

Analysing Case in Spoken Language: A Corpus Study on East Franconian Dialects

Sophie Ellsäßer
University of Münster

Due to phonological and morphological processes, there has been an excessive reduction of case forms in the development of the German case system. Thus, syncretism is frequently found in Standard German and even more so in German dialects, where the process of case levelling is more advanced. In this sense, analysing dialects permits a diachronic perspective on German case systems since they show a more innovative behaviour compared to Standard German. Nevertheless, little research has been done on German dialectal case systems to date. This paper presents the method and initial results of a usage-based corpus study of case marking systems in East Franconian dialects, their interaction with animacy and their involvement in argument marking. One of the main issues of the study is the frequency of syncretic patterns which builds a basis for a comparative measurement of similarity between dialectal systems. This paper discusses a method of quantifying case distinctivity by relating it to a canonical system of case distinctions. The initial results of the study as well as the influence of syncretic case marking on argument marking will be discussed.¹

1. Introduction

As morphological phenomena of German dialects have not been well-researched, little is known about the structure as well as the spatial distribution of dialectal case systems. This is particularly surprising since distinct case marking performs a basic function in expressing argument structure. Since, however, levelling processes in inflection constantly shape German morphology in its diachrony, there is a large number of syncretic patterns in dialectal case systems. To evaluate the influence of syncretic patterns on the linguistic system of German dialects, I will investigate the role of distinct case forms in marking argument structure by analysing the frequency of syncretic versus distinctive patterns in different morphosyntactic and semantic contexts. The analysis is based on empirical data gathered from transcripts of a corpus of spoken dialects, which belong to the Upper German dialect region of East Franconian.² In this it will be possible not only to analyse the structure of dialectal case systems but also to evaluate the relevance of individual syncretic tendencies by examining their frequency of use in spoken language. As a result, any remaining case distinctions in German dialects can be quantified and ranked according to their frequency in spoken language.

This paper's primary focus is a dialectal system that is located in the East Franconian dialect area³ and has supposedly retained a rather high amount of case distinctions. Based on data from

¹ I wish to thank Antje Dammel, Jürg Fleischer, Sander Lestrade, Jeffrey Pheiff and the two other anonymous reviewers for their constructive comments on previous versions of this paper.

² The Upper German dialect area includes the dialects spoken in the south of Germany, German-speaking Switzerland and Austria. Depending on the definition and factors used in the classification, those are parts of the East Franconian, the Bavarian and the Alemannic dialects (see Wiesinger 1983).

³ The dialectal classification of the transcripts is taken from Ruoff (1984). Since most traditional approaches classifying dialectological areas are based primarily on phonological factors rather than morphosyntactic parameters, 'East Franconian' can be considered an umbrella term, which has the benefit of being a well-known term for the area rather than as an exact designation for a morphologically based region. Therefore, although some of the transcripts in the

this area, I outline my design of a usage-based analysis of case systems and introduce a method that is still untested in German dialectology when investigating morphological issues. Thus giving detailed information on syncretic and distinct case marking in certain contexts (with respect to word classes, word order patterns or semantic contexts) only plays a secondary role in this paper. The primary goal of this paper is to exemplify the possibilities in dialectological analysis which arise from the method used in the study.

The paper is organized as follows: To start with, I give a short overview on case, its role in argument marking (Section 1), and its expression in Standard German and in German dialects (Section 2). Section 3 describes the design of the usage-based corpus analysis, which is the basis of my empirical data. After giving an overview of some initial associations found in the sample in Section 4, I will discuss my findings and give an outlook for future research (Section 5).

2. Case syncretism in German dialects

2.1 Case as a marker of argument structure

Argument marking can be investigated on a semantic as well as on a syntactic level by analysing either semantic roles or syntactic functions. In prototypical contexts in German, there is a correlation between the agent and the function of the subject, the patient and the function of the direct object and the recipient and the indirect object. However, there are also several exceptions to this, such as predicatives or subjects in passive constructions (showing patient characteristics). Thus, as has been shown for several other languages (see Blake 2001: 3), there is no direct equivalence between syntactic function and semantic roles and case.

Occasionally, certain co-occurrences are frequently manifested: Since German is a nominative-accusative language, the nominative mostly expresses subject functions,⁴ while accusative expresses the (direct) object functions in transitive as well as in intransitive sentences. Alber & Rabanus (2011) investigate the mismatches of case distinction and syntactic as well as semantic argument structure in a comparative survey of pronominal paradigms in Germanic languages and dialects. By integrating animacy into their model, they add a semantic level that might influence argument marking as well as the morphological case system itself. They summarise the prototypical correspondences they found as follows:

thematic role	agent	patient	recipient, benefactive... ⁵
syntactic function	subject	object	indirect object
case	nominative	accusative	dative
animacy	animate	animate or inanimate	animate

Table 1: Prototypical correspondences of thematic role, syntactic function, case and animacy in Germanic pronominal case systems (from Alber & Rabanus 2011: 36)

Following Krifka (2009: 143), they argue that non-animates are canonically associated with the patient role while animates can express an agent as well as a patient role. They assume that inanimate referents provide a better context for syncretism than animate ones, since the factor of

sample might be classified as being Swabian, if different standards were applied, I am going to stick to the classification by Ruoff (1984) in this paper.

⁴ For a overview over the functions expressed by the cases in Standard German see Dürscheid (1999: 24–25). On the issue of relating semantic roles and syntactic functions in Standard German, see Primus (1999).

⁵ In the further statements as well as in the analysis, I will not refer to distinct semantic roles, but to proto-roles as first defined in Dowty (1991).

inanimacy is sufficient to predict argument structure – which allows levelling processes in morphological marking (Alber & Rabanus 2011: 32–33).

Hence, a hypothesis regarding the German dialectal system under investigation is that more distinct case marking will be found in expressing the agent and the subject, than in expressing the patient and the object. It can also be assumed that a certain association will be found between animacy and the distinctiveness of case marking.

2.2 Case syncretism in German: State of research

As mentioned above, the German case system contains a wide range of different syncretic tendencies determined by word class, gender, number and person. Due to phonological processes, only little case marking is left on the noun. Instead, case is mainly marked on pronouns, adjectives and determiners (see Blake 2001: 101). While there are more distinctions left on pronouns, syncretism has progressed further on determiners as well as on adjectives. Personal pronouns in particular are considered to resist case levelling for a longer time. The reason for this lies in their historically strong distinctive forms, structured as portmanteau morphemes and often showing a high degree of suppletion (see Howe 1996: 70).

Not only word class, but also number and gender affect case marking patterns in German. Due to general tendencies of number profiling in the paradigm, which promoted overt plural marking at the expense of the expression of other categories, case distinctions have been levelled in the plural paradigm. A particularly illustrative example is the distinction of plural nominative and accusative forms of demonstrative pronouns and definite articles, which have already levelled their case distinctions in Old High German (Braune & Reiffenstein 2004: 247) and their gender distinctions in Early New High German (Ebert et al 1993: 218) to a single form *die*. Moreover, in these two word classes as well as in personal pronouns, the nominative-accusative distinction has collapsed in all neuter and feminine relations in the New High German system, while it had already been levelled for plural forms of personal pronouns in Old High German (Paul 1959: 173–179). Especially neuter, but also feminine forms, tend to syncretism of nominative and accusative. While neuter has not shown a distinction of nominative and accusative at any stage of German, the strong tendency towards syncretism in feminine gender is remarkable. Since syncretism is often associated with lower animacy (cf. 2.1), one would rather only expect neuter forms to show a high amount of syncretism.⁶

In order to determine the expectations of syncretic tendencies in German case systems, the nominative-accusative syncretisms of neuter and feminine nouns are assumed for this analysis, since those tendencies already prevailed in Middle and Early New High German systems.

Syncretic patterns of German dialects are even more extensive than those of Standard German. Synthetic genitive forms are restricted to a few dialects in the south-western part of the German-speaking areas. In the majority of the German speaking dialect areas, the genitive has been levelled completely. Thus, I generally assume a maximum of three synthetic cases, nominative, accusative and dative, for German dialects. Those cases show different tendencies towards syncretism in different dialect areas, as illustrated in the following map based on Shrier (1965):

⁶ For further discussion on this issue in general see Bittner (2002). Krifka (2009, 155–158) discusses feminine forms being more likely to express inanimate referents in Standard German. Thus, Krifka (2009) suggests this observation to constitute one of the diverse factors promoting the syncretic development of German feminines.

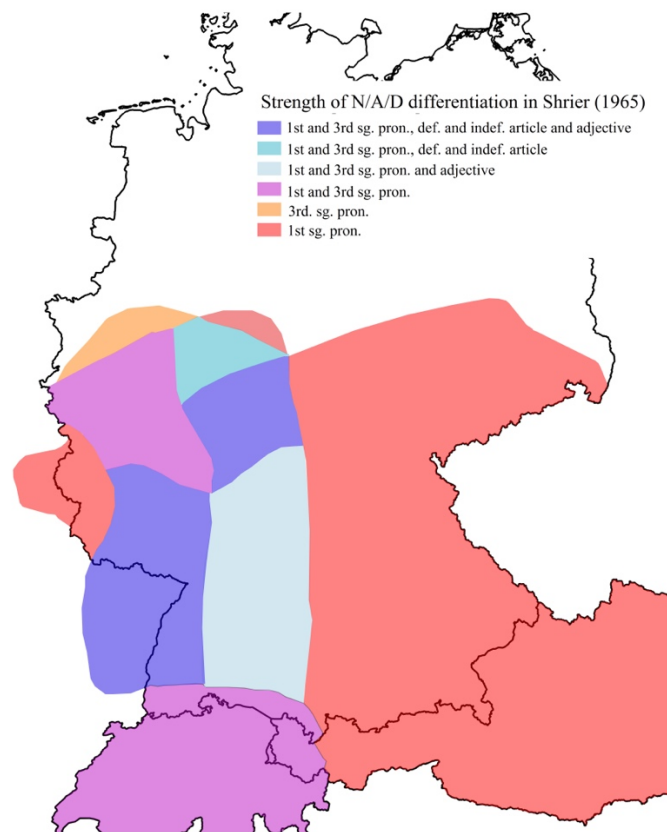


Figure 1: Occurrence of three-way case distinction in German dialects (Shrier 1965: map 13).⁷

Figure 1 illustrates the contexts which show a full distinction of nominative, accusative and dative in the respective areas. Basically, when looking at the information given on the map, a major divide is assumed between the Low German dialects in the north that have a maximum distinction of a nominative and an oblique case, and the High German dialects in the centre and in the south, that retain a distinction of nominative, accusative and dative in some parts of speech. This can be recognized in the transition of white and coloured areas in Figure 1.

The southern regions, however, do not form a homogeneous area with respect to case distinction. In High German dialects, there is an isogloss assumed to run from north to south, approximately along the political border of Baden-Württemberg and Bavaria. It splits up the area into an eastern part, where the three-way case distinction is limited to the first (and second) person singular pronoun and there is a tendency of an accusative-dative syncretism in masculine parts of speech, and a western part, where the three-way case distinction occurs in more pronominal constructions and where a tendency towards nominative-accusative syncretism in masculine items has been observed (see Shrier 1965: 434–435).

⁷ Most studies on German dialectal case systems are based on the article by Shrier (1965). Working with the tendencies in case syncretism stated in the article, special attention has to be paid to her investigation method. The article is based on a rather small amount of secondary data, gathered in an investigation of dialect grammars.

Previous studies suggest a certain interaction of the dialectal case system with syntactic and semantic phenomena such as animacy and agentivity and word order regularities (see Werlen 1990: 170–171, Rohdenburg 1998 and Dal Negro 2004). Hence, in order to capture the sensitivity of the dialectal case system towards these phenomena, it is not sufficient to only analyse a single extract of case marking, such as, e.g., a paradigm in a dialect grammar, but one needs to investigate and quantify a larger amount of case markers occurring in different syntactic and semantic contexts in each linguistic system in use.

3. A usage-based study on case marking in Upper German

3.1 A corpus on language in everyday life: “Alltagstexte” by Ruoff (1984)

The audio recordings that form the basis of the present study were collected and transcribed by Arno Ruoff and Hermann Bausinger from the 1950s up to the 1970s. The sample of transcripts was published in Ruoff (1984) and contains data from the Alemannic, Swabian, East Franconian and Bavarian dialect regions. This article focusses on a sample of East Franconian data. It contains 18 transcripts of audio recordings of dialectal spoken data of between 12 and 15 minutes each. The special property of the corpus is its unique transcription method that is derived from Standard German orthography (which forms a good basis for digital processing) and yet conveys all relevant morphosyntactic information by using diacritics and marking vowel alternation (Ruoff 1973: 129–130).

In working with this corpus, the purpose of the transcripts should also be considered. The publication was designed to show the variety of different social data as well as text types gathered in the survey. Therefore, the Ruoff corpus is not a consistently structured corpus from a sociolinguistic point of view. Moreover, the data were transcribed by different staff members and may contain differences in notational habit for example. Nevertheless, since all of the transcripts analysed in this sample are basically dialectal, i.e., represent the dialect of immobile local speakers, the corpus should meet the requirements for an exploratory study of dialectal case systems. Since the corpus primarily contains initiated monologues, each transcript can be seen as an excerpt of an idiolectal morphosyntactic system of an individual speaker in one location. Hence, in the event of presumed sociolinguistic impacts on the results, it is possible to attribute them to certain speaker’s backgrounds.⁸

Moreover, due to the fact that little is known about the spatial distribution of morphological phenomena this situation offers a good starting point for the analysis. Thus, it is possible to start with a small-scale examination of idiolectal case systems and then consider large-scale connections in a second step.

In addition, in analysing idiolectal systems there is no need to cope with interpersonal variation caused by personal characteristics of different speakers in the transcriptions. Syncretism and distinctions are classified on the basis of information exclusively given from one single idiolect (see also 3.2). Hence, in event of differing transcriptional habits, assuming that transcriptors are properly marking differences and similarities, the precise form of a single marker (which might differ due to personal habits) is less relevant than the indication of its distinction from other

⁸ The initial results presented in Section 4 indicate that these inconsistencies in the sociolinguistic structure of the corpus exert little influence on the use of syncretic and distinct case marking in the sample.

markers. Therefore, by analysing syncretism and distinction based on the internal structure of idiolectal systems, the findings in individual transcripts remain comparable.

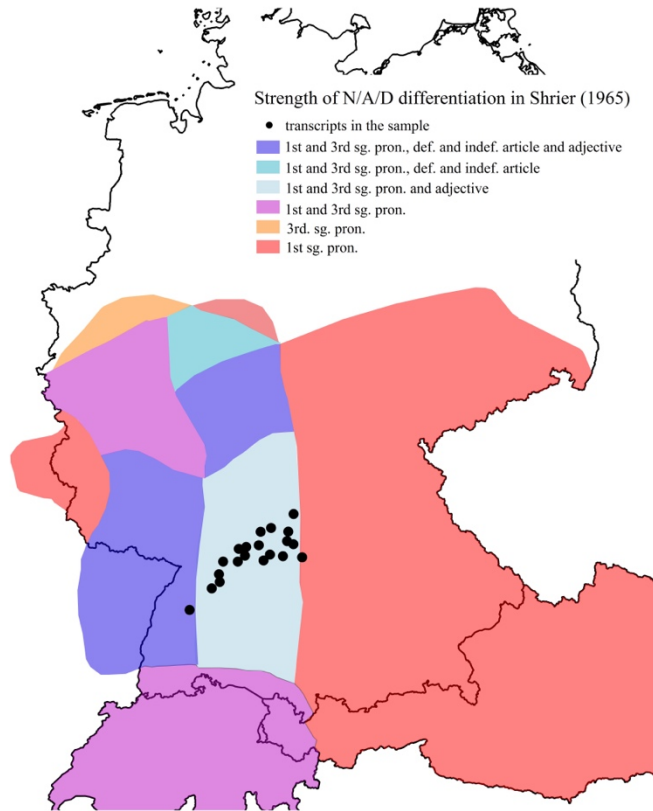


Figure 2: East Franconian transcripts in the Ruoff corpus projected on map 13 from Shrier (1965).

The present sample of East Franconian dialects was chosen for this study because these dialects, according to Shrier (1965), still retain a large amount of case distinctions compared to other Upper German varieties. As can be seen in Figure 2, the system is assumed to show a distinction of all three cases in first and third person singular pronouns as well as in adjectives.

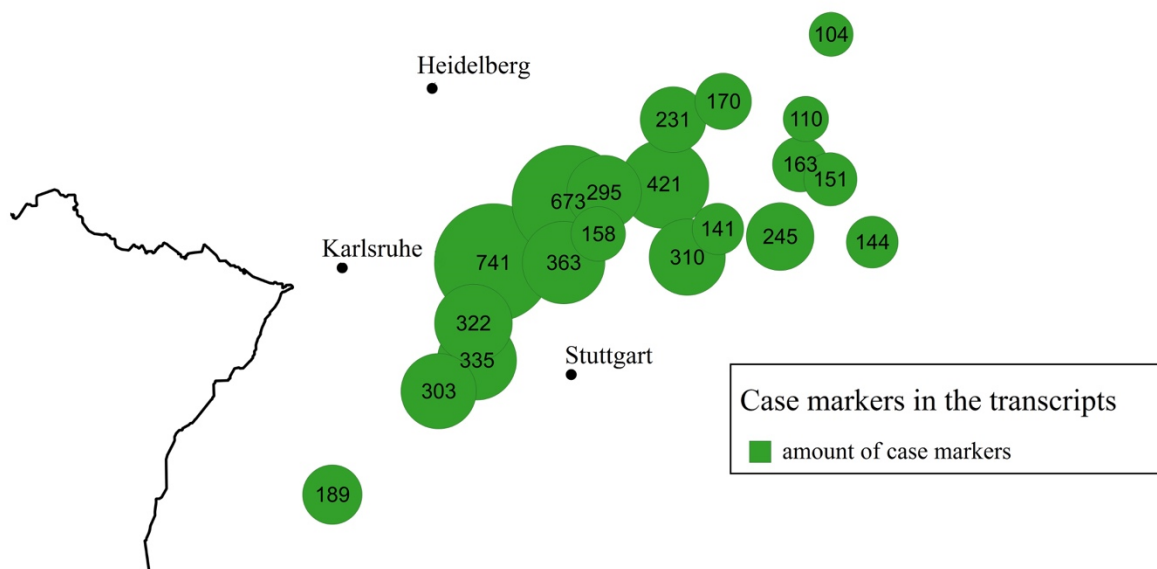


Figure 3: Amount of case marking word forms in the East Franconian transcripts in the Ruoff corpus.

Figure 3 gives an impression of the amount of case markers analysed in the transcripts. The total amount of 5592 markers, which includes personal pronouns, demonstrative pronouns, possessive pronouns, reflexive pronouns, indefinite pronouns, definite and indefinite articles as well as adjectives, provides a solid base for statements on first results and a statistical analysis on general associations.

3.2 Deviation and canonical forms: Comparing syncretic patterns

As sketched above, this paper introduces a usage-based corpus study of transcripts based on spoken dialectal language (see section 3.1). Since it is a word-based study, each word form that can potentially mark case, i.e., pronouns, adjectives or definite or indefinite articles, is included in the database and analysed with respect to the proto-role, the syntactic function that it expresses and the animacy of its referent. In doing so, the syntactic and semantic context of the case marker is also classified.

Besides arguing for a certain correlation of syncretism, patient relations, objects and a less animate referent (see section 2.1), Alber & Rabanus (2011) introduce an interesting approach on quantifying case marking systems: In order to compare different Germanic pronominal paradigms and to investigate the interaction of case marking and animacy, they calculate a syncretism rate for each pronominal paradigm in a language (or dialect) system. As the basis of comparison, they construe a hypothetical paradigm with a distinction of nominative, accusative and dative in each paradigmatic row as a reference system. Each syncretism of two of the three cells in a row is counted as a rate of 50%, while a syncretism of all three cells is counted as 100% (Alber & Rabanus 2011: 27–30). Therefore, it is possible to compare case marking systems by giving empirically based statements on the possibility of syncretism in the systems. This approach proves to be particularly interesting for the analysis presented here, since it develops a method of quantifying syncretism by measuring the system's deviation from a canonical ideal. Nevertheless, since I focus

on the frequency of syncretic patterns in spoken language, rather than on the possibility of syncretism within single paradigms, I am not using syncretism rates but rather ratios of absolute numbers in the analysis.

In order to establish a comparable basis to mark the extent of case distinction left in each idiolectal system, I use the method of *Canonical Typology* (see Brown & Chumakina 2013) in this analysis to create a standardized benchmark. A canonical case system can be used as a constant reference point for the entire area examined in the study, although it will not be mirrored empirically by any of the actual case systems (see Brown & Chumakina 2013: 9).

As this analysis takes up two issues, *Canonical Typology* is used in two ways here: A canonical ideal is used in a first analytical step to analyse canonical case forms to capture the case that is expressed by each case marking form. For methodological reasons, this ideal is not used to assign whether these forms are syncretic or distinct. In fact, a second system that is aligned to the characteristics of the German proto-system and is therefore missing some features of the canonical ideal, is used as a reference point in the second analytical step.

Since it is quite difficult to form a dialectal or diachronic reference system of each idiolect for this kind of classification, I opt for an abstract, canonical ideal with a maximal number of distinctions of accusative, nominative and dative for each word class, number and gender, related to the hypothetical paradigm used by Alber & Rabanus (2011) in investigating Germanic pronouns. According to Baerman et al. (2005: 35) values of a feature are seen as syntactically relevant in canonical syncretism (loss of certain but not all values of the feature in special contexts) as long as they are distinguished in other parts of the morphosyntactic system. Since, despite the syncretic patterns, the three cases are distinguished in some parts of the Upper German case systems, they are regarded as relevant for the morphosyntactic system and therefore are included into the canonical ideal of the case system.

	singular	plural
nominative	f _a	f _d
accusative	f _b	f _e
dative	f _c	f _f

Table 2: Distinction of forms in the canonical case marking system

Hence, a plural article form would be analysed by the canonical case demanded by the morphosyntactic context, although there is no distinct case marking left in any of the plural articles in any contemporary German system. Thus, I attempt to capture each context of syncretism or distinction, even those forming a total syncretism in all contexts and all systems. In this way, implicational hierarchies of case distinction versus syncretism in different grammatical contexts can be recognized. Thus, the canonical ideal constitutes a system which, in each paradigm cell, marks the arguments by a distinct case form.

In the analysis, each case marker is allocated to a “canonical case”. This is best explained by using an example:

- (1) *na habe se halt a sau(e)re Milch gässa,*
 so aux **theyPERS.PRON.NOM/ACC.PL** PART INDEF curdled milk eatPP
 ‘So they ate curdled milk then.’ (Eschenau/Schwäbisch Hall, Ruoff 1984: 18)

- (2) *die hollt se jeden Däch, d'Millich.*

she fetch **it**PERS.PRON.NOM/ACC.F.SG each day DEF=milk
 'She fetches it each day, the milk.' (Bieringen/Künzelsau, Ruoff 1984: 3)

In both examples, (1) and (2), the same form *se* is used. Since the forms are syncretic in both contexts, one cannot tell from the word form whether they express nominative or accusative case. However, since in (1), the morphosyntactic context, especially verbal alignment, demands a nominative, the potential case of the form *se* in (1) is classified as canonical nominative, while the form in (2) is classified as canonical accusative. Thus, the category of canonical case cannot be directly deduced from the phonological representation, but represents a first step of classification in the analysis.

After that, each idiolect's case markers are analysed according to their distinctiveness and then divided into syncretic and distinct markers. If the same form is found at least twice in an idiolect, differing only in the feature of canonical case, it is classified as syncretic. If there is evidence for other forms only differing in canonical case and there is no syncretic form found in the whole idiolectal system, it is classified as distinct. The following example illustrates two determiners that are classified as being distinct. (3) expresses a canonical nominative und (4) a canonical accusative. Although they exhibit different forms, they both belong to the same idiolectal system.

(3) *wo's eersch' Mâl der Zuu(ch) g'fâahra is.*
 when=DEF first time DEF.NOM.M.SG train drivePP AUX
 'when the train ran for the first time' (Bieringen/Künzelsau, Ruoff 1984: 3)

(4) *wail m'r den Dinkel schnaida hât missa.*
 because one DEF.ACC.M.SG spelt cutINF AUX mustINF
 'because one had to cut the spelt' (Bieringen/Künzelsau, Ruoff 1984: 1)

If there is no clear evidence that a particular word form is distinct or syncretic, a problem that may be caused by the context of the corpus, the form is excluded from the analysis to avoid distorting the analysis with assumptions. In order not to exclude too much evidence from the sample in this manner, particular forms are classified as being syncretic, even if there might be no form only differing on canonical case in the idiolect. This exception affects nominative and accusative forms of feminine and neuter definite articles in the singular, as well as definite articles and third person personal pronouns in the plural. This is justified as those items are assumed to be syncretic already in former German systems (see section 2.2). Thus, while the analysis of canonical case is strictly referring to the canonical ideal, the classification of distinctiveness is filtered through a system that is already known to miss certain features of the ideal.

Hence, the canonical case reflects the potential distinctions, regardless of the dialectal system and thus enables one to compare the findings of this analysis to other case systems. In contrast, the analysis of syncretism and distinction refers to the formal representation and therefore is oriented towards a more restricted system that is partly based on syncretism constellations arisen diachronically and that are shared by all German dialects. The latter is therefore used for methodological reasons. When working with a corpus, the finiteness of data is to be considered. Depending on the length of the recordings as well as on the content of the speech recorded, the transcripts only show subsections of the idiolectal systems. These are often incomplete.

Nevertheless, the corpus-based method developed for this study does have some clear benefits when compared to previous studies: apart from being able to analyse distinct and syncretic patterns

in a quite dense geographical grid, no previous assumptions on either the exact distribution of case marking patterns or their interaction with other factors are needed. By investigating a corpus of spoken data, it is possible to capture naturally occurring patterns in spoken language without having to trigger special contexts.

Furthermore, the method allows statements about the frequency of syncretic and distinct patterns and thus, to classify their relevance for argument marking in the respective system on an empirical basis. The rationale underlying this assumption is that the more frequent a syncretic pattern is in the corpus, the more it affects the morphosyntactic system. Hence, when I compare different systems in a further step, the measure of morphosyntactic proximity will rather depend on the proximity of frequent patterns than on less frequent ones. Moreover, especially with regard to the purpose of this article, empirical frequency data can be used to estimate the relevance that the case system has in argument marking.

4. The frequency of distinct case marking: Initial results

Analysing dialectal data as done in the present study yields several types of information. First, the assumptions on the spatial distribution of case marking patterns in certain word classes made based on the maps in Shrier (1965) can be examined more precisely.

Second, based on exact knowledge of the corpus structure, precise information can be gained on the ratio of occurrence of different semantic roles, syntactic functions and case markings found in each transcript as well as in the total corpus. Those findings are of particular interest, because they show us how central to the language system phenomena affecting these units are. In addition, it is possible to conduct statistical analyses on the association of case marking patterns and argument structure in the corpus. Furthermore, the spatial distribution of both corpus structures, which means the frequency of occurrence of certain arguments, canonical cases or word classes within the sample as well as argument marking patterns can be investigated.

To begin, a comparison of the data with Shrier's (1965) conclusion is drawn to typologically classify the present case marking system. Based on Shrier (1965: 437)'s map, one would assume the 3-case-distinction in the 1st and 3rd person singular pronoun⁹ as well as in the adjective in the area under investigation for the most part.

	total	syncretic case marking	distinct case marking	not applicable
adjectives	241	187	24	30
1 st sg. pronoun	641	2	637	2
masc. 3 rd sg.	213	5	204	4
pronoun				
masc. def.	485	31	445	9
article				

Table 3: Distinct and syncretic case marking in the contexts mentioned in Shrier (1965)

But, as can be seen from Table 3, those assumptions cannot completely be verified by the data. While the 1st and 3rd person masculine pronouns do indeed show distinct case marking forms, the adjective rather shows a tendency towards syncretism (even in masculine relations, where there are

⁹ This might be a small inaccuracy in Shrier's (1965: 437) article. Since parts of the case distinction in feminine and neuter relations have already been levelled for the most part, as previously outlined, the distinction can only be assumed in masculine items of the 3rd person singular pronoun.

only 18 distinct vs. 35 syncretic word forms). Additionally, the masculine definite articles do show a tendency towards distinct forms in the area as well, although a certain amount of variation seems to be possible here. These differences in classifying the area's case system might be partly based of a mistake in Shrier's (1965) illustration, since she describes the pattern found for the masculine definite article for the adjacent area in map 13 (Shrier 1965: 437) – which might indicate that she confused the areas – as well as for the area under investigation in map 2 (Shrier 1965: 424), where she concentrates on masculine definite articles. Nevertheless, this analysis can deal with the imprecision and refine the assumptions on the typology of the case marking system for this area.

To gain a deeper insight into the general structure of the corpus, Table 4 provides general information on the occurrence of arguments in the sample:

total	proto-agent	proto-patient	subject	object
5592	1794	2051	3393	827
	32%	36.7%	60.7%	14.8%

Table 4: Evidence for arguments carrying a case marker in the sample

The data suggests that a rather high amount (60.7%) of case exponents mark subject functions, while only 14.8% mark direct objects. This outcome is partly due to the fact that 54.4% of the clauses in the sample are intransitives and partly due to the situation that many objects are object clauses and therefore take the object function without carrying overt case marking as an inflected word form. Table 4 reflects the quantitative difference between the syntactic and semantic dimension of argument marking. While only 32% of markers express the proto-agent, 60.7% express the subject. The difference is based on the asymmetric relation where the syntactic functions correspond to more than one semantic role. The data therefore indicates that syncretic patterns affecting the subject interfere with the case system more profoundly than those affecting the proto-roles or even the direct object.

total	syncretic case marking	distinct case marking	not applicable
5592	2510	2872	149
	44.9%	51.4%	2.7%

Table 5: Frequency of distinct and syncretic case marking in the sample

Table 5 illustrates the quantity of total distinct and syncretic case marking across all gender, number and word class distinctions.¹⁰ The column “not applicable” contains those forms that cannot be classified as either syncretic or distinct (see section 3.2). As can be seen from the data, just over half of the case markers have a distinct form. Considering the canonical ideal of full distinction in case marking, in quantitative terms, the dialectal systems therefore only seem to approach this ideal half way.

¹⁰ In the present article, I focus on formal case marking, argument marking and animacy. Since especially word class, but also gender and number exert significant influence on the distinctiveness of German case marking, those factors are analysed within the broader scope of my thesis.

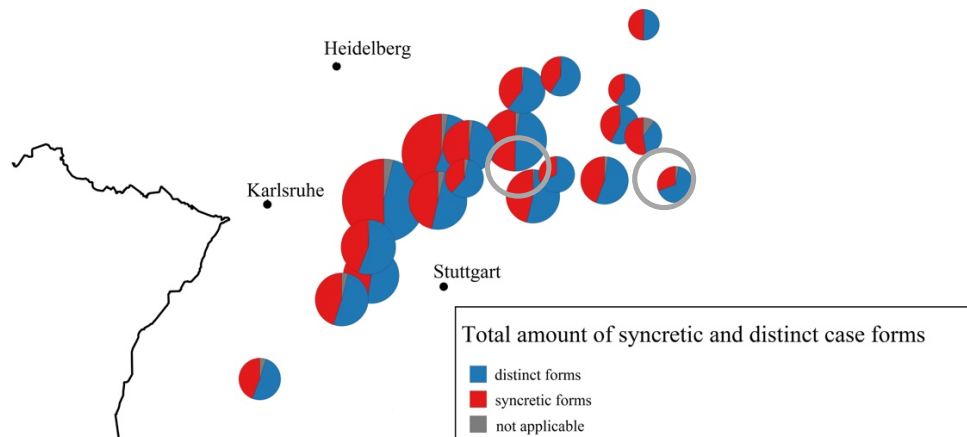


Figure 4: Spatial distribution of syncretic and distinct word forms in the corpus.

Looking at the spatial distribution of distinctiveness and syncretism in case marking in the sample (Figure 4), it is remarkable that, despite the differences in contexts, social data and number of tokens, there is a homogeneous distribution of syncretic and distinct forms over the area of investigation.¹¹ Hence, despite the inconsistency of the data discussed above, the corpus as well as the method prove to be practicable for the analysis. Figure 5 further suggests that a certain consistent minimal amount of distinctiveness might be required in the case systems examined. The amount seems to be fairly even, at least in the systems of the East Franconian dialects analysed here.

The result brings the concept of morphological minimum to mind that was first introduced by Rabanus (2008). It refers to a minimum level of distinct morphological marking needed to express categorical features while producing as little redundancy as possible (Rabanus 2008: 262). He developed this assumption on the base of a comparative study of intransitive clauses across High German dialects. Hence, the central idea of a minimal amount of distinction characteristic for a large dialect area has already been tested in German dialect morphology, but on the interaction of verb and subject pronoun, not on case. Below, I approach the issue whether the basic distinctions found in my sample are based on the need for distinct argument marking by implementing statistical analyses.

In favour of readability, I have chosen to show an overview of the results of the analyses in the tables below. The observed as well as the residual values are in the appendix only. In the following tables, + marks a positive association, i.e., those factors occur significantly often in the same contexts. The – sign marks a negative association, i.e., the respective factors co-occur significantly

¹¹ A first Pearson's Chi-squared-test on the frequency of syncretic and distinct forms in the particular locations confirms this assumption. The test shows no significant outlier, except from lower amounts of distinctions in two locations (residual values: Hengstfeld –2.48 and Häfnerhaslach –2,23) that seem to have no significant impact on the amount of syncretic case forms. There are only two locations (marked by grey circles), which differ from this pattern by showing a significantly higher amount of distinct forms (residual values: 2,58 in Öhringen and 2,49 in Michelfeld) as well as a significantly lower amount of syncretic forms (residual values: –2,23 in Öhringen and –2,62 in Michelfeld). Considering the social data of the speakers here, their comparatively low age is noticeable. Thus, this might be an indication for an emerging change in the system. However there are three additional speakers of the same age in the sample, that do not show this pattern. This issue deserves further investigation.

rarely in the same contexts. 0 marks no significant association and *n.a.* marks the factors that do not meet the frequency necessary for statistical analyses.

Initially, when looking at the tests, there is a remarkably high amount of data which shows significant results in the analysis. Besides permitting significant statements on the association of certain factors, as it is done below, this proves the suitability of both the corpus as well as the method used.

The statistical analyses of the data show that distinct case marking is significantly associated with proto-agent and the function of the subject.¹²

	proto-agent	proto-patient	subject	direct object
syncretic	-	+	-	+
distinct	+	-	+	-

Table 6: Syncretic and distinct case forms and argument marking

As can be seen from Table 6, arguments expressing proto-patient show a high association with syncretic case forms, while those expressing proto-agent show a high association with distinct case marking. Similar observations can be made looking at syntactic functions: The subject shows a high association with distinct forms, while the direct object is associated with syncretic forms. Thus, both levels of argument marking, the semantic as well as the syntactic one, show clear associations with case marking patterns. Therefore, although not every argument is marked by a distinct case form, case still seems to be strongly involved in marking argument structure in East Franconian dialects, despite its nearly 45% of syncretic patterns. At this stage proto-agents and subjects appear to show a positive association to distinct case forms while the opposite pattern is found with proto-patients and direct objects, as assumed in 2.1.

	syncretic	distinct
SELF	-	+
HUMAN	0	0
ANIMATE	+	-
INANIMATE	+	-
MASS	+	-
LOCATION	-	0
ABSTRACT	+	-

Table 7: Levels of animacy and syncretic or distinct argument marking

Table 7 relates levels on the animacy hierarchy (see Silverstein 1976) to the distinctiveness of case forms. As can be seen from the numerous levels, a more fine-grained animacy hierarchy was used here, in contrast to Alber & Rabanus (2011). The results below illustrate the advantage of this more fine grained analysis.

As can be seen from the residual values, non-HUMAN referents are associated with syncretic forms, while SELF-reference are associated with distinct forms. This association is not unexpected,

¹² To find out whether the findings show significant statistical associations, Pearson's Chi²-tests were used. In order not only to figure out whether there is a significant relation between the proto-role, syntactic function and case distinction (which is given in each constellation here, e.g., for syncretism vs. subject/object/agent/patient is: X-squared = 1159.6, df = 6, p-value < 2.2e-16), I do not focus on the p-values, but on the residual values generated for each cell in the cross table. The guidelines for significant deviations are <- 2 for a significantly low association and >+2 for a significantly high association.

since SELF is always expressed by personal pronouns that (as mentioned in Section 2.2) are assumed to maintain case distinctions longest. Interestingly, no association on either side can be seen on the HUMAN level. Regarding the observed values in Appendix 3, HUMAN referents show roughly similar percentages in syncretic as well as in distinct forms. Hence, based on this data, it can be assumed that the level of HUMAN rather than the distinction of ANIMATE versus INANIMATE might mark a turning point with regard to the influence on the distinctiveness of case marking. This also indicates interesting suggestions on the influence that person or speech act participants (see De Lancey 1998) might have on the distinctiveness of case marking. Unfortunately, the corpus does not allow to analyse a sufficient number of second person contexts to further investigate this issue, since it mostly contains monologues in which the only person being addressed is the interviewer, who is moreover spoken to using formal address forms.¹³ This result on the interaction of case marking and animacy leads to the question on whether the patterns are caused by animacy itself or whether they show an effect of pronominality. In fact, when investigating the association of case marking and animacy after factoring out both personal pronouns and indefinite pronouns, no association pattern can be found in the data at all. Therefore, at this stage of the analysis and in the system under investigation, the interaction of case marking and animacy appears to mainly occur on the level of pronouns.

Another interesting observation is the association on the level of LOCATION, which is a rather low level on the animacy hierarchy, with distinct case marking. This can be explained by the relevance of distinct dative marking in several Upper German dialects (see Dal 1971). As LOCATION is prototypically expressed by a dative (and further marked by adpositional case forms) this outcome is not surprising.¹⁴

As it was argued that syncretic patterns are associated with low levels of animacy, they are likely to mark proto-patient relations on a semantic level. It will be interesting to investigate how far the levels of the animacy hierarchy themselves influence argument marking.

	proto-agent	proto-patient	subject	direct object
SELF	+	-	+	-
HUMAN	+	-	+	-
ANIMATE	0	0	-	+
INANIMATE	-	+	-	+
MASS	n.a	n.a.	-	+
LOCATION	n.a	n.a.	n.a.	n.a.
ABSTRACT	-	+	-	+

Table 8: Levels of animacy and argument marking

As can be seen from Table 8, the levels of SELF and HUMAN show largely different associations compared to the lower levels of animacy. While the subject is in a positive association to those high levels, the object is in a negative association. Since no associations were found between the level of ANIMATE (- HUMAN) and the proto-roles, there are no obvious limitations on the semantic level of argument marking. Nevertheless, regarding the overall picture, animacy appears to have an

¹³ There are only 60 second person pronouns in the total sample. Only eleven of them are forms of informal *du* and six of them are forms of informal plural *ihr*.

¹⁴ Since LOCATION is often marked by adpositional dative constructions, this fact strongly affects the distinctiveness of locational referents. Adpositional markings are classified specifically in the data. Nevertheless, the case markers within the adpositional markings are themselves also classified in case and distinctiveness and integrated into the present analysis.

important function in argument marking in East Franconian dialects by distinguishing between humans and non-humans.

However, this analysis largely simplifies the relations. In fact, animacy and agentivity cannot be understood as independent variables, as it is presumed in Pearson's Chi-squared tests, since agentivity requires a certain amount of animacy. To further investigate the interaction of animacy and proto-roles within the limits of the data as well as to consider findings on the interaction of animacy and case marking in disambiguating proto-roles, syncretic and distinct case marking forms expressing proto-agent and proto-patient are also examined separately.

	proto-agent		proto-patient	
	syncretic	distinct	syncretic	distinct
SELF	–	+	–	+
HUMAN	+	–	–	+
ANIMATE	+	n.a.	0	n.a.
INANIMATE	+	–	+	–
MASS	+	–	0	0
LOCATION	+	–	0	n.a.
ABSTRACT	+	–	+	–

Table 9: Animacy and case marking in proto-agent and proto-patient relations

The results in proto-agent relations show a significantly low amount of syncretic case marking and a significantly high amount of distinct marking on the level of SELF, while the other levels show an opposite pattern. When looking at the proto-patient relations, the pattern found on the level of SELF also can be seen on the HUMAN level. Apart from the statistical simplification mentioned before, this allows for an interesting first conclusion. While the level of SELF generally tends to be marked distinctly, the level of HUMAN appears to show distinct case marking when it arises in the less prototypical proto-patient position. This gives initial indications on how animacy and case marking interact to distinctively express arguments.

5. Conclusion and outlook

As a fundamental conclusion, the method and the corpus provide useful generalizations on the emergence of East Franconian case marking patterns in different morphosyntactic contexts, suggested by the statistical significance of the results, as well as in the geographical dimension.

The comparison of the corpus data to the findings in Shrier (1965) proved to be useful, since it has refined the assumptions on the case marking system in the area under investigation and clear up some inconsistencies in Shrier's (1965) maps.

With regard to the development of case marking systems in German dialects, it can be noted that case, despite its tendency towards syncretism in numerous contexts, still performs an important role in expressing grammatical relations. Additionally, animacy, especially on the levels of SELF and HUMAN, is also involved in disambiguating argument structure, at least when it comes to interaction with case marking patterns in patient marking. The issue of the influence of animacy which seems to be reduced to pronominal contexts has only been given a brief account in this article. Still, it is an issue that should be thoroughly studied at a later point.

Further analytical steps can also be deduced from the initial findings: More advanced statistical tests on a larger data volume could allow a more detailed insight into the interaction of the factors under investigation and handle the complex interaction of the variables of animacy and proto-role.

In addition, several other factors are assumed to interact with case systems. In this article, I did not expand on word order, transitivity or word class differences. In order to scrutinize their influence on case marking, they should be included in a exhaustive analysis of a case system and will be integrated in further investigations.

As the present study gave only partial insights into the characteristics of German dialectal case systems, it is an interesting issue to tackle whether adjacent dialect areas show significant differences in the association of case distinction, agentivity and animacy or whether the basic amount of case distinctions across dialects is similar to those found in the East Franconian data – based on a common need to mark grammatical relations. Alemannic dialects, which presumably show a strong tendency towards syncretism of nominative and accusative forms in masculine items, will therefore be of particular interest. Since nominative-accusative syncretism heavily affects subject marking, it is an important issue whether this pattern occurs clearly in the data or whether other morphological, syntactic or semantic means prove to be more relevant here.

Abbreviations

ACC accusative, NOM nominative, PL plural, SG singular, DEF definite article, INDEF indefinite article, M masculine, F feminine, N neuter, PART particle, AUX auxiliary, PERS.PRON personal pronoun, INF infinitive

Appendix

Appendix 1: Observed values for syncretic and distinct argument marking

	proto-patient	proto-agent	object	subject
syncretic	1379 (ca.68%)	526 (ca.29%)	661 (ca.80%)	1332 (ca.39%)
distinct	567 (ca.28%)	1232 (ca.71%)	101 (ca.12%)	1985 (ca.59%)
total	2042	1793	823	3385

Appendix 2: Residual values for syncretic and distinct argument marking

	proto-patient	proto-agent	object	subject
syncretic	11.6961476	-12.3949585	13.7474801	-7.9654391
distinct	-12.4686050	13.2135681	-14.6128793	8.4154530
total	0.2964362	-0.3141474	0.5977665	-0.3106238

Appendix 3: Observed values for animacy and syncretic or distinct argument marking

	syncretic	distinct
SELF	94 (ca.4%)	1027 (ca.36%)
HUMAN	953 (ca.38%)	1099 (ca.28%)
ANIMATE	32 (ca.1%)	4 (ca. 0.1%)
INANIMATE	549 (ca.22%)	286 (ca.10%)
MASS	49 (ca.2%)	19 (ca.1%)
LOCATION	72 (ca.3%)	124 (ca.4%)
ABSTRACT	761 (ca.30%)	313 (ca.11%)
total	2510	2872

Appendix 4: Residual values for animacy and syncretic or distinct argument marking

	syncretic	distinct
SELF	-18.753703	17.5320086
HUMAN	-0.128978	0.1205758
ANIMATE	3.712214	-3.4703849
INANIMATE	8.086759	-7.5599534
MASS	3.069709	-2.8697353
LOCATION	-2.030003	1.89776
ABSTRACT	11.622656	-10.8655082

Appendix 5: Observed values for animacy and argument marking

	proto-agent	proto-patient	subject	object
SELF ¹⁵	525	216	956	34
HUMAN	1142	500	1761	137
ANIMATE	11	21	10	17
INANIMATE	91	5114	245	284
MASS	n.a.	n.a.	26	37
ABSTRACT	25	724	390	313

Appendix 6: Residual values for animacy and argument marking

	proto-agent	proto-patient	subject	object
SELF	9.174016	-8.743537	5.643637	-11.457633
HUMAN	12.892371	-12.287413	5.976724	-12.133859
ANIMATE	-1.084262	1.033384	-2.516062	5.108073
INANIMATE	-11.607287	11.06263	-8.758527	17.781437
MASS	n.a.	n.a.	-3.468834	7.042377
ABSTRACT	-17.557572	16.733706	-7.388589	15.000208

Appendix 7: Observed values for animacy and case marking in proto-agent and proto-patient relations

	proto-agent		proto-patient	
	syncretic	distinct	syncretic	distinct
SELF	7	518	29	187
HUMAN	450	660	283	178
ANIMATE	8	n.a.	19	n.a.
INANIMATE	45	44	406	91
MASS	n.a.	n.a.	44	16
LOCATION	n.a.	n.a.	6	n.a.
ABSTRACT	16	7	592	93
total	526	1232	1379	567

¹⁵ Unfortunately, not each level of the animacy hierarchy has shown sufficient results for each cell. Since Pearson's Chi-squared-test cannot manage datasets lower than 5, the levels MASS and LOCATION had to be taken out of the analysis.

Appendix 8: Residual values for animacy and case marking in proto-agent and proto-patient relations

	proto-agent		proto-patient	
	syncretic	distinct	syncretic	distinct
SELF	-11.974717	7.824431	29	187
HUMAN	6.468592	-4.22667	283	178
ANIMATE	2.595534	n.a.	19	n.a.
INANIMATE	3.560012	-2.32616	406	91
MASS	n.a.	n.a.	44	16
LOCATION	n.a.	n.a.	6	n.a.
ABSTRACT	3.560012	-2.27120	592	93

References

- Alber, Birgit & Stefan Rabanus. 2011. Kasussynkretismus und Belebtheit in germanischen Pronominalparadigmen. *Dynamik des Dialekts – Wandel und Variation*, ed. by Elvira Glaser, Natascha Frey & Jürgen Erich Schmidt, 23–46. *Zeitschrift für Dialektologie und Linguistik Beihefte 144/3*. Stuttgart: Steiner.
- Baerman, Matthew, Dunstan Brown & Greville Corbett. 2005. The Syntax-Morphology Interface: A Study of Syncretism. *Cambridge Studies in Linguistics 109*. Cambridge: Cambridge University Press.
- Bittner, Dagmar. 2002. Semantisches in der pronominalen Flexion des Deutschen. *Zeitschrift für Sprachwissenschaft 21/2*. 196–233.
- Blake, Barry. ²2001. Case. Second edition. *Cambridge Textbooks in Linguistics 227*. Cambridge: Cambridge University Press.
- Braune, Wilhelm & Ingo Reiffenstein. ¹⁵2004. Althochdeutsche Grammatik 1: Laut- und Formenlehre. *Sammlung kurzer Grammatiken Germanischer Dialekte 5/1*. Tübingen: Niemeyer.
- Brown, Dunstan & Marina Chumakina. 2013. What there might be and what there is: an introduction to Canonical Typology. *Canonical Morphology and Syntax*, ed. by Dunstan Brown, Marina Chumakina & Greville G. Corbett, 1–19. Oxford: Oxford University Press.
- Dal Negro, Silvia. 2004. Artikelmorphologie: Walserdeutsch im Vergleich zu anderen alemannischen Dialekten. *Alemannisch im Sprachvergleich*, ed. by Elvira Glaser, Peter Ott & Rudolf Schwarzenbach, 101–111. *Zeitschrift für Dialektologie und Linguistik Beihefte 129*. Wiesbaden: Steiner.
- Dal, Ingerid. 1971. Systemerhaltende Tendenzen in hochdeutschen Mundarten. *Untersuchungen zur germanischen und deutschen Sprachgeschichte*, ed. by Ingerid Dal, 171–180. Oslo/Bergen/Tromsø: Universitetsforlaget.
- De Lancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language 57/3*. 626–57.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language 67/3*. 547–619.
- Dürscheid, Christa. 1999. Die verbalen Kasus des Deutschen: Untersuchungen zur Syntax, Semantik und Perspektive. *Studia Linguistica Germanica 53*. Berlin: De Gruyter.
- Ebert, Robert Peter, Oskar Reichmann, Hans-Joachim Solms & Klaus-Peter Wegera. 1993. Frühneuhochdeutsche Grammatik. *Sammlung kurzer Grammatiken Germanischer Dialekte 12*. Tübingen: Niemeyer.

- Howe, Stephen. 1996. The Personal Pronouns in the Germanic Languages: A study of personal pronoun morphology and change in the Germanic languages from the first records to the present day. *Studia Linguistica Germanica* 43. Berlin: De Gruyter.
- Krifka, Manfred. 2009. Case syncretism in German feminines: Typological, functional and structural aspects. *On Inflection*, ed. by Patrick Steinkrüger & Manfred Krifka, 141–172. *Trends in Linguistics. Studies and Monographs* 184. Berlin: De Gruyter.
- Primus, Beatrice. 1999. Cases and Thematic Roles: Ergative, Accusative and Active. *Linguistische Arbeiten* 393. Tübingen: Niemeyer.
- Paul, Hermann. 1959. Deutsche Grammatik: Zweiter Band: Teil 3: Flexionslehre. Halle (Salle): Niemeyer.
- Rabanus, Stefan. 2008. Morphologisches Minimum: Distinktionen und Synkretismen im Minimalsatz hochdeutscher Dialekte. *Zeitschrift für Dialektologie und Linguistik Beihefte* 134. Stuttgart: Steiner.
- Rohdenburg, Günter. 1998. Zur Umfunktionierung von Kasusoppositionen für referentielle Unterscheidungen bei Pronomen und Substantiven im Nordniederdeutschen. *Zeitschrift für Dialektologie und Linguistik* 65. 293–300.
- Ruoff, Arno. 1973. Grundlagen und Methoden der Untersuchung gesprochener Sprache: Einführung in die Reihe „Idiomatologica“. *Idiomatologica* 1. Tübingen: Niemeyer.
- Ruoff, Arno. 1984. Alltagstexte I: Transkriptionen von Tonbandaufnahmen aus Baden-Württemberg und Bayrisch-Schwaben. *Idiomatologica* 10. Tübingen: Niemeyer.
- Shrier, Martha. 1965. Case Systems in German Dialects. *Language* 41/3. 420–438.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. *Journal of Linguistics* 15. 112–171.
- Stump, Gregory. 2015. Inflectional Paradigms: Content and Form at the Syntax-Morphology Interface. *Cambridge Studies in Linguistics* 149. Cambridge: Cambridge University Press.
- Werlen, Iwar. 1990. Kasus und Wortstellung in alemannischen Dialekten. *Alemannische Dialektologie im Computer-Zeitalter*, ed. by Marthe Philipp, 165–190. *Göppinger Arbeiten zur Germanistik* 535. Göppingen: Kümmerle.
- Wiesinger, Peter. 1983. Die Einteilung der deutschen Dialekte. *Dialektologie: Ein Handbuch zur deutschen und allgemeinen Dialektforschung* 2/1, ed. by Werner Besch, Ulrich Knoop & Herbert Ernst Wiegand, 807–900. Berlin: De Gruyter.