German dialects (Shrier, 1965).

2. Experimental Set-Up

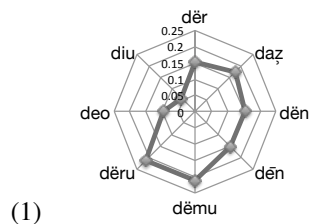
The predictions of the aforementioned hypotheses can be tested through computational experiments. The experiments reported here start with a bidirectional processing model of German in Fluid Construction Grammar (see van Trijp, 2011, 2013 and www.fcg-net.org/demos/german-case/ for an online demo).

2.1. Processes that Create Variation

While the rest of the grammar remains fixed, the experiments generate a space of possible variations by changing the paradigm of definite articles, after which each variation can be evaluated in terms of its linguistic fitness. Since we are interested in explaining how the OHG-paradigm may have evolved into its current form, the experiment starts with a computational reconstruction of the OHG-system.

Variation is caused by two pairs of widely attested processes. The first pair consists of phonological processes that either “erode” or “expand” forms. For instance, the process *apocope* is a force of erosion in which the last sound of a word is dropped, e.g. [dɛmu] → [dɛm]. Phonological expansion works in the other direction, where new sounds can be attached to a form. Phonological processes can leave the distinctions of a paradigm intact, or they can cause forms to collapse.

The second set of processes consists of attraction and repulsion. Attraction may happen when two forms are phonologically close to each other. In this case, the form with the highest type frequency will attract the other form and effectively usurp its functions. For instance, example (1) illustrates the phonological distance from the OHG-article *die* (in the center of the spider chart) to the other articles of the OHG-paradigm. As can be seen, *die* [diə] is phonologically closest to *diu* [diu] and *deo* [deo]. There is thus a high probability that one of these forms will attract the others, and thereby increase the syncretism of the paradigm.



The opposite of attraction is repulsion. Repulsion occurs when one of the cells that is covered by a syncretic form breaks free and gets its own distinct form. Repulsion therefore increases the transparency of a paradigm.

2.2. Linguistic Fitness

Each generated variation is tested for both parsing and production against a corpus of 360 declarative utterances, which exposes the variation to all possible combinations of case, number and gender in transitive and ditransitive patterns. In parsing, the model tries to disambiguate the argument structure that underlies the utterance (i.e. ‘who did what to whom’).

The linguistic fitness of a variation is evaluated as a weighted average of four measures: disambiguation power (based on the amount of utterances that the language is able to disambiguate), processing efficiency (based on the amount of

primitive operations that the language faculty has to perform in order to parse or produce sentences), ease of articulation (based on the amount of movements that articulators such as the lips and tongue have to make when pronouncing sounds) and acoustic distinctiveness (based on the phonological distance between words). All of these measures are formally defined and discussed by van Trijp (2013). Disambiguation power weighs 85%, processing efficiency 13%, and ease of articulation and acoustic distinctiveness weigh 1% each. These weights were experimentally obtained through standard feature weighting methods in order to find the best fit of the model on the empirical data on the evolution of German articles.

3. Experimental Results

The resulting fitness landscape for German is shown in Figure 2. The X-axis shows the average length of the articles in each variation, and the Y-axis shows the amount of distinct case forms (maximum 18 because the genitive case is ignored in the experiments because it is not a core argument role). The height of the landscape corresponds to linguistic fitness.

The first remarkable result is that, perhaps counterintuitively, the language does not require a lot of distinct articles for reaching a high fitness value: two distinct forms already push the fitness beyond 80%, and all other variations are very close to each other with linguistic fitness values between 85 and 92%.

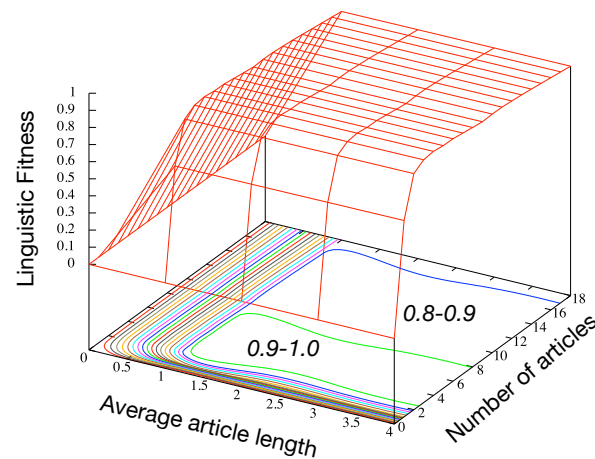


Figure 2. The fitness landscape for German definite articles. The paradigm is quite robust to change and already reaches a linguistic fitness of more than 80% with only two articles. The contour map beneath the landscape shows that the fittest variants have paradigms of 3 to 7 articles.

The large plateau of high fitness values indicates that the language is remarkably robust for changes in its case system, which may explain the enormous variation in the case systems of German dialects (Shrier, 1965). Nevertheless, the contour map beneath the landscape suggests that some variants on the plateau still have higher fitness values, with the best paradigms consisting of three to seven articles.

Figure 3 zooms in on all values above 60% for linguistic fitness, disambiguation power and processing efficiency; and it marks where Old and New High German are situated in the landscape. The Figure reveals that the plateau is in fact what cyclists call a “false flat”, meaning that there is a low-gradient climb. The results thus confirm that the evolution of the German paradigm of definite articles can be conceptualized as an upwards movement in a fitness landscape without any intermediate “valleys” that need to be bridged.

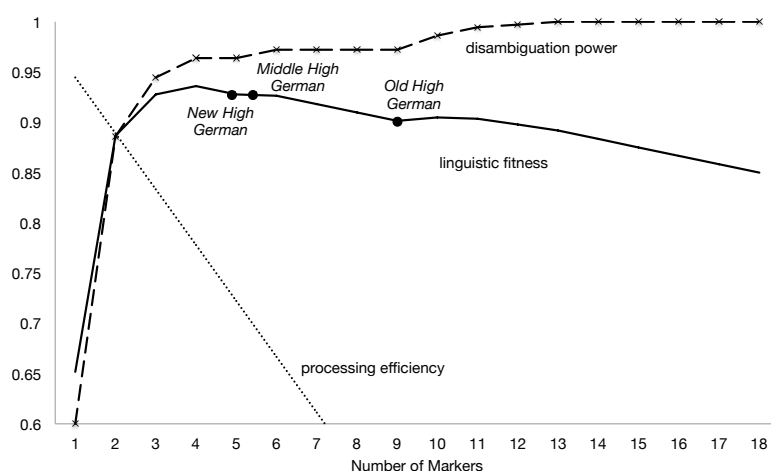


Figure 3. This Figure shows the processing efficiency of different article systems (ranging from a fully syncretic paradigm to a maximally transparent one), their impact on the language’s disambiguation power and its linguistic fitness. The Y-axis zooms in on values of more than 60%. The Figure shows that the language achieves a higher fitness score with the NHG-paradigm than with the OHG-paradigm.

The smooth evolution towards NHG is also confirmed when looking at the fittest linguistic variants generated through the processes of variation. For instance, just as what happened in German, early variants typically lose the gender distinctions in nominative and accusative plural forms (i.e. *die* vs. *deo* vs. *diu*). Collapsing these forms significantly increases the processing efficiency of the paradigm. Since gender and number are also marked on German nouns, the loss of the gender distinction does not affect the language’s disambiguation power.

When looking at the variants that evolve in the other direction – namely increasing the number of distinct forms through repulsion – we first see a re-establishment of the nominative-accusative distinction for singular-neuter nouns and for all plural forms. Twelve articles then suffice for achieving maximum disambiguation power if they are arranged in a proper way. However, each additional form comes at the cost of processing efficiency, which explains why smaller paradigms are preferred over larger ones.

4. Discussion and Conclusion

The experimental results of the case study provide support for the theory of cultural language evolution as proposed by Steels (2011). More specifically, the theory predicts that linguistic variants may become dominant in a population if they offer a selective advantage for communication. The case study has demonstrated that linguistic selectionism can indeed explain even the seemingly erratic evolution of German definite articles. Moreover, the fittest variants that were generated by the computational model correspond to those changes that actually happened in the German language.

The experiments therefore cast serious doubt on the alternative explanations for the evolution of German definite articles as discussed in section 1.2. First, all systematic trends in the increase of syncretism emerge as a side-effect of linguistic selection. The model thus shows that there is no need to posit universal case or gender hierarchies, as hypothesized by Hawkins (2004). Secondly, it is very unlikely that the syncretic forms of the paradigm evolved as a historical accident rather than as a result of selection. Without selection, all variants have an equal chance of “survival”, which results in an explosion of the space of possible variations. In such a large space, it is highly improbable that the language has followed a path that consistently moves upwards in the fitness landscape by sheer accident.

In sum, this paper has demonstrated that fitness landscapes are a valuable tool for computational experiments in cultural language evolution. Fitness landscapes allow researchers to visualize the relations between a space of possible languages and their linguistic fitness, which helps to understand the complex dynamics between cognitive, perceptual and social forces that shape and reshape the structure and knowledge of language.

Acknowledgements

The work presented in this paper was funded by and conducted at Sony CSL Paris. The idea to apply fitness landscapes to cultural language evolution first came up in discussions with Luc Steels (Sony CSL Paris - CSIC-UPF Barcelona) and Pieter Wellens (VUB Artificial Intelligence Laboratory Brussels) in the Summer of 2012, and was later also suggested to me by Robert C. Berwick (MIT) and Dan Dediu (Max Planck Institute Nijmegen) at the Workshop on Language, Cognition and Computational Models (Paris, 28-29 May 2013).

References

- Baerman, M. (2009). Case syncretism. In A. Malchukov & A. Spencer (Eds.), *The Oxford handbook of case* (pp. 219–230). Oxford: OUP.
- Baronchelli, A., Chater, N., Pastor-Satorras, R., & Christiansen, M. (2012). The biological origin of linguistic diversity. *PLoS ONE*, 7(10), e48029.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., & Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning*, 59(s1), 1-26.
- Beuls, K., & Steels, L. (2013). Agent-based models of strategies for the emergence and evolution of grammatical agreement. *PLoS ONE*, 8(3), e58960.
- Coupé, C., Shuai, L., & Gong, T. (2013). Review of the 9th international conference on the evolution of language (evolang9). *Biolinguistics*, 7, 112–131.
- Fay, N., & Ellison, T. M. (2013). The cultural evolution of human communication systems in different sized populations: Usability trumps learnability. *PLoS ONE*, 8(8), e71781.
- Hawkins, J. (2004). *Efficiency and complexity in grammars*. Oxford: OUP.
- Richter, H. (2010). Evolutionary optimization and dynamic fitness landscapes. from reaction-diffusion systems to chaotic CML. In I. Zelinka, S. Celikovský, H. Richter, & G. Chen (Eds.), *Evolutionary algorithms and chaotic systems* (pp. 409–446). Berlin: Springer.
- Shrier, M. (1965). Case systems in German dialects. *Language*, 41(3), 420–438.
- Steels, L. (2000). Language as a complex adaptive system. In M. Schoenauer (Ed.), *Proceedings of ppsn vi: Lectur notes in computer science* (pp. 17–26). Berlin: Springer-Verlag.
- Steels, L. (2011). Modeling the cultural evolution of language. *Physics of Life Reviews*, 8(4), 339–356.
- van Trijp, R. (2011). Feature matrices and agreement: A case study for German case. In L. Steels (Ed.), *Design patterns in Fluid Construction Grammar* (p. 205-235). Amsterdam: John Benjamins.
- van Trijp, R. (2013). Linguistic assessment criteria for explaining language change: A case study on syncretism in German definite articles. *Language Dynamics and Change*, 3(1), 105–132.
- Wright, J. (1906). *An Old High German primer* (2nd ed.). Oxford: Clarendon Press.
- Wright, J. (1916). *A Middle High German primer* (2nd ed.). Oxford: Clarendon Press.
- Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In *Proceedings of the sixth international congress on genetics* (pp. 355–366).