Neurophysiological Evidence for Morphological Underspecification in German Strong Adjective Inflection

Abstract

In the present study we investigate the relevance of the concept of underspecified inflection markers for the processing of language in the human brain. Underspecification is recognized as the main source of syncretism in many current morphological theories. However, relatively little is known about its cognitive status. In underspecification-based theories, a competition among morphological exponents arises systematically. In order to win such a competition, an inflection marker has to meet two requirements: Compatibility and Specificity. If underspecification is real, these two principles should also be an inherent part of the language processing system. One should therefore be able to observe separable effects for the violation of each of the criteria. We used the Event-Related Potential (ERP) violation paradigm to test this hypothesis in the domain of strong adjective inflection in German. We expected differences in brain potentials between two incorrect conditions whenever they represented different types of violation (of Compatibility and Specificity). Our findings strongly support underspecification: An ERP-component related to morpho-syntactic integration (viz., left anterior negativity; LAN) was modulated by violations of Specificity versus Compatibility. Furthermore, the neurophysiological evidence helps to distinguish between two kinds of morphological underspecification that have been proposed: It argues for maximal rather than minimal underspecification. Finally, the observed brain responses indicate increased processing demands for highly specific markers, which suggests that LAN effects may be sensitive not only to morpho-syntactic violations but also to the degree of processing effort.

1. Background

1.1 Syncretism and Economy

The starting point of the present study is a very common morphological phenomenon: syncretism. Syncretic morphological marking is characterized by the use of one phonological form in different morpho-syntactic contexts; thus, one morphological exponent appears in different cells of a paradigm. Consider, for instance, the examples of a small part of the system of Latin nominal declension given in table 1.

[Insert table 1 about here.]

The phonological forms of neuter nouns like *oppidum* are identical in nominative and accusative contexts although these two cases encode different grammatical functions. In Latin, the contrast between nominative and accusative is both syntactically relevant and morphologically marked for other genders/inflection classes (compare a masculine noun like *dominus* – *dominum* in table 1, which otherwise belongs to the same (o-) inflectional class).

The frequent occurrence of syncretism in the world's languages is remarkable in so far as it leads to mismatches between syntax and morphology (see Baerman et al. 2005). Moreover, identical forms for distinct morpho-syntactic configurations may generate severe ambiguity, as the German example in (1) illustrates.

(1) Die Frau sieht das Kind.
the woman$_{\mathrm{nom/acc}}$ see-3.SG the child$_{\mathrm{nom/acc}}$
'The woman sees the child.' *or* 'The child sees the woman.'

The morphological case marking of the two NPs in the example in (1) is ambiguous: Both phrases are morphologically marked as nominative or accusative. Thus, when processing a sentence like (1), there are morphologically ambiguous cues for the assignment of case and theta-roles to the NPs.[1]

In contrast, in a morphological system of argument encoding that does without syncretism, there are only one-to-one relations of form and function. A particular morphological exponent then represents exactly one specification of a syntactic context (and a particular specification of a syntactic context is unambiguously realized by exactly one morphological exponent, which, however, is also usually the case in morphological systems employing syncretism). It is clear that in fusional systems (where morphological exponents often encode more than one grammatical category – e.g., case, number, and gender), a strict one-to-one mapping of morphological exponents and syntactic contexts will increase the number of inflection markers; and the more grammatical categories there are that are simultaneously encoded by an exponent, the more dramatic this increase will be. In view of this, the tendency of many morphological systems towards a reduced marker inventory and the concurrent emergence of syncretism may be assumed to be motivated by reasons of economy: The fewer exponents an inflectional system consists of, the fewer lexical (sub)items have to be stored.

However, it would seem that this economy gain for the storing component of language can only be achieved at the cost of extra effort in the procedural component: When dealing with syncretism, the system has to cope with processing difficulties such as ambiguities in parsing (either of the local type – when they are eventually resolved by unambiguous morphological encoding – or of the global type – when they persist to the end, as in (1)). Furthermore, as we will see momentarily, providing syncretic morphological exponents with a unique function presupposes underspecification; and underspecification in turn typically necessitates two further assumptions concerning the modelling of syncretism in morphological theory that would also seem to be demanding for processing capacities, viz., feature decomposition and competition resolution. We turn to this issue in the next section.

1.2 Underspecification in Morphological Theory

It is not a priori clear to what extent syncretism can be viewed as systematic, and to what extent it might be accidental. However, it is uncontroversial that at least some instances of syncretism are not accidental.[2] It is a common assumption of many current theoretical approaches to morphology that as many instances of syncretism should be accounted for systematically as possible. Most often, this research strategy is implicitly presupposed; sometimes, it is explicitly formulated (see, e.g., Embick 2003, Müller 2007, Pertsova 2007, and references cited in these works). Consequently, given that there is systematic syncretism in abundance, the question arises of how to account for the phenomenon.

One of the most widely pursued approaches to syncretism in morphological theory is underspecification. The basic idea goes back to Jakobson (1962a;b), and the method has been further refined by Bierwisch (1967). Here we will illustrate the mechanics of an underspecification approach to syncretism on the basis of the paradigm of pronominal inflection in German. There are two reasons for this choice. First, many underspecification-based analyses of this paradigm have been proposed in the literature; in addition to Bierwisch's (1967) original approach, syncretism in German pronominal inflection has been addressed in terms of underspecification by Blevins (1995), Sauerland (1996), Wunderlich (1997a), Wiese (1999), Gallmann (2004), and Trommer (2005). And second, whereas the system of pronominal inflection in German provides morphological exponents for determiners like *dies-* ('this'), *jen-* ('that') and, with some stem variation, the definite article *der, die, das* ('the$_{masc}$', 'the$_{fem}$', 'the$_{neut}$'), it is almost identical to the system of strong adjective inflection that forms the empirical basis of our neurophysiological study (and that provides morphological exponents for adjectives in NPs where there is no initial determiner). As a matter of fact, the two paradigms are identical except for genitive masculine/neuter singular contexts, where pronominal inflection has an exponent *-es* and strong adjective inflection has an exponent *-en*. Since this variation is a negligible factor (see note 10 below), and since our own study focuses on accusative markers where there is no difference, we can thus look at underspecification-based analyses of pronominal inflection as analyses of strong adjective inflection by extension; and this is what we will do throughout the remainder of the paper.[3]

Against this background, consider the paradigm of pronominal inflection in German in table 2, with four instantiations of the grammatical category case, three instantiations of the grammatical category gender, and two instanatiations of the grammatical catgory number.

[Insert table 2 about here.]

There are no gender differences in morphological exponence in the plural. Depending on whether or not one assumes that this implies that the pertinent paradigm cells are conflated not only morphologigally but also syntactically, there are thus either sixteen or twenty-four different paradigm cells in this paradigm. However, there are only five distinct morphological forms: *-e*, *-er*, *-en*, *-es*, *-em*. For instance, there is only one marker for the two morpho-syntactic feature specifications [dative masculine singular] and [dative neuter singular], viz., *-em*; and this marker is different from all the other markers in the paradigm.

The most important consequence of the underspecification approach is that natural classes of instantiations of grammatical categories can be expressed. Thus, a common basis of the instances of a given syncretism is sought – a property that the different contexts exhibiting an identical morphological exponent have in common. This property then characterizes a natural class of morpho-syntactic specifications. In the case at hand, [dative masculine singular] and [dative neuter singular] contexts (or paradigm cells) differ only with respect to gender information. This leads to the assumption that [masculine] and [neuter] form a natural class. In an approach that relies on morpho-syntactic features, this could be expressed by attributing the two contexts some extra feature that the information associated with an inflection marker can then refer to. Alternatively, and this is the strategy pursued by Jakobson, Bierwisch, and many others, one can derive natural classes of instantiations of a grammatical category (e.g., masculine and neuter as two instantiations of the grammatical category gender) from a *decomposition* of the standard morpho-syntactic features into combinations of more abstract primitive features. Thus, the instantiations of gender in German can be assumed to have the fine structure in (2), with the masculine, feminine and neuter arising from a cross-classification of the primitive gender features [$\pm$masc] and [$\pm$fem].[4]

(2) *Decomposition of gender features in German*:

a. masculine = [+masc,–fem]
b. feminine = [–masc,+fem]
c. neuter = [–masc,–fem]
d. [ ] = [+masc,+fem]

The idea then is that morphological exponents do not have to be characterized by fully specified morpho-syntactic features; they can also be characterized by *underspecified* morpho-syntactic information. For instance, the dative singular marker *-em* is not characterized as [+masc,–fem] (= masculine) or as

[–masc,–fem] (= neuter). Rather, this exponent is characterized by a feature specification that is underspecified with respect to gender: [–fem].

Exactly the same situation arises with case features. As shown in table 2, the marker *-es* is employed for both nominative neuter and accusative neuter contexts. This syncretism is in line with a basic Indo-European principle (see note 2), and thus certainly not accidental. The syncretism with *-e* in nominative feminine and accusative feminine contexts in German looks systematic in the same way; and so on. This can be taken to indicate that cases should be decomposed into combinations of primitive features in the same way as genders. Bierwisch's (1967) proposal (which builds on Jakobson's general assumptions for Slavic but modifies his specific suggestions substantially) is based on the cross-classification of the two primitive case features [±obj(ect)] and [±obl(ique)]; it is given in (3).[5]

(3) *Decomposition of case features in German*:

a. nominative = [–obj,–obl]
b. accusative = [+obj,–obl]
c. dative = [+obj,+obl]
d. genitive = [–obj,+obl]

From this it follows that nominative and accusative form a natural class in German (captured by the feature [–obl], which they exclusively share); as do genitive and dative ([+obl]), accusative and dative ([+obj]), and nominative and genitive ([–obj]). In contrast, nominative and dative do not form a natural class, and neither do accusative and genitive. Thus, one does not expect systematic instances of syncretism to affect exclusively these latter two case pairs.

By assuming underspecification with respect to case, gender, and number information, natural classes of these instantiations of grammatical categories can be defined, and instances of syncretism that document these natural classes can be derived. Of all the above-mentioned analyses that pursue this general kind of approach to pronominal inflection marking in German, we will focus on a version of Blevins's (1995) account in what follows, both for the illustration of the mechanics of undespecification and for hypothesis formation for our empirical investigation.[6]

The essentials of Blevins's (1995) analysis of the system of pronominal inflection marking in German are given in (4). On this view, there are eight morphological exponents, all of which are underspecified with respect to morpho-syntactic features (case, gender, number). In fact, one marker (/e/, realized as ə) is fully underspecified with respect to all such information; it emerges as an *elsewhere* marker.[7]

(4) *Inventory of exponents for pronominal inflection in Blevins (1995)*

a.	/n/ ↔ [+pl,+obj,+obl]	(dat.pl.)
b.	/m/ ↔ [–fem,+obj,+obl]	(dat.masc.sg./neut.sg.)
c.	/s/ ↔ [–fem,+obl]	(gen.masc.sg./neut.sg.)
d.	/r/ ↔ [+obl]	(dat./gen.fem.sg., gen.pl.)
e.	/n/ ↔ [+masc,–fem,+obj,–obl]	(acc.masc.sg.)
f.	/r/ ↔ [+masc,–fem,–obl]	(nom.masc.sg.)
g.	/s/ ↔ [–fem,–obl]	(nom./acc.neut.sg.)
h.	/e/ ↔ [ ]	(nom./acc.fem.sg./pl.)

At this point, what is needed is a method to correctly match the underspecified exponents in (4) to the fully specified syntactic contexts (or paradigm cells) that they show up in; i.e., it remains to be shown how the paradigm in table 2 can be derived from the marker inventory in (4). To achieve this, underspecification-based approaches employ two general requirements that we will refer to as *Compatibility* and *Specificity*. These requirements can then be combined into the general principle governing morphological realization (and thus the correct filling of paradigm cells) given in (5):[8]

(5) *Morphological Realization*:
A morphological exponent M is chosen for a syntactic context (or paradigm cell) S if (a) and (b) hold.

a. M is compatible with S.
b. M is the most specific exponent among those that satisfy (5-a).

Compatibility can be understood as in (6).

(6) *Compatibility*:
A morphological exponent M is compatible with a syntactic context (or paradigm cell) S if M realizes a *subset* of the morpho-syntactic feature/value pairs of S.

From Compatibility, it follows that, e.g., the exponent /m/ (which is specified as [–fem,+obj,+obl]) cannot be used in accusative contexts (which are specified as [+obj,–obl], and therefore cannot form a superset of the former specification). It also follows from Compatibility that /m/ cannot be used in dative feminine contexts (since the latter are specified as [+fem,+obl,+obj], which is incompatible with the marker's [–fem,+obj,+obl] specification). In contrast, /m/ is compatible with dative masculine singular contexts and dative neuter singular contexts, where it does indeed occur.

However, Compatibility alone does not suffice to ensure correct marker choice in underspecification-based theories. In many cases (in fact, in almost all syntactic contexts identified in table 2, except for non-oblique feminine or plural specifications), more than one exponent is compatible (in the sense of (6)) with the syntactic target specification. Consider, e.g., accusative masculine singular contexts. It turns out that four exponents in (4) are compatible with the syntactic specification [+masc,–fem,+obj,–obl,–pl]. This is illustrated in (7).

(7) <u>*Different compatible markers for accusative masculine singular contexts*</u>

			syntactic context				
			[+masc,	–fem,	−pl,	+object,	−oblique]
marker	/n/	↔	[+masc,	–fem,		+object,	−oblique]
	/r/	↔	[+masc,	–fem,			−oblique]
	/s/	↔	[	–fem,			−oblique]
	/e/	↔	[				]

More generally, underspecification typically has the effect of producing a *competition* of different markers for one and the same morpho-syntactic context. The competition can be resolved by invoking Specificity in addition to Compatibility, as it is done in (5-b). A simple definition of Specificity is given in (8).[9]

(8) *Specificity*:
A morphological exponent M_1 is more specific than a morphological exponent M_2 if M_1 realizes *more* features than M_2.

Given the principle of Morphological Realization in (5), it follows from (8) that /n/ wins the competition in (7); it emerges as the most specific marker that is available. Similarly, it can easily be verified that the inventory of exponents in (4) fully derives the paradigm of pronominal inflection marking in German, and thus accounts for a number of instances of syncretism in a principled way.[10]

At this point, we may note that given the Compatibility and Specificity requirements for morphological realization, there may in fact be two distinct reasons underlying the failure of morphological exponents to qualify as optimal for some particular syntactic context. In accusative neuter singular contexts, for instance, /s/ (with the specification [–fem,–obl]) is chosen, given (5). The exponent /n/ (with the specification [+masc,–fem,+obj,–obl]), which is chosen in accusative masculine singular contexts, is excluded here because of Compatibility (via the feature [+masc]); in contrast, the exponent /e/ (with the empty specification [ ]), which is chosen in accusative feminine singular contexts, is excluded here not because of Compatibility (there are no features that could conflict) but because of Specificity. This will become relevant later.

Underspecification-based approaches to syncretism that rely on feature decomposition and competition resolution by Specificity can be viewed as well established and successful. These concepts play an important role in many current linguistic theories, among them Distributed Morphology (Noyer 1992; Halle & Marantz 1993; Harley & Noyer 2003), Minimalist Morphology (Wunderlich 1996; 1997b; 2004), A-morphous Morphology (Anderson 1992; Aronoff 1994), Paradigm Function Morphology (Stump 2001), Network Morphology (Corbett & Fraser 1993), and various accounts in Optimality Theory (Grimshaw 2001, Trommer 2001; 2006).[11]

Still, it seems clear that if morphological realization centers around (a principle like) (5) and relies on concepts like underspecification, feature decomposition, and competition resolution, this will create an extra layer of morphological complexity and implies additional processing routines (note that underspecification and feature decomposition do not seem to be independently motivated in syntax, thereby essentially creating purely "morphomic" structure, i.e., what Aronoff (1994) calls "morphology by itself"). Therefore, as mentioned earlier, it would seem that reducing the effort in the storing component (by having morphological exponents with multiple functions, i.e., syncretism) increases computational effort in the procedural parts of the grammar not only in syntax (by giving rise to ambiguity) but also in morphology (by requiring additional means to manage syncretism in the grammar), and vice versa. At present, there is still a very lively, ongoing debate whether reducing the effort for the lexicon or the procedural component of the grammar should be a preferred strategy in the modelling of grammatical knowledge (cf., e.g., construction grammar as laid out in Tomasello 2003 and Goldberg 2006 vs. classical rule-based approaches), as well as in the modelling of language processing in the human mind in psycholinguistics (see the next section).

1.3 Underspecification in Psycholinguistics

Psycholinguistic research on inflectional morphology has predominantly focused on whether inflected word forms are processed and stored as whole words or decomposed into more basic components (e.g., stem and inflectional affix) and on how mental representations of inflected word forms are related to each other. There are many accounts where the full-listing hypothesis is adopted (Butterworth 1983), and there are also many accounts arguing for morphological substructure (Taft & Forster 1975). Finally, there are dual-route models in which both mechanisms may operate more or less simultaneously (Caramazza et al. 1988; Pinker 1991; Clahsen et al. 2003). For overviews see Schriefers 1999 and Marslen-Wilson 2007.

However, at present relatively little is known about the psychological reality of more fine-grained analyses of inflectional phenomena as they have been advanced in morphological theory. As should be clear from the previous section, these kinds of analyses are very common in theoretical linguistics. They establish a radically different view on the nature of morpho-syntactic features and categories. As seen above, morphology may not necessarily operate with holistic instantiations of grammatical categories as they are known from syntax like, e.g., masculine vs. feminine vs. neuter gender, or nominative vs. accusative vs. dative vs. genitive case. Instead, an underlying fine-structure of these labels has been argued for that reflects systematic and structural relations between them.

Strongly related to this notion of feature decomposition is the question of the relation of single inflection markers in the lexicon and how they may differ from each other with respect to the number and kind of features they represent. More precisely, the question arises whether the fine-grained structures assumed by analyses in theoretical linguistics are also relevant for human language comprehension, and whether inflection markers are processed in the way it is predicted by underspecification models.

Until now only a few psycholinguistic studies have addressed this issue; see, e.g., Janssen & Penke 2002; Clahsen et al. 2001; Penke et al. 2004. In Clahsen et al. 2001 the authors report a study on German adjectival inflection. They show that inflected words with a highly specific affix (*-em*) lead to longer reaction times in a lexical decision task than words with less specific affixes (*-es*) irrespective of the frequency of the whole word form. As these results may also be interpreted as an effect of the frequency of the relevant affixes, the authors additionally report a cross-modal priming experiment. They argue that this specific experimental task is less prone to possible frequency effects of the phonological word forms. Again the results of this experiment are compatible with the notion of differences in specificity of different inflection markers. But, crucially, the results of both experiments may be confounded by phonological factors. This holds, for instance, for the observed asymmetry in priming of *-e* $\rightarrow$ *-(e)s* (= reduced priming) versus *-(e)s* $\rightarrow$ *-e* (= full priming). In theses cases the degree of phonological overlap could well be a critical factor, especially in priming experiments, which are known to be susceptible to such effects.

Penke et al. (2004) also focus on feature specifications of inflected adjectives and determiners in German. They use a sentence-matching task to compare reading times of correct and incorrect sentences. This experimental paradigm is prone to the so-called 'grammaticality effect', i.e., when two sentences have to be compared and it needs to be determined whether they are identical or not, participants' reaction times are sensitive to the grammaticality of these sentences. Response times typically slow down for the judgment of incorrect compared to correct sentences. The authors report significant differences for this grammaticality effect for different types of incorrect markers on adjectives and determiners, concluding that the feature specifications of these markers conflict differently with the features of the syntactic context. For markers with underspecified but non-conflicting features they obtained no grammaticality effect, whereas whenever a (positively valued) feature was in conflict with a (negatively valued) feature of the context, reaction times slowed down. This study thus reports first evidence that the two above-mentioned distinct principles of Specificity and Compatibility lead to separable effects in behavioral experimental tasks.

Moreover, Penke and colleagues presuppose a fundamental, qualitative distinction between positive and negative feature specifications. The authors claim that the former 'are a part of the representations of morphologically complex word forms or affixes, whereas negative features are assigned on the basis of paradigmatic contrasts' (Penke et al. 2004: p. 432). Therefore they suggest that only positive feature values may conflict with features of the syntactic context and hence slow down participants' reaction times.

Another study that addresses the issue of underspecified inflection markers is Janssen & Penke (2002). The authors investigate verb-subject agreement errors of elicited speech-production data of German agrammatic aphasics. The basic assumption is that verb-subject agreement is not impaired in principle in agrammatic aphasics but that they produce slightly more often erroneous forms than an unimpaired control group would do. The authors compare the results of two different elicitation tasks with predictions drawn from two different morphological models of the representation of inflectional affixes: (i) the Network Morphology model that relies on default inheritance trees (Corbett & Fraser 1993) and (ii) the Minimalist Morphology model, an account based on underspecified paradigms by Wunderlich (1996). Janssen & Penke hypothesize that the different conceptual representations of affixes in the two accounts should be manifested in different types of substitution errors in language production. The results reported by Janssen & Penke support Wunderlich's account.

So far, all empirical evidence that underspecification might be involved in the processing of language comes from studies (of both healthy and aphasic participants) that use *behavioral* methods to measure responses to linguistic stimuli. These studies cannot display the exact timing of how and when morphological information is processed. Moreover, some of the studies presented inflected words in isolation (see Clahsen et al. 2001). Here it is not clear to what extent the obtained effects are induced by these tasks, and how inflected word forms are processed without any (at least minimal) morpho-syntactic context information.

In order to investigate the processing of morphological information online, we conducted an ERP (event-related potential) experiment in which correctly and incorrectly inflected word forms were embedded in plausible morpho-syntactic contexts. An important advantage of the ERP method is that it allows for a measurement of highly automatic and even sub-conscious processes relevant to language, with a high temporal resolution (cf. Friederici 2002, Kaan 2007). The ERP violation paradigm that was used for the present study compares EEG responses to correct versus incorrect linguistic stimuli. This comparison reveals typical brain potential patterns that are usually defined as ERP components according to their electrophysiological and functional characteristics. These characteristics comprise the latency, amplitude, scalp distribution and sensitivity of an evoked potential. Language-related ERP components are, e.g., a left anterior negativity (LAN) and a late positivity (P600), which have been shown to be sensitive to morpho-syntactic information processing. The LAN usually emerges around 300-500 ms after stimulus presentation, with unilateral or bilateral distribution on anterior sites. A LAN is often evoked when a participant encounters a morpho-syntactic violation; it is commonly associated with error detection on a phrase-structural or morpho-syntactic level (see, e.g., King & Kutas 1995, Kluender & Kutas 1993 and Münte et al. 1998). This component is sensitive to grammatical violations that hinder integration of the stimulus into the morpho-syntactic context, such as word category or agreement errors. On the other hand, the P600 consists of a stronger deviation of wave forms towards positivity. It emerges between 500-900ms and displays a centroparietal scalp distribution. The P600 component has been predominantly found for syntactic anomalies (see, e.g., Friederici et al. 1993, Friederici & Meyer 2004, and Rossi et al. 2005), but also for syntactically complex or ambiguous sentence structures (see Friederici et al. 2002, Friederici & Weissenborn 2007), as well as for semantic-pragmatic anomalies (see Regel et al. 2011). The P600 effect in response to syntactic manipulations has been interpreted as a reflection of controlled processes enganged in reanalysis or repair of sentences.

2. Experiment

2.1 Setting the Stage

In the present study we examined whether underspecification of inflection markers, assumed as a core principle in morphological theory, is relevant for processing inflection in language comprehension. In order to approach this issue, we used the ERP violation paradigm to investigate brain responses to morphologically correctly and incorrectly marked phrases in German.

The underlying idea is that if the assumption of underspecified morphological markers is correct, we expect a modulation of the LAN component in response to the two different types of illicit morphological exponents. More specifically, these processing differences should depend on the *kind of violation* an incorrect morphological marker represents. Basically, the question was whether incorrect markers with *incompatible* features (i.e., a violation of Compatibility) trigger different responses than markers that are *compatible* but ruled out due to their insufficient specificity (i.e., a violation of Specificity). These differences should be mirrored by diverging amplitudes of the error-detection component (LAN) and/or the component reflecting reanalysis and repair (P600).

In the experiment reported here we investigated the processing of case-marked and gender-marked adjectives in complex prepositional phrases in German. All phrases consisted of a preposition followed by an adjective and a noun. As seen above, (pre-nominal) adjectives in German agree with a noun with respect to case, gender, and number. These categories are expressed morphologically as (fusional) suffixes on the adjective. We focused on the morphological exponents of the strong adjective inflection (which, as noted above, is nearly identical to the pronominal inflection paradigm). More specifically, we compared the three inflection markers in the accusative singular for each of the three genders (i.e., feminine, masculine, and neuter): see (9).

(9) a. durch schlicht-e Struktur
by plain-FEM.SG.ACC structure.FEM

b. durch schlicht-en Geschmack
by plain-MASC.SG.ACC taste.MASC

c. durch schlicht-es Design
by plain-NEUT.SG.ACC design.NEUT

The choice of exactly these three morphological exponents (/e/, /n/, /s/) is due to two considerations. First, the three markers in the accusative singular domain of the paradigm are very likely to differ with respect to the kind and number of their morpho-syntactic features. The feature specifications of these three exponents according to Blevins's (1995) analysis are repeated here from (4); see (10).

(10) *Underspecified accusative markers* (Blevins 1995)

e. /n/ $\leftrightarrow$ [+masc,–fem,–obl,+obj]

g. /s/ $\leftrightarrow$ [–fem,–obl]

h. /e/ $\leftrightarrow$ [ ]

In this analysis, /n/ is the most specific of the three markers while /e/ is the least specific (elsewhere) marker of the whole system (it is radically underspecified).[12] Showing up in syntactic accusative contexts, the morphological exponent /e/ should therefore always meet the Compatibility criterion while /n/ and /s/ may give rise to a feature conflict.

The second reason for choosing these three markers in particular is that, when comparing different forms in different conditions in an ERP experiment, as many other factors as possible should be identical; i.e., the syntactic contexts should differ only in the feature specifications immediately relevant to the experiment. This requirement could be met by manipulating the gender of the noun that the adjective agrees with. All other factors of the context such as number (singular) and case (accusative) could be kept identical. Thus, in the accusative domain of the paradigm we found three good candidates to test our hypothesis: First, many existing theoretical approaches to pronominal inflection and strong adjective inflection in German converge on the assumption that the feature specifications of these exponents may differ with respect to both Specificity and Compatibility if chosen in incorrect contexts. And second, other factors could be kept nearly identical in the different conditions.

As mentioned earlier, the basic idea of the experimental design was that, whenever gender agreement was manipulated, the incorrect versions should represent different types of violation depending on the nature of their illegitimacy (viz., incompatibility of features versus a lack of specificity).

This basic idea is illustrated in (11) for phrases containing neuter nouns (more precisely, accusative neuter singular contexts). Depending on whether the preceding adjective is incorrectly marked with /e/ (i.e., a feminine form) or /n/ (i.e., a masculine form), two different types of violation occur.[13]

(11) *Two types of illicit neuter agreement* (based on marker specifications in Blevins 1995)

a. *correct*

durch	schlichtes	Design
by	plain.NEUT	design.NEUT
	$\left[-fem, -obl \right]$	$\begin{bmatrix} -masc, -fem \\ -obl, +obj \end{bmatrix}$

b. *incorrect 1*

compatible (excluded by Specificity)

*durch	schlichte	Design
by	plain.FEM	design.NEUT
	$\left[\quad \right]$	$\begin{bmatrix} -masc, -fem \\ -obl, +obj \end{bmatrix}$

c. *incorrect 2*

incompatible (excluded by Compatibility)

*durch	schlichten	Design
by	plain.MASC	design.NEUT
	$\begin{bmatrix} \textbf{+masc}, -fem \\ -obl, +obj \end{bmatrix}$	$\begin{bmatrix} -\textbf{masc}, -fem \\ -obl, +obj \end{bmatrix}$

Assuming that all potentially intervening factors can be excluded, this difference of a violation of Compatibility versus Specificity should be mirrored in a systematically different pattern in relevant ERP components when comparing the two incorrect conditions with the correct condition. More precisely, regarding error detection we hypothesized that whenever a marker has conflicting features this could be a stronger cue for error detection than an incorrect but compatible marker. Under this view, an incompatible marker can be regarded as a major violation compared to a compatible (but still incorrect) marker, which in turn should represent a minor violation. On the other hand, if morphological underspecification does not play a role in language processing and if the difference between correct and incorrect conditions is a categorical one instead, both incorrect conditions should induce similar brain activity for detecting the mismatch.

Concerning the experimental design it is important to emphasize that we did not compare the forms of differently inflected adjectives in isolation. When incrementally presented, all phrases are correct until the noun is encountered, and the system cannot detect any error up to this point. Only when the (gender) features of the noun are available can the wrong morphological exponent of the adjective become detectable. We therefore compared ERPs during the processing of the *noun*, not the adjective, since it is at this position that the morpho-syntactic agreement between the adjective and the noun becomes manifest, and agreement actually has to be evaluated. An additional advantage of this experimental set-up is that one does not have to compare different phonological forms for critical epochs: In all three conditions within one gender (i.e., one correct condition and two incorrect conditions) identical forms are compared (e.g., the word *Design* in the example above).

2.2 Maximal vs. Minimal Underspecification

Next, in addition to distinguishing between a traditional, categorical account and underspecification-based accounts, the present investigation was supposed to shed new light on a fundamental question that has only recently begun to come to the fore in theoretical linguistics: Assuming that morphological underspecification exists, it remains an open question *to what extent* exponents should be taken to be underspecified. If a homogeneous approach is to be pursued here, it seems that the two options are that inflection markers should be either *minimally* or *maximally* underspecified. Many analyses, like that of Blevins (1995), try to reduce feature specifications as much as possible, yielding maximally underspecified markers. On the other hand, it is also possible to retain as many features as possible by omitting only those that give rise to inconsistencies among the occurrences of some marker. This latter kind of approach relies on *minimal* (or ‘contrastive’) underspecification. The learning algorithm for acquiring underspecified morphological exponents that is sketched in Harley 2001 and developed in some detail in Pertsova 2007 (where it is called the No-Homonymy learner) implies such a minimal underspecification: Underspecified feature structures of exponents are learned by *intersecting* the sets of the different (fully specified) environments; as soon as a minimally underspecified exponent can be postulated, the algorithm stops, with no option to proceed to a more substantially (or even maximally) underspecified marker. Consequently, in an analysis that envisages minimal underspecification, markers may still contain redundant features.

The two strategies yield different conceptual benefits. Analyses relying on *maximal underspecification* may be regarded as more elegant from a purely formal point of view. As a side effect, if feature sets of markers are kept as small as possible, the complexity of the lexical component of such systems is further reduced: Fewer features have to be stored for each vocabulary item. In contrast, analyses based on *minimal underspecification* keep the feature sets of each vocabulary item as rich in content as possible; features are only reduced as much as necessary to cover patterns of syncretism. This strategy may leave lexical entries of morphological exponents rather complex (viz., as specific as possible); however, it reduces processing efforts since the number of competing markers for each cell in the paradigm is minimized. Thus, even in underspecification-based approaches, the lexical–procedural dichotomy still emerges as relevant. An empirical argument in favor of minimal or maximal underspecifiaction can thus be regarded as an interesting result. It emphasizes either the role of the lexicon by reducing demands on the procedural part or vice versa.

As a case in point, consider Blevins's (1995) analysis again. Closer inspection reveals that the specifications of markers in (4) still contain redundant features, despite an overall orientation towards maximal underspecification. For concreteness, it is the [–obl] feature in the specifications associated with the markers /r/ (for nominative masculine singular contexts, in (4-f)) and /s/ (for nominative and accusative neuter singular contexts, in (4-g)) that is not strictly necessary in order to fully derive the paradigm. A strengthened, maximally underspecified version of Blevins' system that does without these two occurrences of [–obl] is given in (12).

(12) *Inventory of exponents in Blevins 1995, modified to maximal underspecification*

a.	/n/ ↔ [+pl,+obj,+obl]	(dat.pl.)
b.	/m/ ↔ [–fem,+obj,+obl]	(dat.masc.sg./neut.sg.)
c.	/s/ ↔ [–fem,+obl]	(gen.masc.sg./neut.sg.)
d.	/r/ ↔ [+obl]	(dat./gen.fem.sg., gen.pl.)
e.	/n/ ↔ [+masc,–fem,+obj,–obl]	(acc.masc.sg.)
f.	/r/ ↔ [+masc,–fem]	(nom.masc.sg.)
g.	/s/ ↔ [–fem]	(nom./acc.neut.sg.)
h.	/e/ ↔ [ ]	(nom./acc.fem.sg./pl.)

In (12), the set of features associated with each marker – and thus, the information that has to be stored for each item – is maximally reduced.

On the other hand, an analysis that maintains underspecification but tries to keep the feature configurations of markers as specific as possible admits redundant features for markers. The list of exponents in (13) illustrates an analysis in terms of minimal underspecification on the basis of a modified version of Blevins' original account. Here, all redundant features that do not create a conflict with possible contexts of insertion are retained in the markers' specification.[14]

(13) *Inventory of exponents in Blevins 1995, modified to minimal underspecification*

a.	/n/ ↔ [+pl,+obj,+obl]	(dat.pl.)
b.	/m/ ↔ [–fem,+obj,+obl,*–pl*]	(dat.masc.sg./neut.sg.)
c.	/s/ ↔ [–fem,+obl,*–obj*,*–pl*]	(gen.masc.sg./neut.sg.)
d.	/r/ ↔ [+obl]	(dat./gen.fem.sg., gen.pl.)
e.	/n/ ↔ [+masc,–fem,+obj,–obl,*–pl*]	(acc.masc.sg.)
f.	/r/ ↔ [+masc,–fem,–obl,*–obj*,*–pl*]	(nom.masc.sg.)
g.	/s/ ↔ [*–masc*,–fem,–obl,*–pl*]	(nom./acc.neut.sg.)
h.	/e/ ↔ [*–obl*][15]	(nom./acc.fem.sg./pl.)

The two (or three, given the difference between our original version of Blevins' system and the one employing maximal underspecification) analyses are extensionally equivalent; it is not clear whether any independent evidence can be adduced for or against either of them purely on the basis of a theoretical approach. The situation is different if a neurophysiological perspective is taken.

2.3 Hypotheses and Predictions

Our hypothesis is that if the underspecification approach is on the right track, there should be two different kinds of morphosyntactic violations when gender agreement on adjectives is manipulated. The first kind of violation is induced by *incompatible features*; thus, it gives rise to a Compatibility problem. In contrast, incorrectness due to *insufficient specificity* represents a second kind of violation, viz., a violation of Specificity. While, on the one hand, a violation of Compatibility can be regarded as a *major* violation for the system (i.e., a relatively strong cue for error detection, and larger LAN), violating Specificity, on the other hand, can be regarded as a *minor* violation (i.e., a less salient cue for error detection). It is important to note that for each gender the predictions regarding the type of violation of the two incorrect versions differ. Moreover, the two versions of underspecification-based accounts (employing maximal vs. minimal underspecification) lead to different predictions. Based on the maximal underspecification account, we expect the pattern in (14).

(14) *Two types of illicit agreement, with maximal underspecification (as in (12))*:

a. for feminine phrases:
identical kind of violation (Compatibility)

context features:		[–masc, +fem, –obl, +obj]
correct marker:	-e	[ ]
incompatible (incorr.1):	-(e)n	[**+masc, –fem**, –obl, +obj]
incompatible (incorr.2):	-(e)s	[**–fem**]

b. for masculine phrases:
identical kind of violation (Specificity)

context features:		[+masc, –fem, –obl, +obj]
correct:	-(e)n	[+masc, –fem, –obl, +obj]
compatible (incorr.1):	-(e)s	[–fem]
compatible (incorr.2):	-e	[ ]

c. for neuter phrases:
different kind of violation (Compatibility vs. Specificity)

context features:		[–masc, –fem, –obl, +obj]
correct:	-(e)s	[–fem]
incompatible (incorr.1):	-(e)n	[**+masc**, –fem, –obl, +obj]
compatible (incorr.2):	-e	[ ]

Thus, in an approach that relies on maximal underspecification (see (14)) predictions for incorrect conditions only differ for neuter nouns. In this context, an /n/-marked adjective triggers a feature conflict (i.e., it violates Compatibility) whereas an /e/-marked adjective does not bear conflicting features but violates Specificity. In contrast, in an approach that relies on minimal underspecifiaction (see (15)), there are also different predictions for masculine contexts.

(15) *Two types of illicit agreement, with minimal underspecification (as in (13))*:

a. for feminine phrases:
identical kind of violation (Compatibility)

context features:		[–masc, +fem, –obl, +obj, –pl]
correct marker:	-e	[–obl]
incompatible (incorr.1):	-(e)n	[**+masc, –fem**, –obl, +obj, –pl]
incompatible (incorr.2):	-(e)s	[–masc, **–fem**, –obl, –pl]

b. for masculine phrases:
different kind of violation (Compatibility vs. Specificity)

context features:		[+masc, –fem, –obl, +obj, –pl]
correct:	-(e)n	[+masc, –fem, –obl, +obj, –pl]
incompatible (incorr.1):	-(e)s	[**–masc**, –fem, –obl, –pl]
compatible (incorr.2):	-e	[–obl]

c. for neuter phrases:
different kind of violation (Compatibility vs. Specificity)

context features:		[–masc, –fem, –obl, +obj, –pl]
correct:	-(e)s	[–masc, –fem, –obl, –pl]
incompatible (incorr.1):	-(e)n	[**+masc**, –fem, –obl, +obj, –pl]
compatible (incorr.2):	-e	[–obl]

The difference between (14-b) and (15-b) is that in the approach based on minimal underspecification (i.e., (15-b)) there is a redundant [−masc] feature of the marker /s/. This feature is predicted to yield a conflict with the syntactic context, and therefore to induce a qualitatively different error compared to the illicit but compatible /e/ marking. Thus, the two models differ in their predictions regarding masculine phrases. While with maximal underspecification both incorrect exponents represent the same kind of violation (satisfying Compatibility and fatally violating Specificity), with minimal underspecification they differ: The /e/ marker is still compatible but the /s/ marker now violates Compatibility.

To sum up, the predictions for the processing of different incorrect markers for a categorical versus two different versions of underspecification-based accounts are given in table 3.[16]

[Insert table 3 about here.]

2.4 Method

2.4.1 Participants

Forty-two native German-speaking students (21 female, mean age 26.0 years (SD 2.7)) participated in the experiment. All subjects were right-handed, had a normal or corrected-to-normal vision and were paid for their participation.

2.4.2 Stimulus Material

As stimulus material a set of 180 experimental items and 60 filler items was used. All items were German prepositional phrases consisting of a preposition followed by an inflected adjective and a noun, e.g. *durch gute Arbeit* (‘by good work’). Four prepositions that uniquely govern accusative case were chosen: *durch* (‘by’), *ohne* (‘without’), *für* (‘for’), and *gegen* (‘against’). Singular neuter, feminine and masculine adjective noun combinations (20 noun phrases each) were used as completion of the prepositional phrase. All nouns were matched for frequency (half frequency class 11, half frequency class 12 according to http://wortschatz.uni-leipzig.de), length (on average length of 6 letters) and syllable structure (only disyllabic words), and differed only in gender. The nouns of each of the three genders included morphologically simple as well as derived forms. The overall proportion of derived forms within each gender group was kept as similar as possible across the three gropus (i.e., masculine: ca. 40%; feminine: ca. 50%; neuter: ca. 30%).[17] To ensure that potential ERP effects did not result from differences in plausibility and familiarity, two pretests were carried out on the stimuli. In a rating study, the familiarity of the prepositional phrases was examined. In this pretest, 18 participants rated the prepositional phrases on 6-point scale from unfamiliar to very familiar. On average, experimental items were rated with 3.6; differences between the items (F(1,179)=0.06, n.s.) were not found. In a further pretest, the degree of plausibility was tested in a rating using a 6-point scale from implausible to very plausible. Average plausibility of the items was 4.1. Differences in plausibility were not found (F(1,179)=0.47, n.s.). A complete list of the experimental material (correct conditions) is given in the Appendix.

For the morphosyntactic violation conditions, two additional incorrect versions were created for each of the 180 correct phrases: The morphological marker at the adjectival ending, i.e. /e/, /n/ or /s/, was manipulated in such a way that it resulted in a gender disagreement between the adjective and the noun. All three inflectional markers /n/, /e/ and /s/ were fully crossed, leading to nine experimental conditions (see table 4).

[Insert table 4 about here.]

Thus, the noun was critical for the grammaticality of the phrase since at this position it became clear whether the phrase structure was morpho-syntactically correct or not. Filler items were similar to experimental items and were all morpho-syntactically correct in order to balance the ratio of correct and incorrect items.

2.4.3 Procedure

Participants were tested individually in a 35 minutes session during which they were seated in a sound-proof, dimly lit cabin. Participants were asked to read attentively all prepositional phrases and to reply to the experimental task (a grammaticality judgment) as accurately as possible with *yes* or *no*. Participants had to judge whether the phrase was grammatical or not. Half of the trials were correct and half were incorrect. All responses were given via a button press. Prior to the experiment, participants received an instruction and a short training phase of six trials. For experimental presentation, the stimulus material was pseudo-randomized and divided into three versions of 180 items and 60 filler items each. In this way, the critical word of each experimental item was only presented once within each version. Experimental conditions were equally divided across all versions (i.e., 20 items of each condition). A trial sequence started off with the presentation of a fixation cross for 300 ms in the middle of the monitor. After an interstimulus interval (ISI) of 400 ms, the prepositional phrase was presented visually, word-by-word, centered on the screen with 400 ms for each word and an ISI of 300 ms. After stimulus presentation and an additional interval of 1500 ms subjects had to perform the experimental task. When the response was given (response time of maximal 3000 ms), and an intertrial-interval of 1000 ms ended, the next trial started. *Yes* and *no* answers were completely balanced across all experimental conditions.

2.4.4 Data Recording and Analysis

Behavioral data were analyzed in a repeated-measures ANOVA with the factors *Gender* (neuter/feminine/masculine) and *Marker* (correct/incorrect1/incorrect2). ERP data were extracted from the electroencephalogram (EEG) recorded from 52 Ag-AgCl electrodes according to the International 10-20 System of Electrode Placement. During the signal recording the EEG was referred to the right mastoid and afterwards re-referenced to the average of the left and right mastoids. The recording included bipolar horizontal and vertical EOG. Resistance of all electrodes was kept under 5 kΩ. EEG and EOG signals were recorded continuously with a band pass between DC and 70 Hz and a sampling rate of 500 Hz. Average ERPs were computed for the critical word (i.e., the noun) for each electrode position for each of the nine experimental conditions. Averages were aligned to a 200 ms pre-stimulus baseline and calculated for a period of 1000 ms after stimulus onset, and included only correctly answered trials. Approximately 8% of the trials were excluded from the averages due to ocular artifacts (automatic EOG rejection $\pm 40\,\mu$V).

For statistical analysis of ERP data two latency windows were employed: 300-550 ms (LAN) and 600-900 ms (P600). Repeated measures analyses of variance (ANOVAs) were performed on all dependent variables. For distributional ERP analyses two topographical factors *Site* (anterior/posterior) and *Hemisphere* (left/right) were defined and completely crossed, yielding 4 different Regions of Interest (ROIs) each containing nine electrodes.[18] Midline electrode positions were analyzed separately. In sum, analyses were carried out on four within-subject factors: *Site* (anterior/posterior), *Hemisphere* (left/right), *Gender* (neuter/feminine/masculine) and *Marker* (correct/incorrect1/incorrect2). In order to avoid potential violations of sphericity, the Greenhouse-Geisser correction (LIT) was applied to all effects with more than one degree of freedom. Whenever interactions between two or more factors were found, further analyses were carried out. Post-hoc tests using t-test for pairwise comparison were applied to analyze main effects of *Marker* and *Gender*. Due to multiple comparison all effects observed in post-hoc testing are corrected by the Bonferroni Holm procedure.

2.5 Results

2.5.1 Behavioral Data

Behavioral performance of participants on the grammaticality judgment was excellent and revealed a mean accuracy rate of 96.5% (SD 5.02). Statistical analysis showed a main effect of *Gender* ($F(2,82)=14.88$, $p<0.001$), a marginally significant effect of *Marker* ($F(2,82)=2.43$, $p<0.09$), and a significant interaction of *Gender* with *Marker* ($F(4,164)=3.41$, $p<0.01$).

Resolving the interaction of *Gender* with *Marker* by the factor *Gender* revealed main effects of *Marker* for feminine ($F(2,82)=3.97$, $p<0.02$) and masculine noun phrases ($F(2,82)=5.69$, $p<0.01$). Pairwise comparison of the different levels of *Marker* effects showed a significant difference between the two incorrect conditions ($F(1,41)=7.83$, $p<0.01$) for feminine noun phrases, which suggests that slightly more errors were made in the incorrect condition with an /n/ marked adjective (0.4% of all responses (SD 0.83)) than in the condition with an /s/ marked adjective (0.2% (SD 0.75)). For masculine noun phrases significant differences between the correct and first incorrect condition ($F(1,41)=8.37$, $p<0.01$), as well as the correct and second incorrect condition ($F(1,41)=9.33$, $p<0.01$) were found, indicating that correctly marked masculine noun phrases were judged better than both types of morpho-syntactically manipulated phrases (0.2% (SD 0.53) errors for correct marking versus 0.5% (SD 1.16) errors for incorrect marking by /e/ and 0.53% (SD 1.05) for incorrect marking by /s/).

2.5.2 Electrophysiological Data

ERPs at the critical word revealed a left anterior negativity (LAN) and late centroparietally distributed positivity (P600) for both morpho-syntactically manipulated conditions (i.e. incorrect1 and incorrect2) compared to the correct condition. Most interestingly, the amplitude of the LAN was modulated by gender. A LAN was evoked for manipulations of both feminine and neuter noun phrases, but not for manipulations of masculine noun phrases. In addition, a graded modulation of the LAN amplitude was seen for neuter noun phrases by showing the largest amplitude for the incorrect condition containing an /n/ marked adjective and medium amplitude for the incorrect condition containing an /e/ marked adjective.

2.5.2.1 The 300-550 Ms Latency Window The main analysis of the LAN time window showed a significant effect of *Marker* ($F(2,82)=20.11$, $p<0.001$) and a marginally significant effect of *Gender* ($F(2,82)=2.98$, $p<0.06$). A four-way interaction of *Site*, *Hemisphere*, *Gender*, and *Marker* ($F(4,164)=2.95$, $p<0.03$) was also significant. Resolving this interaction by *Site* showed for anterior electrode sites a main effect of *Marker* ($F(2,82)=19.37$, $p<0.001$) as well as a marginally significant three-way interaction of *Hemisphere*, *Gender* and *Marker* ($F(4,164)=2.02$, $p<0.10$). For posterior electrode sites a significant effect of *Marker* ($F(2,82)=11.26$, $p<0.001$) and a marginally significant interaction of *Gender* with *Marker* ($F(4,164)=2.37$, $p<0.06$) were present. The interactions found for anterior and posterior sites were further resolved, and the effects observed in the different domains are displayed in table 5.[19]

Our hypothesis clearly predicts different patterns for the processing of incorrect phrases according to the noun's gender. This difference is assumed to be due to different underlying feature structures. This means that the factor *Marker* concerns different manipulations for each gender (i.e., for feminines: correct, incompatible1, incompatible2; for neuters: correct, compatible, incompatible; for masculines: correct, compatible1, compatible2). Hence, the incorrect conditions cannot be compared directly between genders. We therefore analyzed the effects of *Marker* separately for each gender.

[Insert table 5 about here.]

For feminine and neuter noun phrases amplitude differences were observed between the correct and both incorrect phrase markings, validating the presence of a LAN effect for these two genders. This effect emerged on left and right anterior sites, and was also present on midline electrode sites, implying that the LAN was bilaterally distributed. In addition, for neuter noun phrases the negativity was confirmed on posterior sites; this suggests a more widespread distribution of the LAN seen for both morpho-syntactically manipulated conditions compared to the correct condition.

Most importantly, for neuter nouns pairwise comparison of the different marker levels revealed significant differences between the two incorrect conditions ($t(1,41)=4.17$, $p<0.05$), which substantiates a graded modulation of the LAN amplitude in response to manipulations of neuter noun phrases. In comparison to the correct condition, the largest LAN amplitude was seen for the incorrect condition containing an adjective marked with /n/; in contrast, a medium amplitude of LAN emerged for the incorrect condition containing an adjective marked with /e/.

However, while for neuter and feminine noun phrases a LAN was evoked for masculine noun phrases no such effect was present ($F(2,82)=0.23$, n.s.).

2.5.2.2 The 600-900 Ms Latency Window The main analysis of the P600 time window revealed a main effect of *Marker* ($F(2,82)=7.01$, $p<0.003$), and a marginally significant effect of *Gender* ($F(2,82)=2.42$, $p<0.10$). A three-way interaction of *Site*, *Gender* and *Marker* ($F(4,164)=3.36$, $p<0.02$) was also present. All effects yielded by the resolution of this interaction, as well as effects found for the midline electrode sites, are displayed in table 6.[20]

[Insert table 6 about here.]

On posterior sites a late positivity was confirmed for incorrect relative to correct phrase marking. Differences between incorrect phrase markings were not observed. This finding suggests the emergence of a P600 effect for both types of phrase marking violations. However, for masculine noun phrases the P600 in response to the incorrect phrase marking was also present on anterior sites, which suggests that the P600 amplitude was more prominent than P600 responses for equivalent manipulations of neuter and feminine noun phrases.

Figures 1, 2, and 3 visualize all observed effects for each gender. Consider first figure 1, which illustrates the ERP effects that can be detected with the processing of feminine noun phrases; here the correct marker is /e/, drawn as a solid line, and /n/ and /s/ are the incorrect markers, drawn as dashed and dotted lines, respectively).

[Insert figure 1 about here.]

Next, figure 2 documents the ERP effects showing up with the processing of neuter noun phrases. The correct marker is /s/ (solid lines); the incorrect markers are /n/ (dashed lines) and /e/ (dotted lines). Note the diverging dashed and dotted wave forms in the left-anterior area (e.g., at FC5), which will become important.

[Insert figure 2 about here.]

Finally, figure 3 illustrates the ERP effects occurring with masculine noun phrases; /n/ is the correct marker (solid lines), and /s/ and /e/ are the incorrect markers (dashed and dotted lines, respectively).

[Insert figure 3 about here.]

So far, one particular observation requires additional analysis. Phrases containing masculine nouns displayed quite an unusual pattern for correct forms. Correct forms did not differ from the incorrect versions in the left anterior region in the early time window. The obvious question is then: Are correct masculine forms special only in comparison to incorrect forms (with respect to a LAN effect) or do they also differ compared to correct forms of other genders? We therefore carried out an additional comparison for all correct phrases by gender. This comparison was possible since we had carefully matched the items of each of the three groups regarding length and frequency, and since we had also controlled for familiarity and semantic/pragmatic plausibility of the correct phrases. For the comparison of correct phrases between grammatical genders we obtained a significant difference for *Gender* for the anterior ROI in the early time window ($F(2,82)=7.55$, $p<0.001$). Subsequent analyses revealed that this effect was caused by a difference between masculine and neuter phrases ($F(1,41)=11.65$, $p<0.01$) as well as between masculine and feminine phrases ($F(1,41)=9.93$, $p<0.01$). However, feminine phrases did not differ from neuters ($F(1,41)=0.32$, n.s.).

Thus, our data indicate that correct masculine phrases show a greater anterior negativity in an early time window compared to other genders. This observation is illustrated in figure 4.

[Insert figure 4 about here.]

2.6 Discussion

2.6.1 Behavioral Data

As indicated by the results of the grammaticality judgment task, participants read the stimuli carefully and had no difficulty judging their correctness. The percentage of incorrect responses was marginal (about 3.5%). An effect of *Gender* was still found, which suggests that masculine noun phrases were more difficult to judge than feminine and neuter noun phrases. Note that the nouns did neither differ with respect to their frequency or word length, nor did the whole phrases differ in plausibility or familiarity. On the basis of the pretest data, what caused the effect of *Gender* must remain open to speculation. More importantly, though, the marking on the adjective in masculine noun phrases did not result in behavioral differences between the two incorrect conditions; this implies that judgments of grammaticality did not differ between these two manipulations. Differences between the two incorrect conditions were only seen for feminine noun phrases. However, for feminine noun phrases, no differences were observed between the correct marking and both types of incorrect marking. Therefore, this finding is not directly relevant for the question of underspecification. More generally, since the experimental task of providing grammaticality judgments can only detect delayed responses, differences between the two types of incorrect condition (which might then indicate differences between Compatiblity and Specificity) are not to be expected. It should be emphasized again that the behavioral task was included into the experiment solely to guarantee that participants read the stimuli thoroughly. Due to the particular characteristics of the task (e.g., the delayed response), the results here are not expected to be crucially influenced by the experimental manipulation. The case is different with online processing, where we expect that such effects might show up.

2.6.2 Electrophysiological Data

In our experiment we found two different kinds of ERP-evidence that support the hypothesis that underspecification of inflection markers is relevant in the actual language processing in the human brain.

2.6.2.1 Comparing Incorrect Conditions for Each Gender The first, more obvious, kind of evidence concerns differences in amplitudes between incorrect phrases for LAN and P600. Left anterior negativity modulated among different incorrect conditions only, but most crucially, for neuter nouns while effects of the centro-parietal P600 did not differ between incorrect conditions of any gender. The P600 effect that was invariantly present for all incorrect conditions for all genders implies a different processing of correct versus incorrect phrases. However, we did not find evidence that the functional processes underlying the P600 (i.e., reanalysis and repair) were differently affected by different genders or different types of violation.

The LAN effects, however, revealed a highly compelling pattern of differences for the three genders. As table 7 summarizes, the results clearly speak in favor of an underspecification-based model of morphological processing.[21]

[Insert table 7 about here.]

Closer scrutiny of the results reveals that the predictions of the maximally underspecified account fit the data perfectly.

For *feminine* nouns no model predicts any differences in the processing of incorrect forms. Accordingly, we did not observe any differences in participants' brain responses.[22]

The crucial case in our experiment, however, is that of *neuter* nouns. Here, predictions of categorical and underspecification-based accounts differ radically. As mentioned above, in a system of underspecified and competing inflection markers, there are two different criteria that have to be met for correct markers: Compatibility and Specificity. For neuter nouns, marker /n/ violates Compatibility while marker /e/ violates Specificity. The two criteria reflect fundamentally different principles and therefore induce different responses in brain activity when violated: Participants' EEG wave forms exhibit a greater negative amplitude when there is a conflicting feature (i.e., a violation of Compatibility), compared to an incorrect marker that is compatible in principle but excluded because it is not specific enough (i.e., that triggers a violation of Specificity).[23]

The processing of *masculine* nouns revealed no difference between the two incorrect conditions regarding the LAN. This was correctly predicted by a categorical account as well as by an account based on maximal underspecification since both markers /n/ and /e/ are compatible with the context. Note, however, that the brain responses for the two incorrect conditions in the LAN latency window did not differ significantly from the correct condition. Let us put this initially surprising result aside for the moment; we will come back to this issue in the next subsection. For the time being, we conclude that the results of the LAN effects for the incorrect conditions clearly mirror patterns predicted by underspecification-based accounts.

The pattern of the P600, which has been found to reflect processes of reanalysis and repair (see Neville et al. 1991, Friederici et al. 1993, and Friederici et al. 2001, among others), is different from the LAN. It could be interpreted to be compatible with a categorical distinction. As a first step towards accounting for the difference between LAN and P600 effects, it should be emphasized again that the two ERP components subsume *different functional processes* (see, e.g., Friederici 2002). The LAN is generally assumed to be sensitive to a first detection of morpho-syntactic violations, and it acts as a reliable indicator of processes of morpho-syntactic agreement, whereas the P600 component is sensitive to parsing problems, and taken to reflect semi-conscious processes of reanalysis and repair. Thus, different patterns for the two components indicate different underlying functional processes. The repair processes of the P600 component may operate on a less abstract level than morpho-syntactic error detection in the LAN component. It is also reasonable to assume that the kind of grammatical violation presented to the participants as well as the additional task (grammaticality judgment) induces complex processes of repair and respond strategies that may not be specific. Indeed, recently P600 effects have been reported for sentence-level semantic/thematic violations (see Kim & Osterhout 2005, Bornkessel-Schlesewsky & Schlesewsky 2008).

Summing up the results so far, the effects of the LAN for the processing of ungrammatical phrases with different sources of violation (Compability vs. Specificity) clearly indicate that underspecification is involved in actual language processing. Only an explanation that relies on two different kinds of incorrectness is consistent with all effects obtained for the LAN. This is crucially demonstrated by different amplitudes of the two incorrect conditions for neuter nouns. Moreover, it is the account based on *maximal underspecification* that is supported by the results because we obtained no such differences for masculine nouns, which would have been expected if the markers were *minimally underspecified* (see table 7). This finding is especially interesting from a theoretical point of view regarding the modelling of inflectional systems within underspecification-based accounts. Until now there were only conceptual and theory-internal reasons for favoring theories relying either on minimal or maximal underspecification of morphological exponents. Thus, not only is there evidence that actual language processing relies on underspecified inflection markers but, furthermore, it is a system with maximally reduced feature sets that seems to be favored by natural language processing.

2.6.2.2 An Effect of Feature Matching A second finding that can eventually be shown to support the notion of underspecified inflection markers (albeit in a somewhat more indirect way) is the very unusual pattern found for phrases comprising masculine nouns. Recall that comparing different morphological markers on adjectives preceding a masculine noun, we obtained no differences in left anterior regions for the early time-window. In contrast, for the two other genders there were reliable LAN effects for correct versus incorrect conditions. There are at least two possible explanations for this fact. First, it might be the case that the morpho-syntactic error was not detected by the system at all. This explanation, however, is quite unlikely and can be ruled out immediately by the fact that we obtained reliable P600 effects for incorrect conditions even for masculines. Furthermore, the grammaticality judgments indicated that participants had no difficulties detecting the error. A second possibility is that there is no observable difference because for the correct condition the amplitude of the negativity is also increased due to an independent factor, thereby masking the difference in processing between correct and incorrect forms. On this view, the two factors effectively cancel each other out.

A straightforward candidate for such a factor is the complexity of the marker's feature specification. As can be seen in (10), /n/ is a highly specific marker that is associated with a large set of features. We would thus like to propose an explanation of the missing LAN differences between correct and incorrect masculine phrases that is based on the feature specification of the inflection marker. More precisely, it is the high degree of specificity of the marker /n/ that induces the observed stronger negativity.[24]

Agreement evaluation of a highly specific marker, thus, may use up more processing resources than agreement evaluation of a less specific marker. This increased processing demand for highly specific markers neutralizes differences that otherwise would occur between correct and incorrect phrases.

We would like to consider this as an *effect of feature matching*. We suggest that the very process of feature matching is mirrored by the amplitude of the LAN. In what follows we illustrate this idea in more detail by providing a brief sketch of the mechanism we want to propose. We contend that the process of binary agreement in language comprehension consists of at least two functionally different steps. In a first step, the process has to check whether successively available pieces of information (i.e., features) are compatible. In a second step, missing information has to be integrated in order to derive a fully specified morpho-syntactic structure. More specifically, the mechanism may operate as follows in examples with a structure [$_{\text{PP}}$ P [$_{\text{NP}}$ A N]] that form the empirical basis of our study. First, the preposition is parsed. Since all P items we considered unambigiously assign accusative case, the feature specification [–obl,+obj] (= *accusative*) is available at this point (as government or selection information). Second, the pre-nominal adjective is encountered. Now the gender information associated with the adjective's inflection marker (if there is any such information in the first place) becomes available; this is [+masc,–fem] with /n/, [–fem] with /s/, and no gender feature with /e/. Furthemore, the adjective's case specification (if there is any) is now accessible. To carry out agreement (and ensure matching of the case specifications of P and A), the case features of the adjective are compared with the case features of the preposition.[25] If there are no conflicts, potentially missing case features of the preposition are copied onto the adjective, yielding full case specifications on A throughout from this point onwards. Third and finally, the noun enters the structure. The noun brings with it fully specified gender information and (irrelevantly for the present reasoning) fully specified number information, but presumably no case information whatsoever yet.[26] As before, the morpho-syntactic features of the new item – here, the noun – must be matched with the morpho-syntactic features of the preceding item – now, the adjective – so as to carry out agreement. The stage of the parsing process that captures this configuration is shown for masculine, neuter, and feminine NPs in (16-abc), respectively.[27]

(16) *Sizes of feature sets in well-formed NPs before A-N agreement*

a.	durch	schlichten	Geschmack	
	by	plain.MASC	taste.MASC	
	$\left[-obl, +obj \right]$	$\begin{bmatrix} +masc, -fem \\ -obl, +obj \end{bmatrix}$	$\left[+masc, -fem \right]$	
				← *most features compared:* 4/2
b.	durch	schlichtes	Design	
	by	plain.NEUT	design.NEUT	
	$\left[-obl, +obj \right]$	$\begin{bmatrix} -fem \\ \underline{-obl}, \underline{+obj} \end{bmatrix}$	$\left[-masc, -fem \right]$	
				← *fewer features compared:* 3/2
c.	durch	schlichte	Struktur	
	by	plain.FEM	structure.FEM	
	$\left[-obl, +obj \right]$	$\begin{bmatrix} \\ \underline{-obl}, \underline{+obj} \end{bmatrix}$	$\left[-masc, +fem \right]$	
				← *fewest features compared:* 2/2

With masculine NPs in (16-a), four features of the adjective (all of them inherent) must be compared with two gender features of the noun upon encountering the latter. In contrast, with neuter NPs in (16-b), only three features of the adjective (one inherent, two acquired from P, given strict maximal underspecification) must be compared with the two inherent noun features. Finally, with feminine NPs in (16-c), only two features of the adjective (none of them inherent) need to be compared with the noun's gender features. Assuming a categorical split between the first and the second/third configurations (which may plausibly further be increased by the qualitative difference pertaining to inherent vs. acquired features, with the latter being easier to process due to recent activation by matching of P and A), the LAN effect with well-formed masculine NPs can be accounted for as a result of additional morpho-syntactic processing effort.

The final step of morpho-syntactic feature integration that follows the matching operations just sketched then consists of copying the case features of the adjective onto the noun, and copying the gender (and number) features of the noun onto the adjective (where necessary), yielding fully specified structures throughout.

Note that this simple model of agreement evaluation may ultimately also account for the observed pattern of different kinds of incorrect phrases. The first stage of the evaluation (i.e., the comparison of the feature sets) is the most complex part of the mechanism. It may be assumed to be highly demanding on processing resources as it consists of symbolic computations, i.e., a pairwise comparison of feature sets. We may therefore expect that error detection at this stage can be regarded as a major violation. This is exactly in line with our finding that a violation of Compatibility leads to a stronger negativity in left anterior regions. In contrast, we may assume that the matching algorithm initially succeeds with illicit morphological exponents on the adjective that are compatible but insufficiently specific, with the choice of the wrong exponent only detectable by invoking Specificity. Assuming that Specificity is sensitive to the difference between inherent and acquired features, it can systematically block feature copying when matching exponents with fewer features are involved. Thus, violations of Specificity have a qualitatively different source in parsing than violations of Compatibility, and they may therefore be less demanding on processing resources. Consequently, a violation of Specificity is expected to show a reduced LAN compared to a violation of Compatibility.

On a more general note, we can conclude that the effects of LAN in our experiment does not reflect correctness or incorrectness of the phrase:[28] Rather, the present data indicate that the LAN is also amplified depending on the amount of features that have to be processed in order to establish an agreement relation. The left anterior negativity is thus observed even for the processing of correct agreement, modulated by the complexity of the morpho-syntactic feature structure. This effect emerges when a masculine adjective is followed by a masculine noun and their agreement has to be brought about.[29]

2.6.2.3 Underspecification and Optimal Design To sum up, the results of our experiment yield two different kinds of evidence for the notion of underspecified inflection markers. First, the differences with respect to LAN between two kinds of incorrect versions of neuter phrases can be explained by an underspecification-based approach, at least in a version that involves maximal underspecification. The distinct amplitudes mirror a difference in *quality*: a violation of Compatibility versus Specificity (i.e., an incorrect marker contains either compatible or incompatible features). Second, the unexpected effects of LANs for correct masculine forms additionally may be taken to suggest that underspecification is also crucially involved in morphological processing. This second finding reflects a difference in *quantity*: the number of features that have to be checked in adjective-noun agreement.

A straightforward consequence of this reasoning is that the degree of specificity of inflection markers is directly correlated with processing load in human language comprehension. This particular observation provides an interesting insight into the question why underspecified morphological exponents occur so frequently in the languages of the world. As mentioned in section 1, a system of underspecified markers at first sight would seem to increase processing demands. However, the results of the present study indicate that an underspecification-based system also leads to a reduction of computational effort – at least in certain contexts, viz., whenever (strongly) underspecified markers are processed. From this perspective, the processing of less specific inflection markers is – at least at a certain stage during language comprehension – less demanding for cognitive resources. Therefore, an inflectional system that consists of a set of underspecified markers can arguably be viewed as *optimal* from an optimal design perspective on language (cf. Chomsky 2005). This is so for at least two reasons: First, underspecification optimizes memory load, as the number of stored markers can be reduced. Second, the processing component of the system also benefits from less specific markers: Computational effort is reduced when the number of features that have to be compared/matched is kept small.

Notwithstanding the fact that there are several reasons to retain more specific markers in the system, such as, for instance, the option of an unambiguous encoding of grammatical information in certain contexts, there are benefits if the number of highly specific markers is reduced in a system in favor of less specific ones. Facilitated processing and memory effort with underspecified markers may be considered as a major motivation for having such markers in the system. This should be observable in at least two dimensions: diachronically and synchronically. On the one hand, there should be an incentive for less specific markers to spread in the system over time.[30] On the other hand, less specific markers should also be more persistent in impaired processing. These implications provide fascinating research topics; but we will leave the matter here.

3. Conclusion

Underspecification of inflection markers is a widely used concept in morphological theories. The main goal of the investigation reported in the present article was to find out whether this concept can be independently corroborated by neurophysiological evidence of language processing in the human brain.

The experiment on strong adjective inflection in German that we carried out shows that whenever agreement between an adjective and a noun has to be evaluated, early components in the EEG (i.e., the LAN) are sensitive to different kinds of violations that are inherently grounded in underspecification-based accounts: Compatibility versus Specificity. Furthermore, the results of our experiment provide additional evidence bearing on the question whether underspecification is maximal or minimal: The pattern of the LAN effects supports a modelling that relies on maximally underspecified inflection markers.

Furthermore, additional evidence from the processing of well-formed phrases supports the assumption that functional processes underlying the LAN may also be sensitive to different degrees of specificity of inflection markers, thereby providing a second argument for underspecification: Our results can be taken to indicate that more specific markers induce a stronger LAN than less specific markers.

To conclude, underspecification is not merely a well-supported concept in morphological theory; it may also explain neurophysiological responses of the human brain when processing language. Thus, it seems that underspecification is real, and that it is not an imperfection: It may qualify as optimal from an optimal design perspective on language as it actually reduces complexity for both the lexicon (minimizing storage costs for marker inventories) and the procedural part (minimizing computational effort in agreement rules).

Appendix: List of 3x60 Experimental Items (Correct Conditions)

	Prep	Adjective	neut NP	Adjective	fem NP	Adjective	masc NP
1	durch	leichtes	Fieber	leichte	Neigung	leichten	Akzent
2	durch	unerwartetes	Indiz	unerwartete	Störung	unerwarteten	Alarm
3	ohne	tägliches	Gehen	tägliche	Hölle	täglichen	Nebel
4	gegen	grausames	Elend	grausame	Seuche	grausamen	Befehl
5	ohne	echtes	Depot	echte	Achse	echten	Segen
6	ohne	geheimes	Rezept	geheime	Akte	geheimen	Bescheid
7	durch	festes	Terrain	feste	Bindung	festen	Beton
8	ohne	zusätzliches	Ressort	zusätzliche	Schwelle	zusätzlichen	Bonus
9	durch	andauerndes	Lernen	andauernde	Schande	andauernden	Boykott
10	ohne	bekanntes	Porträt	bekannte	Miene	bekannten	Doktor
11	gegen	erneutes	Exil	erneute	Panne	erneuten	Eklat
12	durch	umstrittenes	Plakat	umstrittene	Route	umstrittenen	Erlass
13	gegen	übermäßiges	Wollen	übermäßige	Härte	übermäßigen	Genuss
14	ohne	historisches	Archiv	historische	Predigt	historischen	Irrtum
15	durch	tatsächliches	Geschlecht	tatsächliche	Schwere	tatsächlichen	Tarif
16	gegen	modernes	Cockpit	moderne	Grafik	modernen	Kontext
17	für	deutliches	Echo	deutliche	Kopie	deutlichen	Kontrast
18	durch	langes	Tabu	lange	Allee	langen	Konvoi
19	durch	ungewohntes	Fehlen	ungewohnte	Handlung	ungewohnten	Luxus
20	durch	gemeinsames	Hobby	gemeinsame	Spende	gemeinsamen	Tenor
21	ohne	neues	Genre	neue	Sonde	neuen	Rabatt
22	für	berühmtes	Idol	berühmte	Kanzlei	berühmten	Rektor
23	durch	falsches	Futter	falsche	Treue	falschen	Report
24	ohne	weiteres	Shuttle	weitere	Übung	weiteren	Beleg
25	durch	staatliches	Lotto	staatliche	Haftung	staatlichen	Sektor
26	ohne	altes	Sofa	alte	Schüssel	alten	Slogan
27	für	ungewöhnliches	Zeugnis	ungewöhnliche	Ernte	ungewöhnlichen	Käse
28	gegen	geplantes	Logo	geplante	Zensur	geplanten	Transfer
29	gegen	großes	Klischee	große	Achtung	großen	Zirkus
30	ohne	goldenes	Cabrio	goldene	Krone	goldenen	Globus
31	ohne	vorgesehenes	Areal	vorgesehene	Kürzung	vorgesehenen	Balkon
32	gegen	teures	Benzin	teure	Substanz	teuren	Diesel
33	gegen	beliebtes	Cafe	beliebte	Beute	beliebten	Express
34	durch	genaues	Datum	genaue	Quote	genauen	Status
35	durch	erfolgreiches	Debüt	erfolgreiche	Gala	erfolgreichen	Kongress
36	für	politisches	Denken	politische	These	politischen	Konsens
37	durch	schlichtes	Design	schlichte	Struktur	schlichten	Geschmack
38	ohne	besonderes	Detail	besondere	Geste	besonderen	Applaus
39	ohne	kritisches	Drama	kritische	Lehre	kritischen	Faktor
40	durch	kulturelles	Erbe	kulturelle	Schwäche	kulturellen	Nutzen
41	ohne	allgemeines	Fazit	allgemeine	Panik	allgemeinen	Jubel
42	durch	einfaches	Format	einfache	Meldung	einfachen	Anzug
43	ohne	öffentliches	Forum	öffentliche	Moral	öffentlichen	Appell
44	durch	unverschämtes	Handeln	unverschämte	Bitte	unverschämten	Betrug
45	ohne	aktuelles	Journal	aktuelle	Fassung	aktuellen	Profi
46	durch	starkes	Rauchen	starke	Drohung	starken	Beifall
47	für	angenehmes	Klima	angenehme	Wärme	angenehmen	Atem
48	für	einzigartiges	Können	einzigartige	Rettung	einzigartigen	Triumph
49	durch	schnelles	Votum	schnelle	Warnung	schnellen	Kredit
50	durch	sonniges	Lachen	sonnige	Laune	sonnigen	Frühling
51	durch	fehlendes	Motiv	fehlende	Probe	fehlenden	Humor
52	ohne	auffälliges	Objekt	auffällige	Tendenz	auffälligen	Teppich
53	durch	beschlossenes	Paket	beschlossene	Sperre	beschlossenen	Etat
54	für	individuelles	Profil	individuelle	Note	individuellen	Fortschritt
55	für	bevorstehendes	Ringen	bevorstehende	Tagung	bevorstehenden	Export
56	durch	grenzenloses	Schweigen	grenzenlose	Stille	grenzenlosen	Hunger
57	durch	nationales	Symbol	nationale	Trauer	nationalen	Mythos
58	durch	beachtliches	Talent	beachtliche	Masse	beachtlichen	Effekt
59	ohne	begehrtes	Ticket	begehrte	Sparte	begehrten	Pokal
60	durch	offizielles	Labor	offizielle	Instanz	offiziellen	Transport

Bibliography

Anderson, Stephen (1992): *A-Morphous Morphology*. Cambridge University Press, Cambridge.

Aronoff, Mark (1994): *Morphology by Itself*. MIT Press, Cambridge, Mass.

Bader, Markus & Michael Meng (1999): Subject-Object Ambiguities in German Embedded Clauses: An Across-the-Board Comparison, *Journal of Psycholinguistic Research* 28, 121–143.

Baerman, Matthew, Dunstan Brown & Greville Corbett (2005): *The Syntax-Morphology Interface. A Study of Syncretism*. Cambridge University Press, Cambridge.

Bierwisch, Manfred (1967): Syntactic Features in Morphology: General Problems of So-Called Pronominal Inflection in German. In: *To Honor Roman Jakobson*. Mouton, The Hague/Paris, pp. 239–270.

Blevins, James (1995): Syncretism and Paradigmatic Opposition, *Linguistics and Philosophy* 18, 113–152.

Bornkessel-Schlesewsky, Ina & Matthias Schlesewsky (2008): An Alternative Perspective on Semantic P600 Effects in Language Comprehension, *Brain Research Reviews* 59, 55–73.

Butterworth, Brian (1983): *Language production: vol. 2, development, writing and other language processes*. Lonodon: Academic Press, chapter Lexical representation, pp. 257–294.

Caramazza, Alfonso, Alessandro Laudanna & Cristina Romani (1988): Lexical acess and inflectional morphology, *Cognition* 28, 297–332.

Chomsky, Noam (2005): Three Factors in Language Design, *Linguistic Inquiry* 36, 1–22.

Clahsen, Harald, Ingrid Sonnenstuhl & James Blevins (2003): Derivational Morphology in the German Mental Lexicon: a Dual-Mechanism Account. . Mouton de Gruyter, Berlin, pp. 125–155.

Clahsen, Harald, Sonja Eisenbeiss, Meike Hadler & Ingrid Sonnenstuhl (2001): The mental representation of inflected words: an experimental study of adjectives and verbs in German, *Language* 77, 510–543.

Corbett, Greville & Norman Fraser (1993): Network Morphology: A DATR Account of Russian Nominal Inflection, *Journal of Linguistics* 29, 113–142.

Di Sciullo, Anna Maria & Edwin Williams (1987): *On the Definition of Word*. MIT Press, Cambridge, Mass.

Eisenberg, Peter (2000): *Grundriß der deutschen Grammatik. Band 1: Das Wort*. Metzler, Stuttgart.

Embick, David (2003): Locality, Listedness, and Morphological Identity, *Studia Linguistica* 57, 143–169.

Fanselow, Gisbert (1991): Minimale Syntax. Habilitation thesis, Universität Passau.

Fraser, Norman & Greville Corbett (1994): Gender, Animacy, and Declensional Class Assignment: A Unified Account for Russian. In: G. Booij & J. van Marle, eds., *Yearbook of Morphology 1994*. Kluwer, Dordrecht, pp. 123–150.

Friederici, Angela D. (2002): Towards a Neural Basis of Auditory Sentence Comprehension, *Trends in Cognitive Sciences* 6, 78–84.

Friederici, Angela D. & Jürgen Weissenborn (2007): Mapping Sentence Form onto Meaning: The Syntax-Semantics Interface, *Brain Research* 1146, 50–58.

Friederici, Angela D. & Martin Meyer (2004): The Brain Knows the Difference: Two Types of Grammatical Violations, *Brain Research* 1000, 72–77.

Friederici, Angela D., Anja Hahne & Douglas Saddy (2002): Distinct Neurophysiological Patterns Reflecting Aspects of Syntactic Complexity and Syntactic Repair, *Journal of Psycholinguistic Research* 31(1), 45–63.

Friederici, Angela D., Axel Mecklinger, Kevin M. Spencer, Karsten Steinhauer & Emanuel Donchin (2001): Syntactic Parsing Preferences and Their On-Line Revisions: A Spatio-Temporal Analysis of Event-Related Brain Potentials, *Cognitive Brain Research* 11(2), 305–323.

Friederici, Angela D., Erdmut Pfeifer & Anja Hahne (1993): Event-Related Brain Potentials During Natural Speech Processing: Effects of Semantic, Morphological and Syntactic Violations, *Cognitive Brain Research* 1, 183–192.

Fuß, Eric (2005): *The Rise of Agreement*. Benjamins, Amsterdam.

Gallmann, Peter (2004): Feature Sharing in DPs. In: G. Müller, L. Gunkel & G. Zifonun, eds., *Explorations in Nominal Inflection*. Mouton de Gruyter, Berlin, pp. 121–160.

Goldberg, Adele E. (2006): *Constructions at Work*. Oxford University Press, Oxford.

Grimshaw, Jane (2001): Optimal Clitic Positions and the Lexicon in Romance Clitic Systems. In: G. Legendre, J. Grimshaw & S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 205–240.

Halle, Morris (1997): Distributed Morphology: Impoverishment and Fission. In: B. Bruening, Y. Kang & M. McGinnis, eds., *Papers at the Interface*. Vol. 30, MITWPL, pp. 425–449.

Halle, Morris & Alec Marantz (1993): Distributed Morphology and the Pieces of Inflection. In: K. Hale & S. J. Keyser, eds., *The View from Building 20*. MIT Press, Cambridge, Mass., pp. 111–176.

Harley, Heidi (2001): Lecture 11: Distributed Morphology. Halle & Marantz & Potawatomi Inflection. Ms., University of Arizona. Available from:
http://dingo.sbs.arizona.edu/ hharley/courses/ABRALIN/Lecture2Processes.pdf.

Harley, Heidi & Rolf Noyer (2003): Distributed Morphology. In: L. Cheng & R. Sybesma, eds., *The Second GLOT International State-of-the-Article Book*. Mouton de Gruyter, Berlin, pp. 463–496.

Jakobson, Roman (1962a): Beitrag zur allgemeinen Kasuslehre. Gesamtbedeutungen der russischen Kasus. In: *Selected Writings*. Vol. 2, Mouton, The Hague and Paris, pp. 23–71.

Jakobson, Roman (1962b): Morfologičeskije Nabljudenija. In: *Selected Writings*. Vol. 2, Mouton, The Hague and Paris, pp. 154–181.

Janssen, Ulrike & Martina Penke (2002): How are Inflectional Affixes Organized in the Mental Lexicon? Evidence from the Investigation of Agreement Errors in Agrammatic Aphasics, *Brain and Language* 81, 180–191.

Kaan, Edith (2007): Event-Related Potentials and Language Processing: A Brief Overview, *Language and Linguistics Compass* 1, 571–591.

Kim, Albert & Lee Osterhout (2005): The Independence of Combinatory Semantic Processing: Evidence from Event-Related Potentials, *Journal of Memory and Language* 52, 205–225.

King, Jonathan W. & Marta Kutas (1995): Who Did What and When? Using Word- and Clause-Level ERPs to Monitor Working Memory Usage in Reading, *Journal of Cognitive Neuroscience* 7(3), 376–395.

Kiparsky, Paul (1973): ‘Elsewhere’ in Phonology. In: S. Anderson & P. Kiparsky, eds., *A Festschrift for Morris Halle*. Academic Press, New York, pp. 93–106.

Kluender, Robert & Marta Kutas (1993): Bridging the Gap: Evidence from ERPs on the Processing of Unbounded Dependencies, *Journal of Cognitive Neuroscience* 5(2), 196–215.

Lumsden, John (1992): Underspecification in Grammatical and Natural Gender, *Linguistic Inquiry* 23, 469–486.

Marslen-Wilson, William (2007): Morphological Processes in Language Comprehension. In: G. Gaskell, ed., *Oxford Handbook of Psycholinguistics*. Oxford, pp. 175–193.

Mecklinger, Axel, Herbert Schriefers, Karsten Steinhauer & Angela D. Friederici (1995): Processing Relative Clauses Varying on Syntactic and Semantic Dimensions: An Analysis with Event-Related Potentials, *Memory and Cognition* 23, 477–495.

Müller, Gereon (2002): Remarks on Nominal Inflection in German. In: I. Kaufmann & B. Stiebels, eds., *More than Words: A Festschrift for Dieter Wunderlich*. Akademie Verlag, Berlin, pp. 113–145.

Müller, Gereon (2007): Notes on Paradigm Economy, *Morphology* 17, 1–38.

Müller, Gereon (2008): A Review of "The Syntax-Morphology Interface. A Study of Syncretism" by Matthew Baerman, Dunstan Brown, and Greville G. Corbett, *Word Structure* 1, 199–232.

Münte, Thomas F., Kolja Schiltz & Marta Kutas (1998): When Temporal Terms Belie Conceptual Order, *Nature* 395(6697), 71–73.

Neville, Helen J., Janet L. Nicol, Andrew Barss, Kenneth I. Forster & Merril F. Garrett (1991): Syntactically Based Sentence Processing Classes: Evidence from Event-Related Brain Potentials, *Journal of Cognitive Neuroscience* 3(2), 151–165.

Noyer, Rolf (1992): Features, Positions, and Affixes in Autonomous Morphological Structure. PhD thesis, MIT, Cambridge, Mass.

Noyer, Rolf (1997): *Features, Positions and Affixes in Autonomous Morphological Structure*. Garland Publishing, New York.

Penke, Martine, Ulrike Janssen & Sonja Eisenbeiss (2004): Psycholinguistic Evidence for the Underspecification of Morphosyntactic Features, *Brain and Language* 90, 423–433.

Pertsova, Katya (2007): Learning Form-Meaning Mappings in Presence of Homonymy: A Linguistically Motivated Model of Learning Inflection. PhD thesis, UCLA, Los Angeles.

Pinker, Steven (1991): Rules of Language, *Science* 253, 530–535.

Regel, Stefanie, Thomas C. Gunter, & Angela D. Friederici (2011): Isn't It Ironic? An Electrophysiological Exploration of Figurative Language Processing, *Journal of Cognitive Neuroscience* 23(2), 277–293.

Rossi, Sonja, Manfred F. Gugler, Anja Hahne & Angela D. Friederici (2005): When Word Category Information Encounters Morphosyntax: An ERP Study, *Neuroscience Letters* 384(3), 228–233.

Sauerland, Uli (1996): The Late Insertion of Germanic Inflection. Generals paper, MIT.

Schlesewsky, Matthias, Gisbert Fanselow, Reinhold Kliegl & Josef Krems (2000): The Subject Preference in the Processing of Locally Ambiguous Wh-Questions in German.. In: B. Hemforth & L. Konieczny, eds., *German Sentence Processing*. Kluwer, Dordrecht, Boston, pp. 65–93.

Schriefers, Herbert (1999): Morphology and Word Recognition. In: A. D. Friederici, ed., *Language Comprehension: A Biological Perspective*. 2nd edn, Berlin: Springer, pp. 101–132.

Schriefers, Herbert, Angela D. Friederici & Katja Kühn (1995): The Processing of Locally Ambiguous Relative Clauses in German., *Journal of Memory and Language* 34, 499–520.

Stolterfoht, Britta, Helga Gese & Claudia Maienborn (2010): Word Category Conversion Causes Processing Costs: Evidence from Adjectival Passives, *Psychonomic Bulletin and Review* 17, 651–656.

Stump, Gregory (2001): *Inflectional Morphology*. Cambridge University Press, Cambridge.

Taft, Marcus & Kenneth Forster (1975): Lexical storage and retrieval of prefixed words, *Journal of Verbal Learning and Verbal Behavior* 4, 638–647.

Tomasello, Michael (2003): *Constructing a Language. A Usage-Based Theory of Language Acquisition*. Harvard University Press, Cambridge, Mass.

Trommer, Jochen (2001): Distributed Optimality. PhD thesis, Universität Potsdam.

Trommer, Jochen (2005): Markiertheit und Verarmung. Ms., Universität Leipzig. Presented at the Honorary Doctorate Colloquium for Manfred Bierwisch, Leipzig 2005.

Trommer, Jochen (2006): Person and Number Agreement in Dumi, *Linguistics* 44, 1011–1057.

Wiese, Bernd (1996): Iconicity and Syncretism. On Pronominal Inflection in Modern German. In: R. Sackmann, ed., *Theoretical Linguistics and Grammatical Description*. Benjamins, Amsterdam, pp. 323–344.

Wiese, Bernd (1999): Unterspezifizierte Paradigmen. Form und Funktion in der pronominalen Deklination, *Linguistik Online* 4. (www.linguistik-online.de/3_99).

Williams, Edwin (1994): Remarks on Lexical Knowledge, *Lingua* 92, 7–34.

Williams, Edwin (1997): Blocking and Anaphora, *Linguistic Inquiry* 28, 577–628.

Wunderlich, Dieter (1996): Minimalist Morphology: The Role of Paradigms. In: G. Booij & J. van Marle, eds., *Yearbook of Morphology 1995*. Kluwer, Dordrecht, pp. 93–114.

Wunderlich, Dieter (1997a): Der unterspezifizierte Artikel. In: C. Dürscheid, K. H. Ramers & M. Schwarz, eds., *Sprache im Fokus*. Niemeyer, Tübingen, pp. 47–55.

Wunderlich, Dieter (1997b): A Minimalist Model of Inflectional Morphology. In: C. Wilder, H.-M. Gärtner & M. Bierwisch, eds., *The Role of Economy Principles in Linguistic Theory*. Akademie Verlag, Berlin, pp. 267–298.

Wunderlich, Dieter (2004): Is There Any Need for the Concept of Directional Syncretism?. In: G. Müller, L. Gunkel & G. Zifonun, eds., *Explorations in Nominal Inflection*. Mouton de Gruyter, Berlin, pp. 373–395.

Zifonun, Gisela (2001): Grammatik des Deutschen im europäischen Vergleich: Das Pronomen, Teil 1: Überblick und Personalpronomen. Amades working paper 4/01, IDS Mannheim.

Notes

[1]Since German is a verb-second language with variable placement of items in the pre-verb-second ('Vorfeld') position, there is no strict subject-verb-object surface word order. Hence, an object-verb-subject interpretation of the sentence is in principle possible, although this is typically not the canonical word order and it would often have to be prosodically marked. Moreover, there seem to be general parsing strategies such as a subject preference for the first argument that reduce the impact of such ambiguities (cf. Bader & Meng 1999; Mecklinger et al. 1995; Schlesewsky et al. 2000; Schriefers et al. 1995).

[2]This certainly holds for the nominativ/accusative syncretism with all neuters in Latin (cf. table 1) and virtually all other Indo-European languages (though see Corbett & Fraser 1993 and Fraser & Corbett 1994 on the marginal exception of animate neuter plural forms in Russian).

[3]Note that since pronominal inflection involves only closed-class items which are presumably stored as full forms in the mental lexicon, the experimental design made the choice of the strong adjective paradigm mandatory.

[4]The fourth combination that is logically possible – [+masc,+fem] – may be assumed to be unavailable due to conflicting gender information.

[5]Bierwisch originally had [±gov(erned)] instead of [±object]. This difference is purely terminological. We use [±object] here mainly because it may be a bit more widespread in the more recent literature.

[6]Here are the reasons for choosing a version of Blevins' approach over others. First, it reduces the number of different inflection markers to eight; other accounts (like that of Wiese (1999)) may have to envisage more exponents. Second, Blevins' account can easily be construed as involving *maximal* (rather than *minimal*) underspecification; this will become relevant below. Third, this approach avoids *trans-categorial* features, where a feature counts as trans-categorial if it simultaneously encodes information of two (or more) grammatical categories (e.g., gender and number features). As shown by Wiese (1999) and Trommer (2005), it is in principle possible to further reduce the number of inflection markers to the logical minimum of five exponents if one is prepared to adopt such features (e.g., all occurrences of *-er* could be assumed to go back to a single feature, call it [α], if [α] were to exclusively encode the contexts [nominative masculine singular], [dative/genitive feminine singular], and [genitive plural]). However, if trans-categorial features are permitted, the theory by necessity becomes completely unconstrained. Fourth, Blevins's analysis does not involve *disjunctive feature specifications*, in contrast to, e.g., Bierwisch 1967 and Wunderlich 1997a. Disjunctive specifications state (rather than derive) instances of syncretism. Fifth and finally, Blevins's approach does not have to stipulate an extrinsic order of markers (as it is done in Bierwisch 1967). – All that said, it should be pointed out that we have tacitly modified a number of minor aspects of Blevins' analysis; here the rationale was to adjust the gist of his approach to a more theory-neutral system (e.g., by dispensing with inheritance hierarchies).

[7]Some remarks on notation. First, the presentation of the association of form and content of an inflectional exponent by $\leftrightarrow$ is as in Distributed Morphology (see Halle & Marantz 1993), but this is purely for presentational purposes. It should be kept in mind that all that follows can equally well be expressed in any of the other current morphological theories that recognize underspecification (as a matter of fact, Blevins' analysis is originally formulated in a different morphological framework). Crucial differences between the various underspecification-based morphological theories (like the distinction between incremental and realizational organization, or the lexical/inferential divide; see Stump 2001) are completely orthogonal to our concerns here. Furthermore, note that the morphological exponents are rendered in an abstract, phonemic form (as indicated by the / / notation); thus, ə insertion that produces, say, *-en* from underlying /n/, is phonologically predictable, and should thus not be stated as a property of the marker form.

[8]Versions of this principle go under the names of *Subset Principle* (accompanied by a notion of specificity), *Elsewhere Principle, Blocking Principle, Pāṇini's Principle*, and *Proper Inclusion Principle*, to name but a few. See Kiparsky 1973, Di Sciullo & Williams 1987, Fanselow 1991, Anderson 1992, Lumsden 1992, Noyer 1992, Halle & Marantz 1993, Williams 1994, Halle 1997, Williams 1997, Wunderlich 1997b; 2004, Wiese 1999, Stump 2001, among many others.

[9]It has sometimes been argued that the notion of specificity should be taken to be more complex, e.g., by incorporating a hierarchy of features or feature types; see Lumsden 1992, Noyer 1997, and Wiese 1999, among others. For present purposes, we can abstract away from these complications, though.

[10]Three remarks are in order. First, the system as it is laid out here presupposes that gender is indeed neutralized in plural contexts; i.e., no grammatical gender information is present in the plural. This implies that exponents bearing gender features will always violate Compatibility in plural contexts.

Second, the analysis still leaves a number of marker homonymies: There are two /n/ exponents, two /s/ exponents, and two /r/ exponents, with different specifications in each case. As noted above, it is hard to see how these remaining kinds of unresolved syncretism could be accounted for in an underspecification-based approach if trans-categorial features are avoided. However, see Wunderlich 1997a, Wiese 1996, Eisenberg 2000, and Zifonun 2001 for arguments (based on distributional evidence) that exactly these instances of syncretism should in fact *not* be derived systematically; if these arguments are correct, a

Blevins-type approach that leaves three instances of syncretism unresolved is in fact optimal.

Third and finally, since we will mainly be concerned with strong adjective inflection in what follows, the question arises of how to integrate the /n/ exponent that replaces /s/ of the pronominal inflection there in genitive masculine singular and genitive neuter singular contexts. On the one hand, there does not seem to be a way to systematically derive the syncretism with accusative masculine singular /n/; accusative and genitive do not even form a natural class. (Of course, similar problems would arise with dative plural /n/.) On the other hand, distributional evidence as it has just been mentioned also argues against any attempts to treat the two kinds of /n/ markers as one; to wit, the syncretism breaks down when whole NPs are taken into account: *gut-en Wein-es* ('good$_{gen}$ wine$_{gen}$') vs. *gut-en Wein-Ø* ('good$_{acc}$ wine$_{acc}$). Thus, the obvious way to integrate this genitive /n/ into the system of exponents is to simply let it replace, and take over the feature specification of, /s/ in (4-c).

[11]Whereas this conclusion would seem to be entirely uncontroversial for underspecification and competition, it may not be completely obvious at first sight for feature decomposition, especially in view of the fact that underspecification can in principle take place without feature decomposition. However, it is shown in Müller 2008 that all the theories just mentioned regularly employ feature decomposition.

[12]Note that the /n/ marker is the most specific of the three markers not only in the analysis of Blevins (1995). The following short overview summarizes the number of features assigned to each of the markers in Bierwisch 1967; Blevins 1995; Wiese 1999 and Wunderlich 1997a:

	/n/	/s/	/e/
Bierwisch	2	0	1
Blevins	4	2	0
Wiese	2	2	0
Wunderlich	2	0	2

Comparing these four different analyses with respect to the number of features they assign to each of the three markers, a general trend can be detected: /n/ tends to be more specific than /s/ and /e/, and /e/ tends to be the most underspecified (default) marker of the system, although its status relative to /s/ is less clear – Bierwisch and Wunderlich identify /s/ as the elsewhere exponent. Arguments against such a view based on considerations related to iconicity are given in Wiese 1999 and Müller 2002.

[13]Incompatible features that give rise to Compatibility violations are set in bold face, here and in what follows.

[14]The additional features are set in italics.

[15]Note that /e/ could in principle also additionally be specified for gender features (e.g., [+fem]) in order to avoid competition with markers in the masculine and neuter cells of the paradigm. However, such a step would produce complications for the occurrence of /e/ in the nominative and accusative plural. There, /e/ is used for all genders. A correct insertion of /e/ for masculine and neuter forms would then be ruled out, even if we assume that gender features are not active in plural contexts in German, because of a Compatibility violation. This issue could be addressed by a trans-categorial feature decomposition that affects both feminine and plural (see Wiese 1999, Trommer 2005). However, such an approach would be ad hoc as it would blur conceptually well-founded feature categories (i.e., gender and number), and lead to a completely unrestricted theory (see above). In the analysis given here we follow the more conservative account and assume that gender and number features are only available separately. Thus, /e/ cannot be specified for gender features.

[16]Some notational conventions in table 3: ^{1} signals a violation of Specificity; ^{2} signals a violation of Compatibility; $\alpha{=}\beta$ indicates the same type of violation/the same processing; and $\alpha{<}\beta$ indicates a different type of violation/different processing.

[17]Note that the predicted effects were expected to be observed *within each gender* according to the manipulation of the adjectival form preceding the noun. Thus, although there is evidence that morphological derivation (i.e., category conversion) causes processing costs (see Stolterfoht et al. 2010), the minimally different proportions of derivations among the gender groups should not interefere when experimental conditions within each gender are compared.

[18]The four ROIs included the following electrodes: left anterior = F7, F5, F3, FT7, FC5, FC3, T7, C5, C3; right anterior = F4, F6, F8, FC4, FC6, FT8, C4, C6, T8; left parietal = TP7, CP5, CP3, P7, P5, P3, PO7, PO3, O1; right parietal = CP4, CP6, TP8, P4, P6, P8, PO4, PO8, O2; midline = AFZ, FZ, FCZ, CZ, CPZ, PZ, POZ, OZ.

[19]Abbreviations used in this table: cor = correct; incor1 = incorrect1; incor2 = incorrect2; (*) = $p<.10$; * = $p<.05$; ** = $p<.01$; *** = $p<.001$; n.s. = not significant.

[20]Note that, just as for the early latency window, the factor *Marker* had different levels for each gender. Thus we could not compare incorrect conditions independently of the factor *Gender*. Abbreviations are as before: cor = correct; incor1 = incorrect1; incor2 = incorrect2; (*) = $p<.10$; * = $p<.05$; ** = $p<.01$; *** = $p<.001$; n.s. = not significant.

[21]The notational conventions in table 7 are the same as in table 3; see note 16. In addition, * indicates a significantly different amplitude ($p<0.05$).

[22]Note that in underspecification-based approaches, following our hypothesis, both markers exhibit the *same kind of violation*, i.e., they violate Compatibility. But, in contrast to Penke et al. 2004, this violation is induced by positive ([+masc]) and negative ([–fem]) feature values in the same way. Thus, contrary to Penke et al. 2004, we do not find evidence that positive and negative feature values are treated differently in language processing.

[23]A purely phonological/graphematic source for the difference between the two incorrect conditions for neuter nouns can be ruled out. In principle, it is conceivable that the two different incorrect adjectival forms ending in '*-e*' and '*-en*' could induce different brain responses due to the number of their segments alone (forms marked with '*-en*' being one segment longer then forms marked with '*-e*' – at least if ə is maintained and the nasal does not become syllabic in *-en*). Such an account can be dismissed for two reasons: First, we did not compare the processing of the (differently marked) adjectives but the processing of the nouns, which were identical in each condition. A marginal phonological effect of the preceding adjective should certainly have disappeared by the time the noun is to be processed. And second, if there were phonological effects of the adjectives for neuter nouns, the same should hold for masculine nouns, for which both incorrect forms differ in exactly the same way in terms of the number of segments (cf. '*-es*' vs. '*-e*'). However, this phonological difference did not lead to different brain responses.

[24]Note that this reasoning is fully compatible with the observation that we also obtained LAN differences for correct phrases compared by *Gender*. Recall that in this comparison it is also the /n/ marked form on the adjective (i.e., the correct masculine phrase) that induces a stronger negativity on left anterior electrodes in an early time window, when compared with neuter and feminine phrases. It is reasonable to assume that the two observations are closely related.

[25]This step, by assumption, is more costly if more features need to be compared. However, since EEG wave forms were compared for the processing of the *noun* (and not the *adjective*), any such effect cannot be directly observed in the data.

[26]The nouns that were used in our experiment are all massively case-ambiguous between (at least) nominative, accusative, and dative, and therefore uniformly (i.e., across all genders) most likely radically underspecified with respect to case from a purely morphological point of view.

[27]Note that we here adopt the version of Blevins' approach in (12) based on maximal underspecification. However, adopting the approach in (4) would not substantially change things (the difference concerns the status of [–obl] in neuter contexts as derived or inherent). Note also that number features are ignored here to simplify exposition. Inherent features of lexical items and inflection markers are set in italics; features that show up as a result of agreement copying are underlined in addition.

[28]In contrast, correctness is directly mirrored as a function of later and more posterior effects (P600). Here, it is always the incorrect forms that show a larger amplitude compared to the correct condition.

[29]Note in passing that the source of this effect is an *interaction* of the properties of the adjective and the noun. It can neither be the /n/ marked adjective nor the masculine noun alone that is the cause of an increased negativity. On the one hand, if it were the masculine form of the adjective per se that causes the increased negativity, we would also expect to see differences for LAN when incorrect /s/ and /n/ marked forms for feminine nouns are compared. Crucially, this is not the case. On the other hand, we did not observe any main effect for *Gender* in left anterior regions as it would be expected if it were the properties of the masculine nouns alone that triggered the effect. Thus, the unusual pattern observed for correct masculine phrases is determined by an *interaction* of the properties of masculine marked adjectives and masculine nouns; therefore, the pattern has to be grounded in the notion of agreement.

[30]However, as noted by Fuß (2005), this tendency may also be counteracted by explicitness requirements favoring more specific exponents.

List of Tables

Case	*oppidum* 'town'	*dominus* 'master'
Nom sg	$\boxed{\text{oppidum}}$	dominus
Gen sg	oppidī	dominī
Dat sg	oppidō	dominō
Acc sg	$\boxed{\text{oppidum}}$	dominum
Abl sg	oppidō	dominō

Table 1: *Case syncretism in Latin*

	singular			plural		
	masc	neut	fem	masc	neut	fem
Nom	-er	-es	-e	-e	-e	-e
Acc	-en	-es	-e	-e	-e	-e
Dat	-em	-em	-er	-en	-en	-en
Gen	-en	-en	-er	-er	-er	-er

Table 2: *Paradigm of pronominal inflection in German*

		model		
		without underspec.	with underspecification	
noun gender		categorical	maximal	minimal
feminine	(*corr. e*)	$s = n$	$s^2 = n^2$	$s^2 = n^2$
neuter	(*corr. s*)	$e = n$	$e^1 < n^2$	$e^1 < n^2$
masculine	(*corr. n*)	$e = s$	$e^1 = s^1$	$e^1 < s^2$

Table 3: *Predictions: assumed processing differences for different incorrect markers*

	masculine NP 'without new discount'	neuter NP 'without new genre'	feminine NP 'without new probe'
correct	ohne neuen Rabatt	ohne neues Genre	ohne neue Sonde
incorrect 1	ohne neues Rabatt	ohne neuen Genre	ohne neuen Sonde
incorrect 2	ohne neue Rabatt	ohne neue Genre	ohne neues Sonde

Table 4: *Experimental design/conditions*

	Anterior Sites				Posterior Sites		Midline Sites	
	left anterior		right anterior					
	df	F	df	F	df	F	df	F
Marker	2,82	14.23***	2,82	17.71***	2,82	11.26**	2,82	14.25***
Gender x Marker	4,164	6.13**	4,164	2.49(*)	4,164	2.37(*)	4,164	2.39(*)
Feminine								
Marker	2,82	8.44***	2,82	6.19**	2.82	5.01**	2,82	3.02(*)
cor vs incor1	1,41	13.9***	1,41	8.64**	1,41	n.s.	1,41	n.s.
cor vs incor2	1,41	6.40*	1,41	8.74**	1,41	8.41**	1,41	n.s
incor1 vs incor2	1,41	n.s.	1,41	n.s.	1,41	n.s.	1,41	n.s.
Neuter								
Marker	2,82	17.49***	2.82	11.53***	2,82	7.87***	2,82	12.93***
cor vs incor1	1,41	36.59***	1,41	25.54***	1,41	10.71**	1,41	19.41***
cor vs incor2	1,41	12.78**	1,41	9.47**	1,41	10.12**	1,41	16.16**
incor1 vs incor2	1,41	4.17*	1.41	n.s.	1.41	n.s.	1,41	n.s.
Masculine								
Marker	2,82	n.s.	2,82	n.s.	2,82	n.s.	2,82	n.s.

Table 5: *Effects of the step-down ANOVAs for anterior and posterior sites and the ANOVAs for the midline sites of the 300–550 ms latency window*

	Anterior Sites		Posterior Sites		Midline Sites	
	df	F	df	F	df	F
Marker	2,82	n.s.	2,82	10.07***	2,82	6.146**
Gender	2,82	n.s.	2,82	3.45*	2,82	2.62(*)
Gender x Marker	4,164	3.72**	4,164	n.s.	4,164	2.44(*)
Feminine						
Marker	2,82	2.71(*)			2,82	5.42**
cor vs incor1	1,41	n.s.			1,41	n.s.
cor vs incor2	1,41	n.s.			1,41	9.40**
incor1 vs incor2	1,41	n.s.			1,41	n.s.
Neuter						
Marker	2,82	n.s.			2,82	n.s.
Masculine						
Marker	2,82	4.88**			2,82	5.21**
cor vs incor1	1,41	9.25**			1,41	12.73***
cor vs incor2	1,41	n.s.			1,41	n.s.
incor1 vs incor2	1,41	n.s.			1,41	n.s.

Table 6: *Effects of the step-down ANOVAs for anterior and posterior sites and the ANOVAs for the midline sites of the 600-900 ms latency window*

	Predictions			Results	
	without underspec.	with underspecification			
noun gender	categorical	maximal	minimal	LAN	P600
feminine	$s = n$	$s^2 = n^2$	$s^2 = n^2$	$s = n$	$s = n$
neuter	$e = n$	$e^1 < n^2$	$e^1 < n^2$	$e < n^*$	$e = n$
masculine	$e = s$	$e^1 = s^1$	$e^1 < s^2$	$e = s$	$e = s$

Table 7: *Predictions and results*

List of Figures

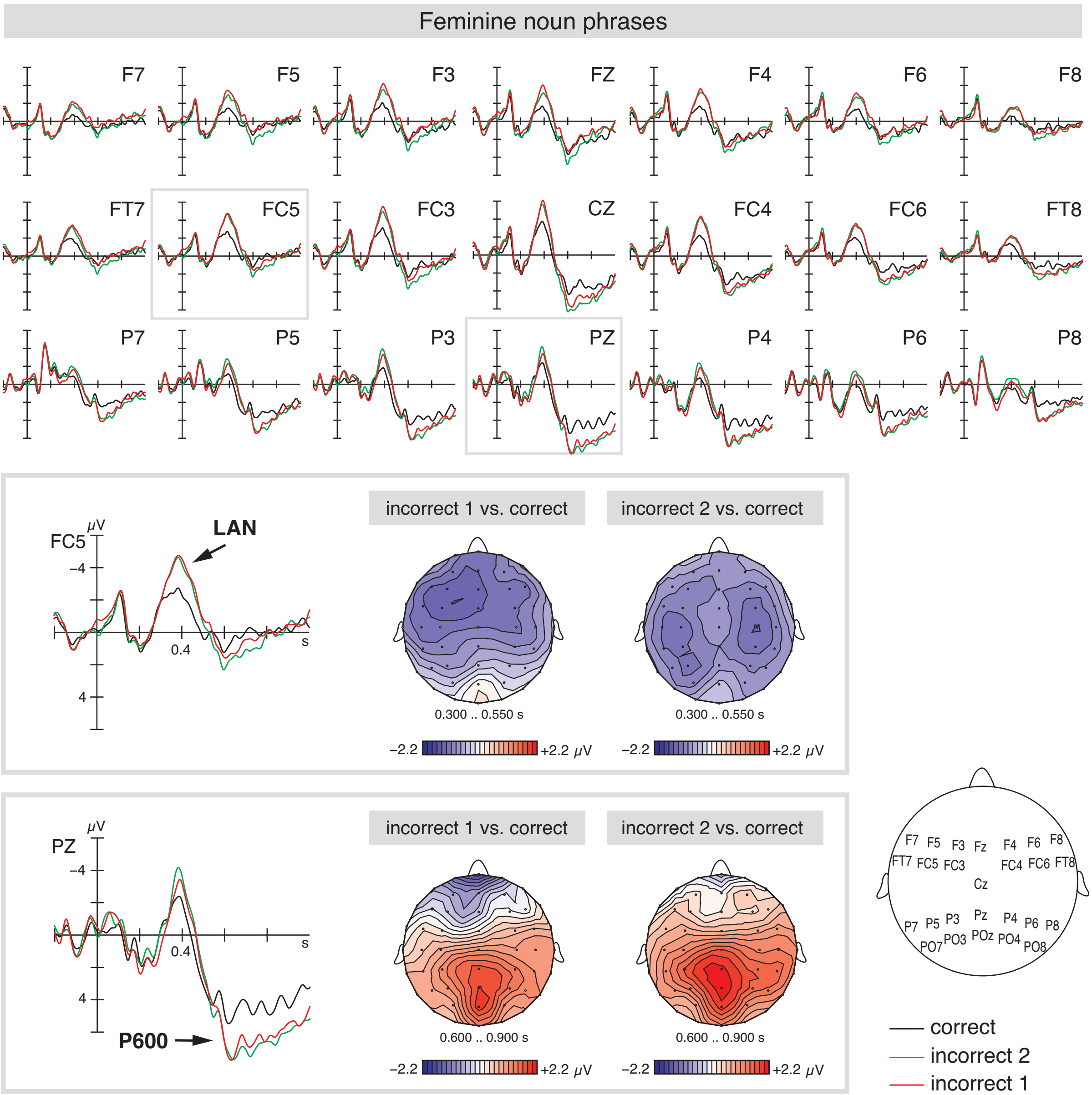

Figure 1: *ERP effects observed for feminine noun phrases*

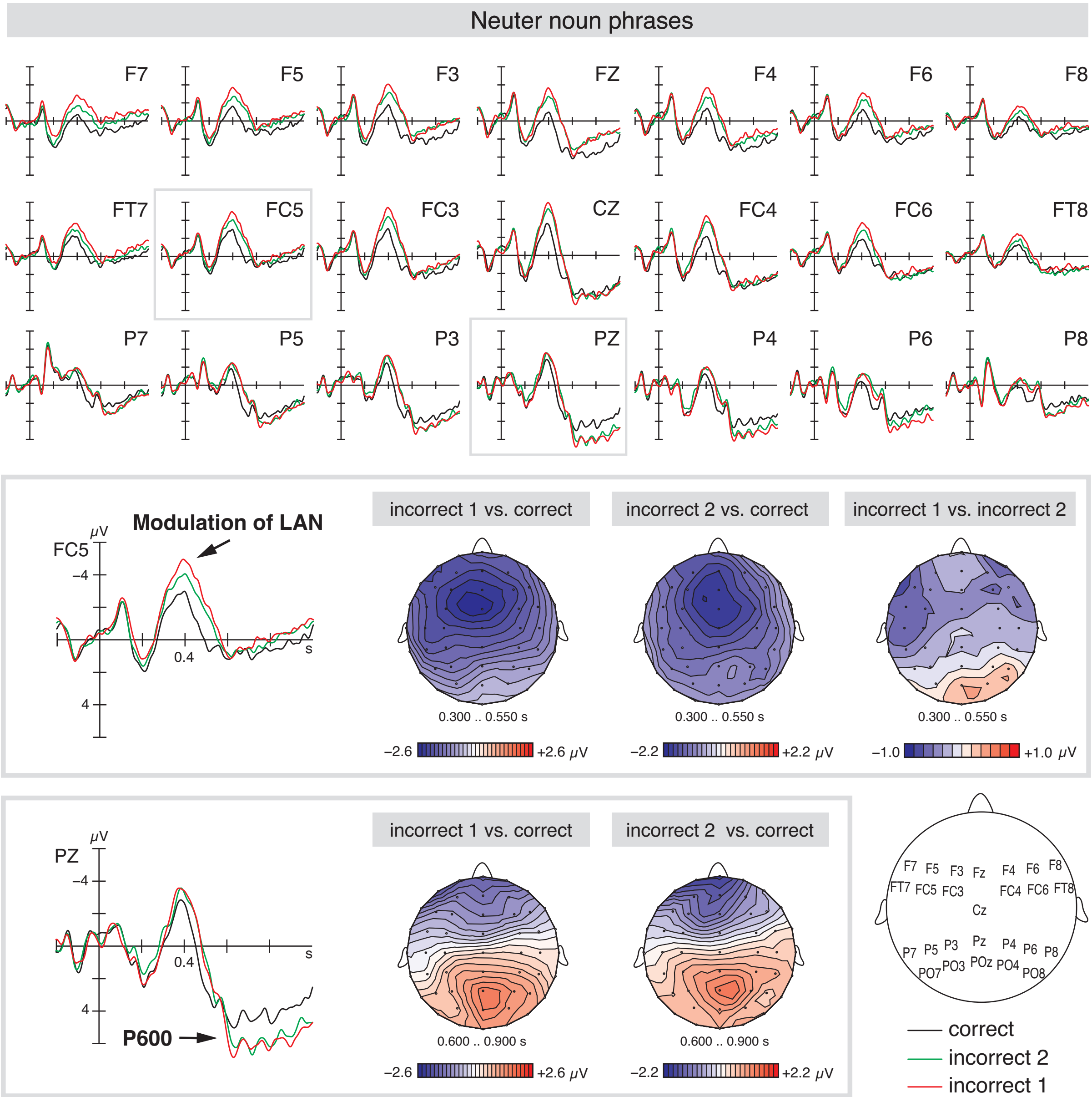

Figure 2: *ERP effects observed for neuter noun phrases*

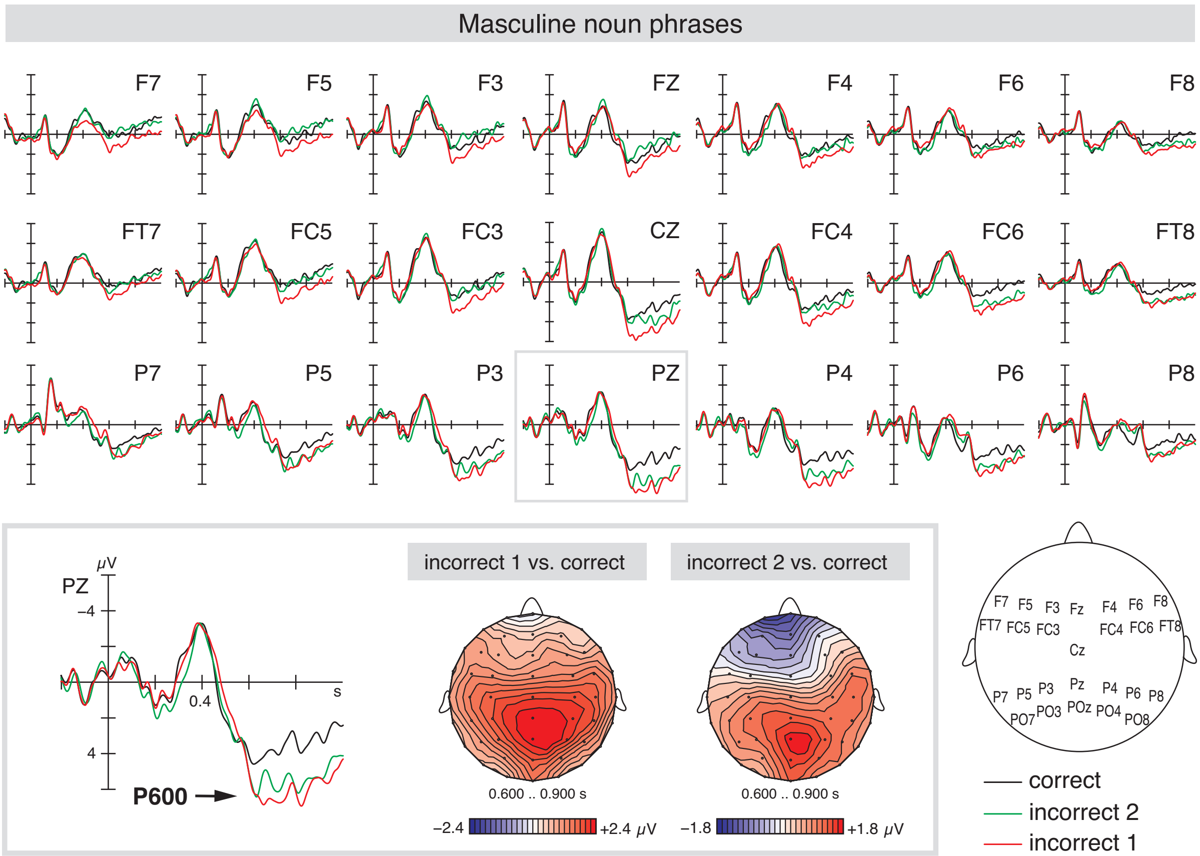

Figure 3: *ERP effects observed for masculine noun phrases*

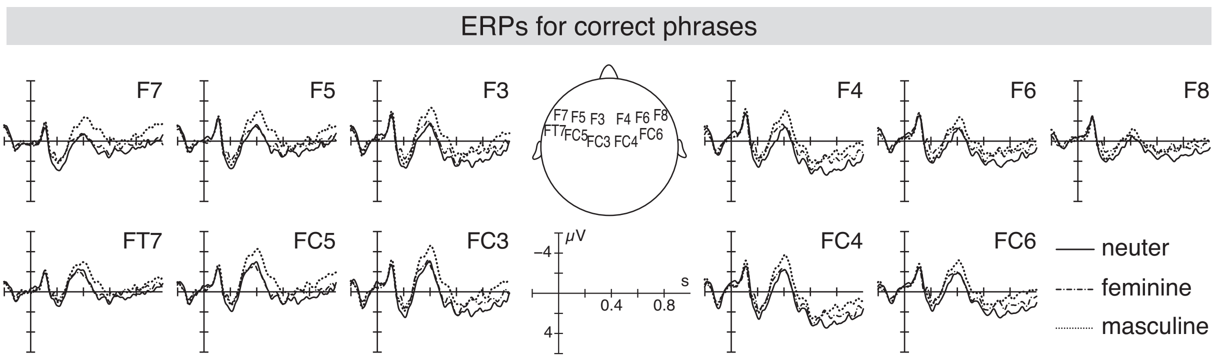

Figure 4: *Differences observed for the processing of correctly inflected phrases by gender.*