Not as Awful as it Seems: Explaining German Case through Computational Experiments in Fluid Construction Grammar

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

Abstract

German case syncretism is often assumed to be the accidental by-product of historical development. This paper contradicts this claim and argues that the evolution of German case is driven by the need to optimize the cognitive effort and memory required for processing and interpretation. This hypothesis is supported by a novel kind of computational experiments that reconstruct and compare attested variations of the German definite article paradigm. The experiments show how the intricate interaction between those variations and the rest of the German 'linguistic landscape' may direct language change.

1 Introduction

In his 1880 essay, Mark Twain famously complained that *The awful German Language* is the most "slipshod and systemless, and so slippery and elusive to grasp" language of all. A brief look at the literature on the German case system seems to provide sufficient evidence for instantly agreeing with the American author. But what if the German case system were not the accidental by-product of diachronic changes as is often assumed? Are there linguistic forces that are not yet fully appreciated in the field, but which may explain the German case paradigm?

This paper demonstrates that there indeed are such forces through a case study on German definite articles. The experiments 'reconstruct' deep language processing models for different variants of this paradigm, and show how the 'linguistic landscape' of German has allowed its speakers to reduce their definite article system without loss in efficiency for processing and interpretation.

2 The Problem of German Case

German articles, adjectives and nouns are marked for gender, number and case through morphological inflection, as illustrated for definite articles in Table 1.

Case	SG-M	SG-F	SG-N	PL
NOM	*der*	*die*	*das*	*die*
ACC	*den*	*die*	*das*	*die*
DAT	*dem*	*der*	*dem*	*den*
GEN	*des*	*der*	*des*	*der*

Table 1: German definite articles.

The system is notorious for its *syncretism* (i.e. the same form can be mapped onto different functions), a riddle that has fascinated many formal and historical linguists looking for explanations.

2.1 Historical Linguistics

Studies in historical linguistics and grammaticalization often propose the following three forces to explain syncretism (Heine and Kuteva, 2005, p. 148):

1. The formal distinction between case markers is lost through phonological changes.
2. One case takes over the functional domain of another case and replaces it.
3. A case marker disappears and its functions are usurped by another marker.

Syncretism is thus considered as the *accidental* by-product of such forces, and German case syncretism is typically analyzed according to these lines (Barðdal, 2009; Baerman, 2009, p. 229). However, these forces are not explanatory: they only describe *what* has happened, but not *why*.

Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics, pages 829–839,
Avignon, France, April 23 - 27 2012.

Another problem for the 'syncretism by accident' hypothesis is the fact that the collapsing of case forms is not randomly distributed over the whole paradigm as would be expected. Hawkins (2004, p. 78) observes that instead there is a systematic tendency for 'lower' cells in the paradigm (e.g. genitive; Table 1) to collapse before cells in 'higher' positions (e.g. nominative) do so.

2.2 Formal Linguistics

Many hidden effects of verbal linguistic theories can be uncovered through explicit formalizations. Unfortunately, formal linguists also typically distinguish between 'systematic' and 'non-systematic' syncretism when analyzing German case. For instance, in his review of a number of studies on German (a.o. Bierwisch, 1967; Blevins, 1995; Wiese, 1996; Wunderlich, 1997), Müller (2002) concludes that none of these approaches is able to rule out accidental syncretism.

There is however one major stone that has been left unturned by formal linguists: processing. Most formal theories, such as HPSG (Ginzburg and Sag, 2000), assume a strict division between 'competence' and 'performance' and therefore represent linguistic knowledge in a purely declarative, process-independent way (Sag and Wasow, 2011). While such an approach may be desirable from a 'mathematical' point of view, it puts the burden of efficient processing on the shoulders of computational linguists, who have to develop more intelligent interpreters.

One example of the gap between description and computational implementation is *disjunctive feature representation*, which became popular in feature-based grammar formalisms in the 1980s (Karttunen, 1984). Disjunctions allow an elegant notation for multiple feature values, as illustrated in example 1 for the German definite article *die*, which is either assigned nominative or accusative case, and which is either feminine-singular or plural. The feature structure (adopted from Karttunen, 1984, p. 30) represents disjunctions by enclosing the alternatives in curly brackets ({ }).

(1)
$$\begin{bmatrix} \text{AGREEMENT} & \left\{ \begin{matrix} \begin{bmatrix} \text{GENDER} & f \\ \text{NUM} & sg \end{bmatrix} \\ \begin{bmatrix} \text{NUM} & pl \end{bmatrix} \end{matrix} \right\} \\ \text{CASE} & \left\{ nom\ acc \right\} \end{bmatrix}$$

However, it is a well-established fact that disjunctions are computationally expensive, which is illustrated in the top of Figure 1. This Figure shows the search tree of a small grammar when parsing the utterance *Die Kinder gaben der Lehrerin die Zeichnung* ('the children gave the drawing to the (female) teacher'), which is unambiguous to German speakers. As can be seen in the Figure, the search tree has to explore several branches before arriving at a valid solution. Most of the splits are caused by disjunctions. For example, when a determiner-noun construction specifies that the case features of the definite article *die* (nominative or accusative) and the noun *Kinder* ('children'; nominative, accusative or genitive) have to unify, the search tree splits into two hypotheses (a nominative and an accusative reading) even though for native speakers of German, the syntactic context unambiguously points to a nominative reading (because it is the only noun phrase that agrees with the main verb).

It should be no surprise, then, that a lot of work has focused on processing disjunctions more efficiently (e.g. Carter, 1990; Ramsay, 1990). As observed by Flickinger (2000), however, most of these studies implicitly assume that the grammar representation has to remain unchanged. He then demonstrates through computational experiments how a different representation can directly impact efficiency, and argues that revisions of the grammar for efficiency should be discussed more thoroughly in the literature.

The impact of representation on processing is illustrated at the bottom of Figure 1, which shows the performance of a grammar that uses the same processing technique for handling the same utterance, but a different representation than the disjunctive grammar. As can be seen, the alternative grammar (whose technical details are disclosed further below) is able to parse the German definite articles without tears, and the resulting search tree arguably better reflects the actual processing performed by native speakers of German.

2.3 Alternative Hypothesis

The effect of processing-friendly representations on search suggests that answers for the unsolved problems concerning case syncretism have to be sought in performance. This paper therefore rejects the processing-independent approach and explores the alternative hypothesis, following

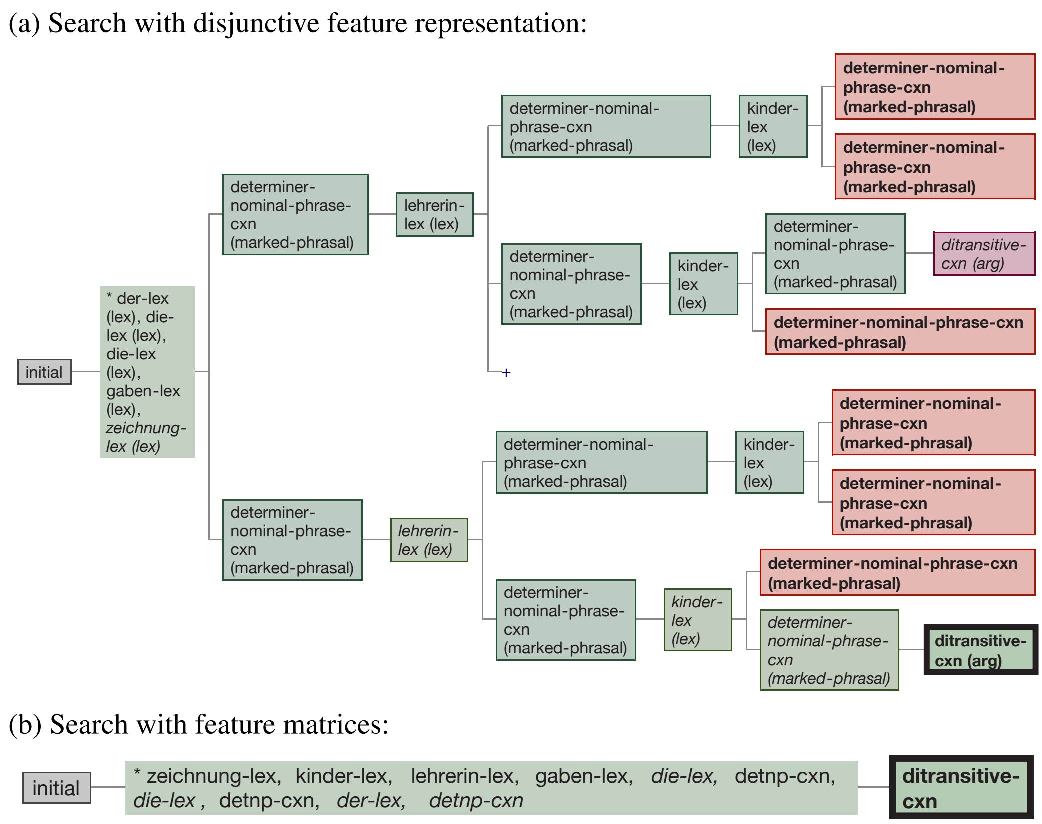


Figure 1: The representation of linguistic information has a direct impact on processing efficiency. The top figure shows a search tree when parsing the unambiguous utterance *Die Kinder gaben der Lehrerin die Zeichnung* ('The children gave the drawing to the (female) teacher') using disjunctive feature representation. The bottom figure shows the search tree using distinctive feature matrices. Labels in the boxes show the names of the applied constructions; boxes with a bold border are successful end nodes. Both grammars have been implemented in Fluid Construction Grammar (FCG; Steels, 2011, 2012a) and are processed using a standard depth-first search algorithm (Bleys et al., 2011) and general unification (*without* optimization for particular types or data structures; Steels and De Beule, 2006; De Beule, 2012). The utterance is assumed to be segmented into words. Interested readers can explore the Figure through an interactive web demonstration at `http://www.fcg-net.org/demos/design-patterns/07-feature-matrices/`.

Steels (2004, 2012b), that grammar evolves in order to optimize communicative success by dampening the search space in linguistic processing and reducing the cognitive effort needed for interpretation, while at the same time minimizing the resources required for doing so. More specifically, this paper explores the following claims:

1. The German definite article system can be processed as efficiently as its Old High German predecessor, which had less syncretism.

2. The presence of other grammatical structures have made it possible to reduce the definite article paradigm without increasing the cognitive effort needed for disambiguating the argument structures that underly German utterances.

3. The decrease of cue-reliability of case for disambiguation encourages the emergence of competing systems (such as word order).

The hypothesis is substantiated through computational experiments that reconstruct three different variants of the German definite article system (the current system, its Old High German predecessor, Wright, 1906; and the Texas German dialect system, Boas, 2009a,b) and compare their performance in terms of processing efficiency and cognitive effort in interpretation.

3 Operationalizing German Case

An adequate operationalization of German case requires a bidirectional grammar (for parsing and production) and easy access to linguistic processing data. All experiments reported in this paper have therefore been implemented in Fluid Construction Grammar (FCG; Steels, 2011, 2012a), a unification-based grammar formalism that comes equipped with an interactive web interface and monitoring tools (Loetzsch, 2012). A second advantage of FCG is that it features *strong* bidirectionality: the FCG-interpreter can achieve both parsing and production using the same linguistic inventory. Other feature structure platforms, such as the lkb-system (Copestake, 2002), require a separate parser and generator for formalizing bidirectional grammars, which make them less suited for substantiating the claims of this paper.

3.1 Distinctive Feature Matrix

German case has become the litmus test for demonstrating how well a feature-based grammar formalism copes with multifunctionality, especially since Ingria (1990) provocatively stated that unification is not the best technique for handling it. People have gone to great lengths to counter Ingria's claim, especially within the HPSG framework (e.g. Müller, 1999; Daniels, 2001; Sag, 2003), and various formalizations have been offered for German case (Heinz and Matiasek, 1994; Müller, 2001; Crysmann, 2005). However, these proposals either do not succeed in avoiding inefficient disjunctions or they require a complex double type hierarchy (Crysmann, 2005).

The experiments in this paper use a more straightforward solution, called a *distinctive feature matrix*, which is based on an idea that was first explored by Ingria (1990) and of which a variation has recently also been proposed for Lexical Functional Grammar (Dalrymple et al., 2009). Instead of treating case as a single-valued feature, it can be represented as an array of features, as shown for the definite article *die* (ignoring the genitive case for the time being):

(2) die: $$\begin{bmatrix} \text{CASE} & \begin{bmatrix} \text{nom} & ?nom \\ \text{acc} & ?acc \\ \text{dat} & - \end{bmatrix} \end{bmatrix}$$

The `case` feature includes a paradigm of three cases (`nom`, `acc` and `dat`), whose values can either be '+' or '–', or left unspecified through a variable (indicated by a question mark). The two variables `?nom` and `?acc` indicate that *die* can potentially be assigned nominative or accusative case, the value '–' for dative means that `die` cannot be assigned dative case. We can do the same for *Kinder* ('children'), which can be nominative or accusative, but not dative:

(3) Kinder: $$\begin{bmatrix} \text{CASE} & \begin{bmatrix} \text{nom} & ?nom \\ \text{acc} & ?acc \\ \text{dat} & - \end{bmatrix} \end{bmatrix}$$

As demonstrated in Figure 1, disjunctive feature representation would cause a split in the search tree when unifying *die* and *Kinder*. Using a feature matrix, however, the choice between a nominative and accusative reading can simply be postponed until enough information from the rest of the utterance is available. Unifying *die* and *Kinder* yields the following feature structure:

(4) die Kinder: $$\begin{bmatrix} \text{CASE} & \begin{bmatrix} \text{nom} & ?nom \\ \text{acc} & ?acc \\ \text{dat} & - \end{bmatrix} \end{bmatrix}$$

3.2 A Three-Dimensional Matrix

The German case paradigm is obviously more complex than the examples shown so far. Let's consider Table 1 again, but this time we replace every cell in the table by a variable. This leads to the following feature matrix for the German definite articles:

Case	**SG-M**	**SG-F**	**SG-N**	**PL**
?NOM	*?n-s-m*	*?n-s-f*	*?n-s-n*	*?n-pl*
?ACC	*?a-s-m*	*?a-s-f*	*?a-s-n*	*?a-pl*
?DAT	*?d-s-m*	*?d-s-f*	*?d-s-n*	*?d-pl*
?GEN	*?g-s-m*	*?g-s-f*	*?g-s-n*	*?g-pl*

Table 2: A distinctive feature matrix for German case.

Each cell in this matrix represents a specific feature bundle that collects the features case, number, and person. For example, the variable `?n-s-m` stands for nominative singular masculine. Note that also the cases themselves have their own variable (`?nom`, `?acc`, `?dat` and `?gen`). This allows us to single out a specific dimension of the matrix for constructions that only care about case distinctions, but abstract away from gender or number. Each linguistic item fills in as much information as possible in this case matrix. For example, Table 3 shows how the definite article *die* underspecifies its potential values and rules out all other options through '–'.

Case	SG-M	SG-F	SG-N	PL
?NOM	–	*?n-s-f*	–	*?n-pl*
?ACC	–	*?a-s-f*	–	*?a-pl*
–	–	–	–	–
–	–	–	–	–

Table 3: The feature matrix of *die*.

The feature matrix of *Kinder* ('children'), which underspecifies for nominative, accusative and genitive, is shown in Table 4. Notice, however, that the same variable names are used for both the column that singles out the case dimension as for the column of the plural feature bundles.

Case	SG-M	SG-F	SG-N	PL
?n-pl	–	–	–	*?n-pl*
?a-pl	–	–	–	*?a-pl*
–	–	–	–	–
?g-pl	–	–	–	*?g-pl*

Table 4: The feature matrix of *Kinder* ('children').

Unification of *die* and *Kinder* can exploit these variable 'equalities' for ruling out a singular value of the definite article. Likewise, the matrix of *die* rules out the genitive reading of *Kinder*, as illustrated in Table 5.

Case	SG-M	SG-F	SG-N	PL
?n-pl	–	–	–	*?n-pl*
?a-pl	–	–	–	*?a-pl*
–	–	–	–	–
–	–	–	–	–

Table 5: The feature matrix of *die Kinder*.

Argument structure constructions (Goldberg, 2006), such as the ditransitive, can then later assign either nominative or accusative case. The main advantage of feature matrices is that linguistic search only has to commit to specific feature-values once sufficient information is available, so the search tree only splits when there is an actual ambiguity. Moreover, they can be handled using standard unification. Interested readers can consult van Trijp (2011) for a thorough description of the approach, as well as a discussion on how the FCG implementation differs from Ingria (1990) and Dalrymple et al. (2009).

4 Experiments

This section describes the experimental set-up and discusses the experimental results.

4.1 Three Paradigms

The experiments compare three different variants of the German definite article paradigm.

Standard German. The Standard German paradigm has been illustrated in Table 1 and its operationalization has been shown in section 3.2. The paradigm has been inherited without significant changes from Middle High German (1050-1350; Walshe, 1974) and features six different forms.

Old High German. The Old High German paradigm is the direct predecessor of the current paradigm of definite articles. It contained at least twelve distinct forms (depending on which variation is taken) that included gender distinctions in plural (Wright, 1906, p. 67). It also included one definite article that marked the now extinct instrumental case, which is ignored in this paper. The variant of the Old High German paradigm that has been implemented in the experiments is summarized in Table 6.

Case	Singular		
	M	**F**	**N**
NOM	*dër*	*diu*	*daȥ*
ACC	*dën*	*die*	*daȥ*
DAT	*dëmu*	*dëru*	*dëmu*
GEN	*dës*	*dëra*	*dës*
	Plural		
	M	**F**	**N**
NOM	*die*	*deo*	*diu*
ACC	*die*	*deo*	*diu*
DAT	*dēm*	*dēm*	*dēm*
GEN	*dëro*	*dëro*	*dëro*

Table 6: The Old High German definite article system.

Texas German. The third variant is an American-German dialect called Texas German (Boas, 2009a,b), which evolved a two-way case distinction between nominative and oblique. This type of case system, in which the accusative and dative case have collapsed, is also a common evolution in the Low German dialects (Shrier, 1965). The implemented paradigm of Texas German is shown in Table 7.

Case	SG-M	SG-F	SG-N	PL
NOM	*der*	*die*	*das*	*die*
ACC/DAT	*den*	*die*	*den*	*die*

Table 7: The Texas German definite article system.

4.2 Production and Parsing Tasks

Each grammar is tested as to how efficiently it can produce and parse utterances in terms of cognitive effort and search (see section 4.3). There are three basic *types* of utterances:

1. Ditransitive: NOM – Verb – DAT – ACC
2. Transitive (a): NOM – Verb – ACC
3. Transitive (b): NOM – Verb – DAT

The argument roles are filled by noun phrases whose head nouns always have a distinct form for singular and plural (e.g. *Mann* vs. *Männer*; 'man' vs. 'men'), but that are unmarked for case. The combinations of arguments is always unique along the dimensions of number and gender, which yields 216 unique utterance types for the ditransitive as follows:

	NOM.S.M	V	DAT.S.M	ACC.S.M
	NOM.S.M	V	DAT.S.F	ACC.S.M
(5)	NOM.S.M	V	DAT.S.N	ACC.S.M
	NOM.S.M	V	DAT.PL.M	ACC.S.M
			etc.	

In transitive utterances, there is an additional distinction based on animacy for noun phrases in the Object position of the utterance, which yields 72 types in the NOM-ACC configuration and 72 in the NOM-DAT configuration. Together, there are 360 unique utterance types. As can be gleaned from the utterance types, the genitive case is not considered by the experiments, as the genitive is not part of basic German argument structures and it has almost disappeared in most dialects of German (Shrier, 1965).

In production, the grammar is presented with a meaning that needs to be verbalized into an utterance. In parsing, the produced utterance has to be analyzed back into a meaning. Every utterance is processed using a full search, that is, all branches and solutions are calculated.

The experiments exploit types because there are three different language systems, hence it is impossible to use a single, real corpus and its token frequencies. It would also be unwarranted to use different corpora because corpus-specific biases would distort the comparative results. Secondly, as the experiments involve models of *deep* language processing (as opposed to stochastic models), the use of types instead of tokens is justified in this phase of the research: the first concern of precision-grammars is descriptive adequacy, for which types are a more reliable source. Obviously, the effect of token frequency needs to be examined in future research.

4.3 Measuring Cognitive Effort

The experiments measure two kinds of cognitive effort: syntactic search and semantic ambiguity.

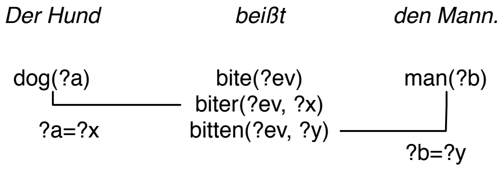
Search. The search measure counts the number of branches in the search process that reach an end node, which can either be a possible solution or a dead end (i.e. no constructions can be applied anymore). Duplicate nodes (for instance, nodes that use the same rules but in a different order) are not counted. The search measure is then used as a 'sanity check' to verify whether the three different paradigms can be processed with the same efficiency in terms of search tree length, as hypothesized by this paper. More specifically, the following conditions have to be met:

1. In production, there should only be one branch.
2. In parsing, search has to be equal to the semantic effort.

The single branch constraint in production checks whether the definite articles are sufficiently distinct from one another. Since there is no ambiguity about which argument plays which role in the utterance, the grammar should only come up with one solution. In parsing, the number of branches has to correspond to 'real' semantic ambiguities and not create additional search, as argued in section 2.2.

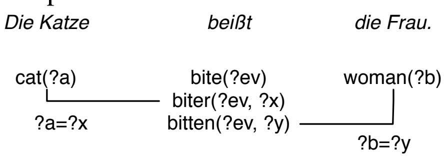
Semantic Ambiguity. Semantic ambiguity equals the number of possible interpretations of an utterance. For instance, the utterance *Der Hund beißt den Mann* 'the dog bites the man' is unambiguous in Modern High German, since *der Hund* can only be nominative singular-masculine, and *den Mann* can only be accusative masculine-singular. There is thus only one possible interpretation in which the dog is the biter and the man is being bitten, illustrated as follows using a logic-based meaning representation (also see Steels, 2004, for this operationalization of cognitive effort):

(6) Interpretation 1:

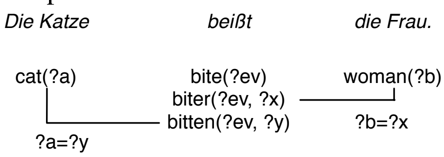


However, an utterance such as *die Katze beißt die Frau* 'the cat bites the woman' is ambiguous because *die* has both a nominative and accusative singular-feminine reading:

(7) a. Interpretation 1:



b. Interpretation 2:



Here, German speakers are likely to use word order, intonation and world knowledge (i.e. cats are more likely to bite a person than the other way round) for disambiguating the utterance.

4.4 Experimental Parameters

The experiments (E1-E4) concern the cue-reliability of the definite articles for disambiguating event structure. In all experiments, the different grammars can exploit the case-number-gender information of definite articles, and also the gender and number specifications of nouns, and the syntactic valence of verbs. For instance, the noun form *Frauen* 'women' is specified as plural-feminine, and verbs like *helfen* 'to help' are specified to take a dative object, whereas verbs like *finden* 'to find' take an accusative object. In other experiments, different combinations of grammatical cues become available or not:

Cue	E1	E2	E3	E4
SV-agreement		+		+
Selection restrictions			+	+

SV-agreement restricts the subject to singular or plural nouns, and semantic selection restrictions can disambiguate utterances in which for example the Agent-role has to be animate (e.g. in perception verbs such as *sehen* 'to see'). All other possible cues, such as word order, are ignored.

5 Results

5.1 Search

In all experiments, the constraints of the *search measure* were satisfied: every grammar only required one branch per utterance in production, and the number of branches in parsing never exceeded the number of possible interpretations. In terms of search length, more syncretism therefore does not automatically harm efficiency, provided that the grammar uses an adequate representation. Arguably, the smaller paradigms are even more efficient because they require less unifications to be performed.

5.2 Semantic Ambiguity

Now that it has been ascertained that more syncretism does not harm processing efficiency, we can compare cue-reliability of the different paradigms for semantic interpretation.

Ambiguous Utterances. Figure 2 shows the number of ambiguous utterances in parsing (in %) per paradigm and per set-up. As can be seen, the Old High German paradigm (black) is the most reliable cue in Experiment 1 (E1; when SV-agreement and selection restrictions are ignored) with 35.56% of ambiguous utterances, as opposed to 55.56% for Modern High German (grey) and 77.78% for Texas German (white).

When SV-agreement is taken into account (E2), the difference between Old and Modern High German becomes smaller, with both paradigms offering a reliability of more than 70%, while Texas German still faces more than 70% of ambiguous utterances.

Ambiguity is even more reduced when using semantic selection restrictions of the verb (set-up

E3). Here, the difference between Old and Modern High German becomes trivial with 4.44% and 6.94% of ambiguous utterances respectively. The difference with Texas German remains apparent, even though its ambiguity is cut by half.

In set-up E4 (case, SV-agreement and selection restrictions), the Old and Modern High German paradigms resolve almost all ambiguities, leaving little difference between them. Using the Texas German dialect, one utterance out of five remains ambiguous and requires additional grammatical cues or inferencing for semantic interpretation.

Number of possible interpretations. Semantic ambiguity can also be measured by counting the number of possible interpretations per utterance. A non-ambiguous language would thus have 1 possible interpretation per utterance. The average number of interpretations per utterance (per paradigm and per set-up) is shown in Table 8.

Paradigm	E1	E2	E3	E4
Old High German	1.56	1.22	1.04	1.03
Modern High German	1.56	1.28	1.07	1.04
Texas German	2.84	2.39	1.36	1.22

Table 8: Average number of interpretations per utterance type.

The Old High German paradigm has the least semantic ambiguity throughout, except in Experiment 1 (E1). Here, Modern High German has the same average effort despite having more ambiguous utterances. This means that the Old High German paradigm provides a better coverage in terms of construction types, but when ambiguity occurs, more possible interpretations exist.

6 Discussion

The experiments compare how well three different paradigms of definite articles perform if they are inserted in the grammar of Modern High German. The results show that, in isolation, Old High German offers the best cue-reliability for retrieving who's doing what to whom in events. However, when other grammatical cues are taken into account, it turns out that Modern High German achieves similar results with respect to syntactic search and semantic ambiguity, with a reduced paradigm (using only six instead of twelve forms).

As for the Texas German dialect, which has collapsed the accusative-dative distinction, the amount of ambiguity remains more than 20% using all available cues. One verifiable prediction of the experiments is therefore that this dialect should show an increase in alternative syntactic restrictions (such as word order) in order to make up for the lost case distinctions. Interestingly, such alternatives have been attested in Low German dialects that have evolved a similar two-way case system (Shrier, 1965). Modern High German, on the other hand, has already recruited word order for other purposes (such as information structure; Lenerz, 1977; Micelli, 2012), which may explain why the current paradigm has been able to survive since the Middle Ages.

Instead of an accidental by-product of phonological and morphological changes, then, a new picture emerges for explaining syncretism in Modern High German definite articles: German speakers have been able to reduce their case paradigm without loss in processing and interpretation efficiency. With cognitive effort as a selection criterion, subsequent generations of speakers found no linguistic pressures for maintaining particular distinctions such as gender in plural articles. Especially forms whose acoustic distinctions are harder to perceive are candidates for collapse if they are no longer functional for processing or interpretation. Other factors, such as frequency, may accelerate this evolution, as also argued by Barðdal (2009). For instance, there may be less benefits for upholding a case distinction for infrequent than for frequent forms.

If case syncretism is not randomly distributed over a grammatical paradigm, but rather functionally motivated, a new explanatory model is needed. One candidate is *evolutionary linguistics* (Steels, 2012b), a framework of cultural evolution in which populations of language users constantly shape and reshape their language in response to their communicative needs. The experiments reported here suggest that this dynamic shaping process is guided by the 'linguistic landscape' of a language. For instance, the presence of grammatical cues such as gender, number and SV-agreement may encourage paradigm reduction. However, reduction may be the start of a self-enforcing loop in which the decreasing cue-reliability of a paradigm may pressure language users into enforcing the alternatives to take on even more of the cognitive load of processing.

The intricate interactions between grammati-

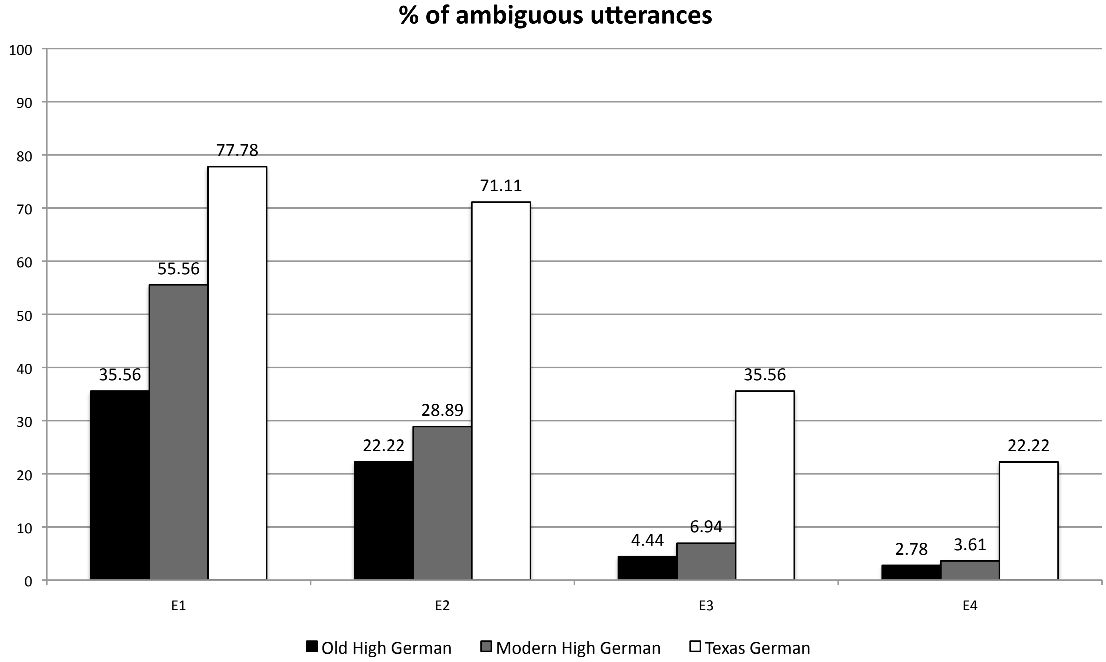


Figure 2: This chart shows the number of ambiguous utterances per paradigm per E(xperimental set-up) in %.

cal systems also requires more sophisticated measures. A promising extension of this paper could lie in an information-theoretic approach to language (Hale, 2003; Jaeger and Tily, 2011), which has recently explored a set of tools for assessing linguistic complexity, processing effort and uncertainty. Unfortunately, only little work has been done on morphological paradigms so far (see e.g. Ackerman et al., 2011), and the approach is typically applied in stochastic or Probabilistic Context Free Grammars, hence it remains unclear how the assumptions of this field fit into models of deep language processing.

7 Conclusions

More than 130 years after Mark Twain's complaints, it seems that the German language is not that awful after all. Through a series of computational experiments, this paper has proposed a different explanation for German case syncretism that answers some of the unsolved riddles of previous studies. First, the experiments have shown that an increase in syncretism does not necessarily lead to an increase in the cognitive effort required for syntactic search, provided that the representation of the grammar is processing-friendly. Secondly, by comparing cue-reliability of different paradigms for semantic disambiguation, the experiments have demonstrated that Modern High German achieves a similar performance as its Old High German predecessor using only half of the forms in its definite article paradigm.

Instead of a series of historical accidents, the German case system thus underwent a systematic and "performance-driven [...] morphological restructuring" (Hawkins, 2004, p. 79), in which linguistic pressures such as cognitive effort decided on the maintenance or loss of certain distinctions. The case study makes clear that formal and computational models of deep language understanding have to reconsider their strict division between competence and performance if the goal is to *explain* individual language development. This paper proposed that new tools and methodologies should be sought in evolutionary linguistics.

Acknowledgements

This research has been conducted at the Sony Computer Science Laboratory Paris. I would like to thank Luc Steels, director of Sony CSL Paris and the VUB AI-Lab of the University of Brussels, for his support and feedback. I also thank Hans Boas, Jóhanna Barðdal, Peter Hanappe, Manfred Hild and the anonymous reviewers for helping to improve this article. All errors remain of course my own.

References

Farrell Ackerman, James P. Blevins, and Robert Malouf. Parts and wholes: Implicative patterns in inflectional paradigms. In J.P. Blevins and J. Blevins, editors, *Analogy in Grammar: Form and Acquisition*, pages 54–81. Oxford University Press, Oxford, 2011.

Matthew Baerman. Case syncretism. In Andrej Malchukov and Andrew Spencer, editors, *The Oxford Handbook of Case*, chapter 14, pages 219–230. Oxford University Press, Oxford, 2009.

J. Barðdal. The development of case in germanic. In J. Barðdal and S. Chelliah, editors, *The Role of Semantics and Pragmatics in the Development of Case*, pages 123–159. John Benjamins, Amsterdam, 2009.

Manfred Bierwisch. Syntactic features in morphology: General problems of so-called pronominal inflection in German. In *To Honour Roman Jakobson*, pages 239–270. Mouton De Gruyter, Berlin, 1967.

James Blevins. Syncretism and paradigmatic opposition. *Linguistics and Philosophy*, 18:113–152, 1995.

Joris Bleys, Kevin Stadler, and Joachim De Beule. Search in linguistic processing. In Luc Steels, editor, *Design Patterns in Fluid Construction Grammar*. John Benjamins, Amsterdam, 2011.

Hans C. Boas. Case loss in Texas German: The influence of semantic and pragmatic factors. In J. Barðdal and S. Chelliah, editors, *The Role of Semantics and Pragmatics in the Development of Case*, pages 347–373. John Benjamins, Amsterdam, 2009a.

Hans C. Boas. *The Life and Death of Texas German*, volume 93 of *Publication of the The American Dialect Society*. Duke University Press, Durham, 2009b.

David Carter. Efficient disjunctive unification for bottom-up parsing. In *Proceedings of the 13th Conference on Computational Linguistics*, pages 70–75. ACL, 1990.

Ann Copestake. *Implementing Typed Feature Structure Grammars*. CSLI Publications, Stanford, 2002.

Berthold Crysmann. Syncretism in german: A unified approach to underspecification, indeterminacy, and likeness of case. In Stefan Müller, editor, *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar*, pages 91–107, Stanford, 2005. CSLI Publications.

Mary Dalrymple, Tracy Holloway King, and Louisa Sadler. Indeterminacy by underspecification. *Journal of Linguistics*, 45:31–68, 2009.

Michael Daniels. On a type-based analysis of feature neutrality and the coordination of unlikes. In *Proceedings of the 8th International Conference on HPSG*, pages 137–147, Stanford, 2001. CSLI.

Joachim De Beule. A formal deconstruction of Fluid Construction Grammar. In Luc Steels, editor, *Computational Issues in Fluid Construction Grammar*. Springer Verlag, Berlin, 2012.

Daniel P. Flickinger. On building a more efficient grammar by exploiting types. *Natural Language Engineering*, 6(1):15–28, 2000.

Jonathan Ginzburg and Ivan A. Sag. *Interrogative Investigations: the Form, the Meaning, and Use of English Interrogatives*. CSLI Publications, Stanford, 2000.

Adele E. Goldberg. *Constructions At Work: The Nature of Generalization in Language*. Oxford University Press, Oxford, 2006.

John T. Hale. The information conveyed by words in sentences. *Journal of Psycholinguistic Research*, 32(2):101–123, 2003.

John A. Hawkins. *Efficiency and Complexity in Grammars*. Oxford University Press, Oxford, 2004.

Bernd Heine and Tania Kuteva. *Language Contact and Grammatical Change*. Cambridge University Press, Cambridge, 2005.

Wolfgang Heinz and Johannes Matiasek. Argument structure and case assignment in german. In John Nerbonne, Klaus Netter, and Carl Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, volume 46 of *CSLI Lecture Notes*, pages 199–236. CSLI Publications, Stanford, 1994.

R.J.P. Ingria. The limits of unification. In *Proceedings of the 28th Annual Meeting of the ACL*, pages 194–204, 1990.

T. Florian Jaeger and Harry Tily. On language 'utility': Processing complexity and communicative efficiency. *WIREs: Cognitive Science*, 2(3):323–335, 2011.

L. Karttunen. Features and values. In *Proceedings of the 10th International Conference on Computational Linguistics*, Stanford, 1984.

Jürgen Lenerz. *Zur Abfolge nominaler Satzglieder im Deutschen*. Narr, Tübingen, 1977.

Martin Loetzsch. Tools for grammar engineering. In Luc Steels, editor, *Computational Issues in Fluid Construction Grammar*. Springer Verlag, Berlin, 2012.

Vanessa Micelli. Field topology and information structure: A case study for German constituent order. In Luc Steels, editor, *Computational Issues in Fluid Construction Grammar*. Springer Verlag, Berlin, 2012.

Gereon Müller. Remarks on nominal inflection in German. In Ingrid Kaufmann and Barbara Stiebels, editors, *More than Words: A Festschrift for Dieter Wunderlich*, pages 113–145. Akademie Verlag, Berlin, 2002.

Stefan Müller. An HPSG-analysis for free relative clauses in german. *Grammars*, 2(1):53–105, 1999.

Stefan Müller. Case in German – towards and HPSG analysis. In Tibor Kiss and Detmar Meurers, editors, *Constraint-Based Approaches to Germanic Syntax*. CSLI, Stanford, 2001.

Allan Ramsay. Disjunction without tears. *Computational Linguistics*, 16(3):171–174, 1990.

Ivan A. Sag. Coordination and underspecification. In Jongbok Kom and Stephen Wechsler, editors, *Proceedings of the Ninth International Conference on HPSG*, Stanford, 2003. CSLI.

Ivan A. Sag and Thomas Wasow. Performance-compatible competence grammar. In Robert D. Borsley and Kersti Börjars, editors, *Non-Transformational Syntax: Formal and Explicit Models of Grammar*. Wiley-Blackwell, Oxford, 2011.

Martha Shrier. Case systems in German dialects. *Language*, 41(3):420–438, 1965.

Luc Steels. Constructivist development of grounded construction grammars. In Walter Daelemans, editor, *Proceedings 42nd Annual Meeting of the Association for Computational Linguistics*, pages 9–19, Barcelona, 2004.

Luc Steels, editor. *Design Patterns in Fluid Construction Grammar*. John Benjamins, Amsterdam, 2011.

Luc Steels, editor. *Computational Issues in Fluid Construction Grammar*. Springer, Berlin, 2012a.

Luc Steels. Self-organization and selection in cultural language evolution. In Luc Steels, editor, *Experiments in Cultural Language Evolution*. John Benjamins, Amsterdam, 2012b.

Luc Steels and Joachim De Beule. Unify and merge in Fluid Construction Grammar. In P. Vogt, Y. Sugita, E. Tuci, and C. Nehaniv, editors, *Symbol Grounding and Beyond.*, LNAI 4211, pages 197–223, Berlin, 2006. Springer.

Remi van Trijp. Feature matrices and agreement: A case study for German case. In Luc Steels, editor, *Design Patterns in Fluid Construction Grammar*. John Benjamins, Amsterdam, 2011.

M. Walshe. *A Middle High German Reader: With Grammar, Notes and Glossary*. Oxford University Press, Oxford, 1974.

Bernd Wiese. Iconicity and syncretism. on pronominal inflection in Modern German. In Robin Sckmann, editor, *Theoretical Linguistics and Grammatical Description*, pages 323–344. John Benjamins, Amsterdam, 1996.

Joseph Wright. *An Old High German Primer*. Clarendon Press, Oxford, 2nd edition, 1906.

Dieter Wunderlich. Der unterspezifizierte Artikel. In Karl Heinz Ramers Dürscheid and Monika Schwarz, editors, *Sprache im Fokus*, pages 47–55. Niemeyer, Tübingen, 1997.