

Defining lexeme types in German

The concept of word type (or part of speech) plays an important role when describing German grammar. Significantly different models have been proposed in this context, assuming between 5 and 51 word types. These models also differ fundamentally in whether they distinguish between lexical words (lexemes) and grammatical words (also simply called words) or whether they negate this central morphological distinction and make use of an ambiguous, mixed concept. Some models only use distinction criteria from a single linguistic subfield, focusing primarily on morphological criteria while others use a mixture of morphological, syntactic, semantic, and possibly other criteria. Traditional grammar (including school grammar) falls into the category of a mixed model while more recent theoretical approaches often, but by no means always, take a more stringent approach. The goal of word type classifications is, in any case, to contribute to an explanation of the grammatical regularities of a language.

The present study is situated in the paradigm of Linguistic Realism and thus conceives of language as an abstract object to be reconstructed as a system by theoretical means. Other linguistic questions concern the knowledge that language users may possess about such a language system and the use that language users may make of this knowledge about a language system. The study takes a language-specific approach in the sense that contemporary German grammar is modeled as an autonomous object without recourse to the grammatical systems of other languages or historical forms of German. Nevertheless, the theory is designed in such a way that it is, in principle, transferable to other languages. Finally, the investigation takes the distinction between lexemes and words as independent units seriously, assuming that the classification systems of both types of units are necessary to model a grammar. The focus of the present study is on the classification of lexemes, with the classification of words only being discussed in a tentative manner. Only the relevant inflectional categories are used to define the lexeme types of German so that the model can be described as criterion-pure. In this way, five lexeme types are distinguished, as in Glinz (1952), which are, nevertheless, subdivided in a somewhat different way and partly have different definitional features compared to Glinz' approach.

Nominal lexemes are defined as inflecting for case and number. An inflectional category of a language system is called 'case' when, for example, the grammatical functions of subject and object are marked by different subunits of this category. In German, there are four cases. An inflectional category of a language system is called 'number' when denoting a single unit is typically marked with a different form than denoting more than one unit. In German, there are two subcategories of number, singular and plural. Nominal lexemes are therefore usually assigned eight words, with the nominative singular word acting as the citation form of the lexeme. The words are realized by word forms in the sense of phonological words. In nominal paradigms, syncretism can be observed, which means that a particular word form is used to realize more than one word of a lexeme. For example, the word form [ha^os] is used in the inflectional paradigm of the lexeme HAUS 'house' for the words nominative singular, accusative singular, and dative singular. To a lesser extent, there is variation in German, meaning that one word of a lexeme can be realized by more than one word form. For example, the dative singular of the lexeme HAUS

can also be expressed by the word form [ha^o.zə]. Importantly, inflection is not understood here, as would be the case in many other approaches, as a change of form relative to an assumed base form but as the filling of cells of an inflectional paradigm. In this sense, a lexeme such as ELTERN ‘parents’ can be considered a nominal lexeme, although it uses only one word form in its entire paradigm. Moreover, this lexeme, exceptionally, only has the sub-paradigm plural, which indicates that it is an irregular nominal lexeme.

Verbal lexemes are defined as inflecting for number, person, tense, and mode. Here, only synthetic forms are considered as inflection. This results, in particular, in a hybrid position for tense: while present tense and past tense are formed synthetically, all other tenses use additional auxiliary verbs in analytic form. This can be modeled in such a way that tense formation is basically regarded as a form type of its own. In the inflectional paradigm of a verb, only the sub-paradigms labelled T1 and T2 are included where, for example, the present tense is formed in such a way that simply the T1 form is used, while in the perfect tense an auxiliary verb occurs in the T1 form with the full verb in the participle II form. Furthermore, it follows from this conception that the non-finite verb forms infinitive, participle I, and participle II do not fall within the inflectional paradigm but are forms of their own kind. In total, a verbal inflectional paradigm includes 26 different words. The formation of the word forms of most verbal lexemes can be described as regular; only some 170 simple verbs behave irregularly in this sense.

Adjectival lexemes are often defined as inflecting for case, number, gender, and comparison. However, since many adjectives are non-comparable (e.g., adjectival compounds such as HELLBLAU ‘bright blue’), comparison is not suitable as a definitional criterion but can rather be viewed as a kind of lexeme formation. The remaining inflectional categories of case, number, and gender are not, however, sufficient to adequately describe the inflectional behavior of adjectives. In the phrases *der lange Rock* ‘the long skirt’ versus *ein langer Rock* ‘a long skirt’, the features for the above categories are identical, but the adjective, nevertheless, occurs in different word forms. Traditionally it is noted in this regard that adjectives have, in addition to inflection proper, two ‘inflectional rows’ called *stark* ‘strong’ and *schwach* ‘weak’. This property can be reconstructed as an original inflectional category, for which the generic term ‘stach’ is proposed (as a blend of *stark* and *schwach*). Thus, adjectival lexemes are defined as inflecting for case, number, gender, and stach, resulting in an inflectional paradigm of 48 words. Among them, in the plural sub-paradigm, there is complete syncretism between the forms for the three genders that are generally assumed. Adjectives ending with an unstressed full vowel (e.g., ROSA ‘pink’), adjectives of origin such as KÖLNER ‘from Cologne’, and cardinal number adjectives such as VIER ‘four’ inflect irregularly, mostly in that they show complete syncretism in their paradigms.

Determiner lexemes are defined as inflecting for case, number, and gender. Determiners include the lexemes DEFINITE ARTICLE and INDEFINITE ARTICLE as well as many lexemes traditionally classified as pronouns such as MEIN ‘my’ and DIESER ‘this’. Traditional pronouns such as personal pronouns are then classified as nominal lexemes. The number of determiner lexemes is small compared to the first three types of lexemes. At the same time, many of these lexemes behave irregularly in their inflectional properties.

The fifth lexeme type in German is characterized by not being subject to any inflectional category. It can, therefore, be meaningfully called *Unflektierbar* ‘uninflectable’. This lexeme type includes units that are traditionally classified as prepositions, conjunctions, or adverbs, among others. Even though these units all carry the same lexeme type, their grammatical differences can still be captured via a classification, namely at the level of the unit word. In such a word type classification, nominal words can be defined as being inflected for case and number, with corresponding definitions applying to verbal, adjectival, and determiner words. Such a conception then requires that the base form of adjectival lexemes has a word type other than adjective, since it is not inflected (and consequently cannot be used attributively). Likewise, the non-finite forms of verbal lexemes must have a word type other than verb. There are different proposals in the literature on this. Some such proposals assume that words can carry more than one word type at the same time. This is a reasonable approach, also with regard to contraction forms such as *im*, which can be analyzed as carrying the word types preposition and determiner at the same time; the word refers simultaneously to the inflectional lexeme IN ‘in’ and the determiner lexeme DEFINITE ARTICLE. Finally, within the group of uninflectable lexemes, appropriate syntactic features can be used to distinguish the word types interjection, particle, and adverb, as well as the word type ‘Fügewort’ ‘linking word’, which includes prepositions, conjunctions, and the lexemes WIE ‘as, like’ and ALS ‘than’, and which is defined by the corresponding words being integrable into syntactic structures, being syntactically expandable, and governing another syntactic element.

Reference

Glinz, Hans (1952): *Die innere Form des Deutschen*. (= Bibliotheca Germanica 4). Bern/München: Francke.