

Leila Behrens

Categorizing Between Lexicon and Grammar The MASS/COUNT Distinction In a Cross-Linguistic Perspective

In dieser Arbeit werden Zusammenhänge zwischen lexikalischer und grammatischer Kategorisierung untersucht. Im ersten Teil werden generelle methodologische Probleme behandelt, die sich durch Kategorien, die schwer eindeutig nur dem Lexikon oder nur der Grammatik zuzuordnen sind ('MASS/COUNT'-Distinktion, ASPEKT/AKTIONSART, Wortklassen), für die Modellierung der Lexikon-Grammatik-Interaktion ergeben. Im zweiten Teil wird ein kontrastiver Vergleich des Kategoriepotentials dreier Sprachen (Englisch, Deutsch, Ungarisch) in der Domäne 'MASS/COUNT'-Distinktion (Quantifikation/Determination) vorgenommen. Dabei konnten signifikante Kategorisierungsunterschiede zwischen diesen Sprachen festgestellt werden, die sich aus der jeweiligen Interaktion von lexikalischen Alternationsmustern und grammatischen (Sub-)paradigmen ergaben. Als entscheidend für den zwischensprachlichen Vergleich erwies sich, daß die Faktoren, die die 'prototypische' MASS/COUNT-Distinktion (die englische), konstituieren, auf mehreren Dimensionen untersucht wurden.

1 Introduction

Lexicon research in the past decade has led to a new way of linguistic thinking. Researchers working in different areas and applications, i.e., theoretical linguists, psycholinguists, lexicologists, lexicographers, cognitive scientists, philosophers, logicians, computational linguists, etc., working on a single language or on cross-linguistic issues (typology, machine translation, multilingual databases), have developed a number of convergent ideas which force us to reevaluate some of the basic linguistic concepts of the last 50 years.

One of the problems commonly addressed by all these people is the identification and representation of categories which cannot be exclusively assigned to either the lexicon or to the grammar, and which result from regularities on both sides of a unified lexico-grammar and from the language-specific interactions of these two components. In a multilingual context, a further problem arises. This concerns the cross-linguistic identification of corresponding categories in a specific domain (for example, in the domain of aspect/causativation or in the domain of quantification/determination). Such categories can be differently distributed between lexicon and grammar in different languages and can have different semantic and pragmatic implications as well as different grammatical implications.

The MASS/COUNT distinction is perhaps the most famous example for this problem. Since the time of Jespersen (cf. 1925), there is an ongoing discussion of the following:

Is the MASS/COUNT distinction to be understood as a lexical category or as a syntactic category¹, or, even as a non-categorial pragmatic distinction? Is the MASS/COUNT distinction primarily a distributionally-defined category with prototypical semantic (ontological or perceptual/conceptual) correlates or is it a semantic category with prototypical distributions? How shall we deal with 'exceptions', that is, with examples having the 'wrong semantics' or having the 'wrong syntax'? How shall we generally address the so-called 'dual nouns' oscillating between MASS and COUNT syntax, as well as between MASS and COUNT semantics? And how do we cope with the diversity of the dual nouns itself?

Those who advocate the syntactic category solution are frequently criticized on the following grounds: Lexical elements have an idiosyncratic bias to either MASS or COUNT type, and the effects of mutual conversion are not always the same. In some cases, highly marked effects appear, often exploited in advertising; in other cases, the use of the 'converted' noun in its 'non-inherent' context is as marked (or unmarked) as the use of the 'non-converted' noun in its 'inherent' context. Those advocates of the lexical category solution are generally criticized supported by examples showing the productivity of dual uses, which would lead to a double categorization of the entire lexicon. A last question is whether or not the MASS/COUNT distinction is a universal distinction in the sense of a substantive universal necessarily occurring in every language. Do, for example, classifier languages like Vietnamese have such a distinction, or do they have only MASS nouns or a different type of distinction in the quantification/determination domain and not to be equated with the English MASS/COUNT distinction at all?

Conceptual, methodological and representational difficulties quite similar to those found in the domain of MASS/COUNT distinction also arise in some other linguistic domains. In this paper, I have decided to examine MASS/COUNT distinction as a model case for two reasons:

- (a) On the one hand, the MASS/COUNT problem is well-known and represents a central topic in very different research contexts (see above).
- (b) On the other hand, nominal categories are, in some respect, not as well investigated as verbal categories and subcategories.

¹ 'Syntactic category' is to be understood here as a (grammatical) category established on the level of syntax, that is, only on the sentence level, rather than being defined by syntactic, i.e., distributional means. A fine-grained classification of different approaches will be given in 3.1.1.

The study of lexical subregularities, i.e., systematic sense relations² which determine the uses of a single lexical form in different sets of contexts in a certain part of the vocabulary, have a longer tradition in the verbal domain than in the nominal domain. There are more detailed descriptions available, both regarding the range of data and regarding the theoretical implications of productive and semi-productive lexical patterns for the cross-classification of verbs and their internal and external structure as lexical entries. In particular, a great deal of attention has been given to such recurrent contrasts or alternations which show a correlation between different senses and different syntactic contexts (syntacto-semantic alternations) (cf. Levin 1993; Pinker 1989; Kilgariff 1993; Basili et al. 1992). This is not to say that systematic metonymies and other sorts of productive and systematic patterns which (potentially) characterize the lexicon of single languages and are often called semantic (lexical) extensions or shifts³ have been neglected for nouns. On the contrary: In cognitive linguistics (of all persuasions), and in other cognition-oriented approaches (cf. Franks/Braisby 1991; Franks 1991; Wilensky 1990), nominal and verbal sense extensions are of presumably equal importance. Here, I simply mean that the interactions between lexicon and grammar are better understood in the case of verbal alternations. This also holds true in a cross-linguistic perspective. Concerning this topic, Talmy's (1985) typological studies are pioneering works and could be fruitfully applied in the nominal domain.

In this paper, I would like to direct the reader's attention to the ubiquity of ambiguity. It is commonly agreed today that lexical forms may show different types of ambiguity. Strong ambiguity/homonymy, weak ambiguity/lexicalized polysemy, systematic ambiguity/polysemy, generality and vagueness are generally distinguished by means of two kinds of tests: The first is based on lexical combinatorics (identity test, etc.), the second on paradigmatic behavior (synonymy, etc.). There is, I believe, a consensus among linguists that grammatical units can also be ambiguous. There is, in fact, no reason to believe that 'function words' and affixes would behave in a way radically different from 'content words' (i.e., showing no ambiguity), if one takes the transitional character of these two domains into account.

Of course, one can commit to a strong version of the one-form-one-meaning principle (biuniqueness). But in this case, homonyms among 'content words' should also be treated,

² What I mean by this term is systematic relations between 'word senses' which are connected with a single lexical form rather than sense relations connected with different lexical forms, i.e. synonymy, antonymy etc.

³ 'Semantic shift' and 'semantic extension' are too strong of terms for the area I would like to cover since both presuppose the existence of a primary sense and a particular direction among the related senses. It may be sometimes appropriate to determine one of the senses as a primary sense, sometimes not. This is exactly the issue which will be discussed later. In the present context, it is sufficient to use these well-known expressions.

firstly, as forms with a single meaning. The indefinite article, in English and in a variety of other languages, is a classical candidate for ambiguity among 'function words' (cf. Gillon 1992: 600). However, this would pose a serious problem for the distributional techniques generally employed for the identification of categories both in the field of lexicon and grammar. That is, for successful application of distributional techniques, it is necessary to take this fact into account. It is precisely this point which was deliberately ignored in official versions of structuralist theory, and, it is probably the legacy of structuralism that this fact is sometimes still ignored. There is a long tradition, reinforced from time to time by new impulses, to not accept lexical ambiguity solely on local distributional grounds. Many linguists would find it very reasonable to say, for example, that an element is in no way ambiguous as long as one can find complementary contexts for different utterance meanings. But context does not always disambiguate: There are neutral contexts and contexts which are themselves ambiguous. The situation is, I believe, much more complicated. I would like to argue that it is a sort of (language-specific) optimization based on syntagmatic and (!) paradigmatic regularities that ultimately determines the distribution of ambiguity/indeterminacy between lexicon and grammar.

In the first part of this paper, I will focus on the historical and intellectual sources of the MASS/COUNT problem. In the last 50 years, mainstream linguistics research was characterized by certain lexicon concepts, namely (a) the concept of lexicon as a (finite) list of irregularities, (b) the concept of the priority of forms for determining lexical categories, and (c) the (later) concept of lexical insertion. I claim that these concepts are conceptually not wholly compatible and, taken together, constitute a background in which it is very difficult to find adequate solutions for the problems mentioned above. In the course of time, linguists developed some standard strategies for coping with problems such as the concept of 'mixed categories' or the concept of 'lexical underspecification'. I would like to address the question as to whether one of these strategies is superior to all others in providing an optimal solution for all languages or whether the optimality of a strategy depends on language-specific factors.

In the second part of this paper, I will present data from three languages. The first one is English, because it has largely determined our view of the MASS/COUNT problem. The situation in English is confronted with that in German and Hungarian. For these two languages, analogous investigations were made in selected semantic fields and with respect to comparable phenomena (i.e., 'multi-words', etc.). Hungarian is of particular interest because the applicability of standard criteria (use of indefinite article, use of plural markers) is problematic here. Depending on the raw estimation of the MASS/COUNT distinction as a semantic category or a syntactic category, one can draw different conclusions regarding the question of whether or not Hungarian has such a distinction. I will argue that only a multi-dimensional approach will be appropriate for all three languages.

One might, of course, ask why we should bother (a) about different views on the MASS/COUNT issue, depending on different views on lexicon, syntax, and different research objects (i.e., logical representation of utterances, mental lexicon, relativity hypothesis, lexicon module in a machine translation system, etc.), and (b) about different categorizations for single languages. The popular argument about the dependency of linguistic description and categorization on a given task is indeed correct. Psychological reality is, for instance, not a necessary requirement for computational systems to work optimally and psycholinguistic results on the mental lexicon and computational representations need not to be congruent. A formal semantician need not care for every idiosyncratic detail of an entire vocabulary in the way a lexicographer would and therefore, it is natural that the two take a different view on categorization.

In practice, however, researchers adopt a lot of assumptions from neighboring disciplines, which do not always correspond very well to the requirements of the goal discipline. While testing (controversial) linguistic hypotheses, psycholinguists sometimes rely on linguistic assumptions which are only valid under the simplifying conditions of a certain linguistic framework. Computational linguists often adopt linguistic tests developed for the description of selected data which are not, however, easily applied in classifying an entire vocabulary under the requirement of predictability. Some researchers (cf. Sampson 1992: 437) consider the inconsistent and vague use of linguistic terms as the most difficult problem in the natural language processing (NLP) discipline and their clarification and standardization as the most urgent need. For years, several groups have been working on 'neutral' lexical databases which are reusable for research groups with different tasks (cf. Boguraev/Briscoe 1989) and on linguistic standardization of corpora and lexical databases (cf. EAGLES). Theoretical linguists might be skeptical about the possibility of standardization. However, times are changing. The very same skeptical linguists are more and more enthusiastic users of huge corpora and lexical databases which are freely available and also annotated as data resource. In my opinion, linguistic annotations, that is, category tags, are entirely useless unless precisely defined and translatable into any given model. Imagine the following situation. Somebody is working on ASPECT and wants to examine the interaction between noun/noun phrase type and the aspectual type of verbs/verb phrases as (a) lexically determined feature and as (b) syntactically generated features in a corpus.⁴ Let us assume that there is a corpus annotated not only with word classes (noun, verb, etc.), but also with subcategories having values such as MASS, COUNT and perhaps MASS-COUNT or NEUTRAL. If the person in question does not know the exact coding principles valid for the entire corpus, these values are useless

⁴ 'Verb' is to be understood here as an uninflected unit of the lexicon, 'verb phrase' as a 'phrasal' or 'constructional' unit of the sentence. That is, 'verb phrase' does not mean a special configuration defined with respect to the distinction between complements and adjuncts.

to him. This kind of tagging problem has been widely discussed in connection with word classes. Even if the MASS/COUNT subcategory is not yet commonly used for corpus annotation, it already provides difficulties in coding multilingual lexical databases (cf. EAGLES).

My aim is here certainly not to make a proposal for standardization in the domain of the MASS/COUNT distinction. Furthermore, I am very skeptical as to whether a linguistic taxonomy in the sense of Sampson (cf. above) would ever work. I see the shared concern which is unifying the members of the interdisciplinary community working on language in developing a 'heuristically adequate' general approach (cf. Wilks 1992). This could help in determining which information types have to be conceptually distinguished, and finding an answer to the following question: Which distinct terms hide very similar or identical concepts, and which terms are used ambiguously? This conceptual task is superordinate to any question of technical representation. In this paper, I will be concerned with questions of a cross-linguistically adequate conceptual design.

2 Categorization Between Lexicon and Grammar

2.1 Complex Categories and the Prototype Approach

It is well known that the most common category symbols used in lexical representations are labels for complex linguistic categories. By complex categories, we mean holistically interpreted units of grammar, i.e., 'grammatical primitives', which are defined on the basis of more than one classificatory dimension (morphology, syntax, semantics), and result in a set of complex properties. Both the major lexical categories and the nominal subcategory MASS/COUNT can be seen - despite apparent differences - as good examples of complex categories (cf. Lyons 1977: Chapter 11.3; Sasse 1993). This phenomenon has often been described as the 'prototypicality of grammatical categories' (cf. Hopper/Thompson 1984; Thompson 1989; Corrigan 1989; Zelinsky-Wibbelt 1989).

The concept of prototypes was founded in psychological research, and there is some general doubt about the legitimacy of its transfer to the lexico-grammatical domain (cf. Tsohatzidis 1990: 8). In my opinion, this is certainly not adequate when employed in the study of universal grammatical categories. The basic idea behind the prototype concept in psychology was that this kind of categorization is more characteristic of the human being than the classical (i.e., Aristotelian) kind. Insofar as categorization is a fundamental cognitive ability and language use is the manifestation of cognitive processes, it is, in principle, correct to claim that linguistic categories in individual languages are organized according to the prototype principle. However, there is simply no such human being, except a linguist, who

would cognitively categorize universal grammatical categories, either as the sum of all language-specific categories or as language-independent abstract categories. 'Prototype' is used in cross-linguistic studies as a metaphor, which is grounded - among other things - on the similarity between difficulties in psychology and linguistics in defining hypothetical categories in terms of necessary and sufficient criteria. However, the criteria for identifying more or less traditional grammatical categories in different languages of the world may simply be mistaken. It is also possible that it is the prototype concept of 'being a good member' (cf. Geeraerts 1989 for different prototype concepts) which is responsible for the transfer of the prototype notion to universal grammatical categories. In this case, one could suppose that the categories (or category configurations) of a specific language or language type represent the 'good members' of the respective universal categories and the categories of other languages the 'poor members'.

However, the concept of prototypicality is also widely used in the description of individual languages. This is the only use of the term which can be seriously considered as in accordance with the original insights of psychological research. It seems evident that the postulation of prototypicality of language-specific and universal categories address entirely different questions.⁵ Language-specific categories can be, for example, fairly discretely organized, whereas universal categories probably never have this property. We are often told that (language-specific) grammatical categories seldom match perfectly with language-independently/philosophically stated ontological categories which constitute their core denotational domain (cf. Lyons 1977). However, the prototype effects caused only by such mismatches do not really bother linguists and rarely lead them to call the legitimacy of the respective categories into question.

If it is said that grammatical categories are simultaneously defined in morphological/syntactic and semantic terms, we certainly cannot conclude that the criteria are randomly used in an unordered fashion (i.e., that sometimes 'formal' criteria, sometimes semantic criteria, are first applied). Presumably all linguists, including those insisting - generally or particularly - on the semantic basis of grammatical categories, give priority to morphological/syntactic criteria (as sufficient) over semantic ones in determining language-specific categories, at least in well-studied (European, etc.) languages, and only later try to assign these categories a semantic or ontological interpretation. Rather, problems of categorization arise, due to the fact that the 'formal' criteria themselves (i.e., the morphological and syntactic criteria) do not constitute necessary and sufficient conditions. The reason for this may lie in the traditional roots of some categories and in the fact - like in the case of MASS/COUNT distinction - that the

⁵ For this argumentation one must, of course, accept that language-specific and universal categories are not identical.

respective categories were first developed in a more restricted semantic/ontological domain and then generalized for other domains. Nevertheless, the resulting 'fuzzy boundary'-effects relevant in language-specific description are to be treated separately from the issue of 'good members' as a subject matter of research on language universals (cf. above).

Let me demonstrate this with some examples. Every linguist would agree that the German nouns *Zerstörung* ('destruction'), *Traum* ('dream') or *Sensibilität* ('sensitivity'), all of them frequently characterized as 'non-prototypical' nouns (cf. Miller 1985, on problems of abstractness and Lyon's classification into first-order, second-order, third-order entities), are, without a doubt, members of the class 'noun'. Likewise, *furniture*, one of the most popular examples for a 'non-prototypical MASS noun' in English, is always without hesitation assigned to the category MASS as long as a binary distinction is assumed, whereas other non-prototypical examples are sometimes estimated as (real) MASS, sometimes as (real) COUNT, sometimes as both, or neither of the two (*weather, fire*, etc.; cf. Allan 1980; Bunt 1985: 15). Similarly, there is a considerable lack of agreement among lexicographers and linguists concerning the categorial status of the German lexical elements *immer* ('always'), *auch* ('also', 'as well'), *auf* ('up', 'open', 'on'), *entzwei* ('in two', 'broken'), *Schweizer* ('Swiss'), *viel* ('much'/'many') (!), corresponding to the especially bad formal distinction of the proposed categories (adjective, adverb, preposition, noun (in opposition to adjective), indefinite pronoun (in opposition to adjective)) (See Bergenholtz 1983, for a good survey of word class problems in German.)

2.2 Paradigms and Lexical Sensitivity

It is useful to establish two logically independent parameters: 'paradigmatic uniformity' and 'lexical sensitivity' for grammatical categories (syntactic functions).

2.2.1 Paradigmatic Uniformity

Paradigm is not a uniform notion. It is systematically used in two overlapping senses: First, as a "systematization of morphological inventories" confined to inflection in a narrow sense and second, as "paradigmatic patterning of grammatical meaning" (cf. Hendrikse 1993: 480ff.). The difference between the two concerns the definition of inflection and on what level of abstractness it is to be interpreted. In a narrow sense, only those formal means (affixes, reduplication, tonal modification) count as inflection, as an endpoint of grammaticalization (cf. Bybee/Pagliuca 1985), which would apply in the scope of the 'phonological word'. The tradition of grammar book writing, by comparison, is somewhat indifferent on this point. Some periphrastic constructions, like those associated with FUTURE, are more commonly

represented as part of the whole verb paradigm, other constructions, consisting of a lexical form and a 'function word' or clitics, like the object clitics in Modern Greek, less commonly (cf. Matthews 1974: 172). The restriction of paradigm to a specific type of grammatical means or to different types operating on a specific part of the lexical form (word initially/word finally) is not very useful in cases where the same grammatical category is distributed over different formal types (cf. the PERSON marking by means of prefixes and suffixes in Kanuri). More generally, it has negative consequences for capturing significant interactions between grammatical categories of a single language (say, between determiners and CASE in German, if article paradigm and CASE paradigm for nouns are presented separately, without explicitly stating that the article paradigm, as such, is a crucial part of the entire noun paradigm) and for capturing typological generalizations.

Thus, a more general paradigmatic notion has to be preferred, which includes the whole set of grammatical means (i.e., 'function words', affixes, tonal modification, etc., and also, such 'formal means of expression' like external agreement) and their configuration, which can contribute to the grammatical value and the sentential interpretation of a type of lexical unit (noun or verb, for example) used in the sentence as a (minimal) phrasal unit.⁶ I will call this extended concept of paradigm 'macroparadigm', indicating that the whole set of grammatical means is allowed to have subparadigms which are not necessarily strongly tied together. I proceed from the assumption that evolution into a unitary macroparadigm, without, or with less, syncretism, is a sort of grammaticalization. Thus, grammaticalization is not necessarily confined to an autonomous movement of lexical elements to grammatical affixes. Furthermore, I assume that paradigmatic balancing of loosely tied subsystems can arise in more than one center, which is reflected in the well-known fact that there are often alternative designs for paradigmatic structures (cf. Hendrikse 1993: 186).

⁶ This is probably the main idea behind the concept and treatment of functional categories in GB. However, I reject any preferential treatment of linear ordering among different grammatical means, focusing, instead, on the paradigmatic (rather than the syntagmatic) interplay between them. Furthermore, I do not want to alter the traditional, formal distinction between, for instance, affixes and reduplication, implying that the latter are also affixes just because they can and should be classed under a general term. Bloomfield used the term "means of expressions" (1914) and later "features of grammatical arrangement" (1933/1973: 163ff.) (cf. Matthews 1993) for some of the formal means mentioned above, excluding grammatical morphemes, which have themselves a 'meaning', and, insisting on the difference between bound and free forms. Contrary to him, I consider both distinctions of secondary importance. The fact that languages mix and organize different means for grammatical categories in a paradigmatic fashion seems to me linguistically more important than the fact that many languages prefer only one type (cf., for example, Bloomfield's argument for dividing strictly between bound and free forms).

For the English system of determiners and NUMBER, more than one such alternative analysis exists. Lønning (1987), for instance, proposes a representation in which the entire system of determiners is cross-classified with NUMBER. According to this, quantifiers and articles⁷, just as demonstratives (*this* vs. *those*), are assigned the binary values SG and PL (*much* (SG) vs. *many* (PL); *the* (SG + PL)). Just as common - or probably more common - is an analysis for the interaction of determiners and NUMBER, which focuses on the similarity of MASS/SG and COUNT/PL. But there seems to be no way of designing a unified paradigm without (or at least only with little) syncretism in which all contrastive values are associated with distinct formal means for all parameters.

The lack of isomorphism is probably more often the rule than the exception in the languages of the world, not only in strong inflectional languages like Latin. The idea of paradigm was developed just on the basis of languages with a great deal of syncretism, and it was precisely such linguistic conditions which served as a main argument for the morphological paradigm model (cf. Matthews 1974). But for a fully developed paradigm, it remains a strong requirement that the identity of grammatical parameters (i.e., categories) must be warranted. That is, semantic and formal variations should not take place simultaneously on the level of grammar. If the semantic interpretation (including semantic implications) of a hypothetical categorial distinction (ASPECT, for instance) changes over different dimensions, it is crucial that the formal means remain the same. Formal variations, on the other hand, presuppose semantic constancy and unambiguous syntactic environment. Both cases are usually modeled in terms of lexical sensitivity. In the first case, semantic shifts are correlated with lexical-semantic classes (see the interpretation of aspectual values in inflected ASPECT systems like Modern Greek or Spanish); in the second case, formal variations are correlated with purely 'formal classes' (see the morphological verb classes in Romance languages).

It is precisely in this way that the English apparatus of determination and quantification does not constitute a fully developed macroparadigm with a categorial bifurcation between MASS and COUNT. It is well known that the English system of determination and quantification with respect to the MASS/COUNT distinction contains several breaks, namely breaks between the domain of generic and non-generic uses and between 'concrete nouns' and 'abstract nouns'. Whereas differences with respect to determination and quantification can be modeled on the basis of lexical classes in the case of 'concrete nouns' and 'abstract nouns', such a solution is not possible for generics and non-generics. Under these circumstances, it is also not very surprising that the configuration *a* + SG is proposed to be treated, in the environment of 'concrete nouns', as COUNT, and in certain environments of 'abstract nouns', as MASS, as

⁷ This demonstrates that the correlation with the word-forms as PL vs. SG is concerned here, rather than the autonomous semantic value of determiners alone.

suggested by Bunt (1985: 14-15). At the same time, however, just such proposal makes the categoriality of the MASS/COUNT distinction doubtful.

2.2.2 Lexical Sensitivity

At a first approximation, we may understand under lexical sensitivity that lexical forms, used in actual utterances (i.e., inflected, etc.), show certain privileges for occurring in a certain position of a particular construction associated with one of the values of a grammatical category rather than with others. Such built-in preferences result in a certain division of that part of the lexicon which constitutes the scope of the respective category. Thus, the lexicon of Tagalog, a language famous for the difficulties it poses with respect to the noun-verb distinction, is not subdivided with respect to the distinction between referential phrase (called topic or subject) and predicate phrase, but it is certainly subdivided in several other ways.

2.2.2.1 Form-Classes, Grammatical Categories and Ambiguity Phenomena

2.2.2.1.1 Bloomfield

For our purposes, it is quite intriguing to study Bloomfield's understanding and usage of the term 'form-classes' (cf. 1933/1973: Chapter 16 "Form-classes and Lexicon"). First, form-classes were classes of the entire lexicon, or of a subpart of the lexicon defined on the basis of single distributional privileges of lexical forms ("functions" in the terminology of Bloomfield). That is, 'each feature' could partition the lexicon and thus establish form-classes. Bloomfield strongly emphasized that different distributional criteria ("functions") "may create overlapping form-classes" (1933/1973: 265). That is, he claimed, on the one hand, that form-classes normally "cross each other and overlap and are included one within the other, and so on" (1933/1973: 269). On the other hand, he registered that some distributional privileges are shared by a lot of lexical forms. This clustering defined, in his opinion, a "large form-class". Large form-classes completely subdividing the lexicon, or part of it, into classes of approximately equal size, in turn, were seen as the essence of grammatical categories, including parts of speech, all inflectional categories and categories marked by 'function words', such as definiteness. This was clearly an intensional definition, distinct from an extensional definition of large form-classes also subdividing the lexicon into classes of equal size, but based on a single distributional criterion. Bloomfield did not touch the question as to whether or not these category-defining classes were mutually exclusive. The reader could also be led to the following conclusion: Whereas criteria examined in pairs often result in non-disjunctive (lexical) classes, they might collectively result in disjunctive (grammatical) classes, provided that we find at least one necessary criterion which can always be applied. Thus, no explicit hints were given concerning the relation (probably an inclusion relation) between the small or large extensionally-defined classes (for instance, the class of irregular plurals), and the large intensionally-defined classes (for instance, the classes of singular and plural as grammatical categories). Between the mutually exclusive but clumsy classes obtained by the intersection of the 'one feature-classes', and the classes constituting categories, there has remained a wide space for different interpretation until today.

While proceeding from the extensionally-defined to the intensionally-defined classes, Bloomfield's perspective shifts from the lexicon, which can be structured in infinite ways, to the paradigmatically organized grammar, where a definitive structure is imposed on the

previously cross-classified lexicon according to salient features.⁸ This is also reflected in his abstract treatment of 'class-cleavage', that is, the assignment of 'dual categories' (MASS and COUNT) to lexical forms, such as *egg* ("*the egg, an egg*") and *salt* ("*Epsom salts*") (examples are taken from Bloomfield (1933/1973: 265).

What Bloomfield does not address explicitly is the interrelation between types of meaning differences associated with single (or phonetically identical) lexical forms and class-cleavage associated with different class-meanings. He draws a distinction between homonymy, "transferred meaning", corresponding to conventionalized and productive/systematic polysemy (metaphors) and cases ("narrowed meaning"), which we would now classify partly as generality, partly as productive/systematic polysemy (metonymies). Is class-cleavage the consequence or instantiation of "narrowed meaning" or does it constitute a new type of 'meaning shift' which is applicable both to the primary ("central") meaning and to the "transferred meaning"? On which entities should the procedure leading to the establishment of large word classes like MASS/COUNT be grounded at all? We can only guess that Bloomfield's intention was to separately examine 'different linguistic forms' (homonyms), in contrast to 'identical linguistic forms with transferred meanings', where the primary meaning, usually a spatio-temporally concrete one, was designated for distributional examination.

2.2.2.1.2 Post-Bloomfieldians

For Bloomfield, the phenomenon of ambiguity did not pose a special problem since he invoked lexical and phrasal meaning, obtained through intuition, translation or whatever, as a very natural criterion, and primarily argued against the ontological foundation of class-meanings. By comparison, the concept of ambiguity was not compatible with the spirit of linguistics of his successors. It was claimed that heuristic studies for mutually considering semantic and formal variations were not necessary. The very idea that systematic semantic variation is less interesting than formal variation and that the latter can be studied independently of the former has had fatal consequences right through to our own times. It is well known that there was a great deal of controversy among structuralists about whether or not it is possible to determine forms and constructions without relying on meaning, and to which degree is it necessary to use meaning (lexical meanings and phrasal meanings) as heuristic 'hints'. Some linguists, such as Fries (1954: 60, 68), explicitly admitted that "a certain control of specific kinds of meaning" are, in fact, "on all levels of linguistic analysis"

⁸ Cf. "For this reason a system of parts of speech in language cannot be set up in any fully satisfactory way: our list of parts of speech will depend upon which functions we take to be most important." (1933/1973: 269).

necessary. All in all, in such a climate, it was inevitable that lexical meaning and phrasal meaning appeared as a random variable (for a very good discussion, see Matthews 1993).

It is significant that Harris (1951/1960) devotes only a few lines to the problem of homonymy. Here, he argues for a 'more rigorous' procedure, by which homonymy will be derived from the results of the distributional analysis for identifying morphemes and morpheme classes, rather than being presupposed. He predicts that homonyms ("meaning differences") will correspond with different environments (1951/1960: 189). If the environment range of a morphemic segment, tentatively seen as one single morpheme, can be partitioned in such a way that one subpart matches with the total environment of a morphemic class already established, and another subpart with the total environment of another class, two different (homonymous) morphemes are allowed to be established. The success of this method depends heavily on the issue of how fine-grained the morpheme classes are. Unfortunately, but consequently, Harris requires that parts of the distribution will be considered unchanged (unambiguous or monomorphemic) for the entire procedure.⁹ However, Harris' approach does not provide heuristic steps which permit treating a form, previously determined as ambiguous (i.e., as two morphemes) for reasons of split distribution, as ambiguous, too, if it appears in new tests as part of the tested environment. Certainly, linguists have no problem (in European languages) in treating nouns and verbs as different entities in the tested environment (i.e., test-frames). However, they do have problems with the hypothetical ambiguity of the English *one* by testing nouns as MASS or COUNT. And the method of relying on forms in the way illustrated above does not tell us what to do with the following sort of ambiguity:

(1) *There is only a little apple in this apple pie.*¹⁰

There is a further problem to be mentioned which conceptually roots itself in structuralism and is still alive. In addition to considering complementary distribution as evidence for meaning difference, we can also make the opposite argument borrowed from phonology: complementary distribution indicates identity of meaning and can be taken as evidence for it.

⁹ "For how do we know that the frame remains unchanged while various would-be morphemes are substitutes in it (cf. the Appendix 4.22), and how do we know that when we test the /beriy/ of *boysenberry* in the utterance *blueberry* we have the 'same' /beriy/? It is necessary therefore to agree that a phonemic sequence will be considered unchanged as to its morphemic segmentation if part of its environment, in one utterance, is replaced by another stretch, making another utterance; that is, given the utterance XY (where X, Y, Z are each phonemic sequences), if we substitute Z for X and obtain the utterance ZY, we will consider that in these two utterances the two Y's are morphemically identical." (1951/1960: 159, Footnotes 6)

¹⁰ Cf. McPherson 1991: 335.

2.2.2.1.3 Contemporary Approaches

Twenty-eight years after the publication of "Methods" (Harris 1951/1960), Dowty (1979) discusses the issue of lexical classification with respect to aspectual classes established in phrasal terms, that is, on the sentence level, without considering whether or not there are morphologically simplex elements for all aspectual classes in a language. He explicitly draws the reader's attention to the problems of ambiguity and that of lexical classification which is certainly not limited to the double categorization vs. underspecification of ACCOMPLISHMENT verbs:

"The problem of distinguishing between lexical verbs which *must* be accomplishments, those which *may* be accomplishments with the right time adverbs, and those which can be accomplishments only under special interpretations is an interesting and difficult one, involving as it does the thorny problems of polysemy versus homonymy."

(Dowty 1979: 62)

In my opinion, two interrelated points are at issue here. The first one concerns the existence of lexical subregularities interweaving the lexicon. Recurrent semantic contrasts correlating with different (sets of) contexts in which a lexical form in the verbal domain is used are likely to be ASPECT-sensitive.¹¹ That is, the different members of a verb alternation can exhibit different aspectual behavior. This may either confirm the classes established on the basis of verb alternations (on a single alternation or on a set of alternations) or it may introduce a cross-classification (cf. Levin 1993 with respect to the interrelation between diathesis alternations and verb classes; cf. also Dowty 1991).¹² The second point concerns the aspectual interpretability on more than one dimension, such as a dimension of 'spatio-temporal concreteness', a dimension of habituality, etc. (cf. Carlson 1981). Here, it is not a priori clear whether we really have to do with semantic shifts or at least with the same type of bidirectional semantic alternations rooted in the lexicon as in the case of usual verb

¹¹ I am assuming here that the alternation patterns resulting in verb classes are not directly based on aspectual behavior. If this is the case, the argument presented above has to be taken for another syntactic category, rather than ASPECT.

¹² There is an interesting class of verbs in German which exhibits a lexically established ambiguity between a 'singulative' sense and a 'durative' sense, including, for instance, *klopfen* ('knock', 'tap'). This class of verbs shares its behavior (a) with verb classes indifferent to the syntactic ASPECT classes ACTIVITY and ACCOMPLISHMENT and (b) with verb classes generally assigned to the syntactic class ACHIEVEMENT (cf. Behrens 1990).

alternations. However, the effects are the same: Different clusters of systematic verb alternations cause us to establish lexical verb classes by cross-classification in the same way as the 'idiosyncratic'¹³ restrictions of a lexical form with respect to different aspectual dimensions. An analogous example from the field of the MASS/COUNT distinction can be given here. The first case (i.e., systematic verb alternations) can be seen as parallel to the lexically established systematic ambiguity in some languages which permits the use of the same lexical form for referring to ANIMALS and to the corresponding FOOD (*lamb, fish*). In the second case, we can think of the different and 'idiosyncratically' restricted compositional devices of English nouns used for referring to SORTS, rather than for referring in concrete space and time:

- (2) a. **these golds / these sorts of gold*
 b. *these milks / these sorts of milk*

2.2.2.2 Some Heuristic Proposals

I take it for granted that the assignment of lexical categories cannot be based on morphemes, but only on 'lexical units' as, for instance, defined by Cruse (1986: 49, 77). That is, lexical categorization should only be applied to (lexical) 'form-meaning complexes', i.e., the 'union of an abstract (uninflected) form and a single stable sense' rather than to forms alone. This implies that tests of ambiguity (identity test, zeugma test, synonymy test, etc.) have to precede tests for behavioral properties indicative of grammatical categories. All those pairs which clearly satisfy the tested criterion of ambiguity (i.e., probably all cases of classical homonymy and a good deal of conventionalized polysemy) have to be separately tested for grammatical

¹³ 'Idiosyncratic' does not necessarily mean 'random' or 'ad hoc', rather than 'not interpretable on the basis of known criteria'.

categories.¹⁴ Compare, for instance, the following 'strongly' ambiguous lexical units (cf. Pinkal 1985 for the use of the terms 'starke Ambiguität' and 'schwache Ambiguität'): ¹⁵

(3) *der Schein* (Genitiv: *(e)s*, Plural: *e*¹⁶)

1. 'light'
2. 'appearance(s)'
3. 'certificate'

Contextually specified senses, proved as generality or vagueness (on the other end of the spectrum of ambiguity), can be combined as single lexical units. However, all the cases in-between, partially discussed above and commonly subsumed under the broad term of systematic ambiguity or polysemy, pose difficulties. As already mentioned, they interweave the lexicon, occurring in smaller lexical fields or in an entire major lexical category; they can also cross-cut major lexical categories. Some of them have suppletive or morphologically-derived correspondences in other languages; others have no correspondences at all. Some of them directly address the relevant concern of how a grammatical category imposes a classification on the lexicon. Take, for instance, the recurrent contrast between the senses instantiated in expressions referring (a) to ANIMALS and (b) to FOOD (*lamb*) or (probably; cf. the discussion in 3.1.2.6) between the senses instantiated in expressions referring to (a) MATERIALS and (b) to natural and discrete occurrences of the same MATERIAL (*stone*). Other

¹⁴ In languages where no major lexical categories have yet been established, we will, of course, not begin with ambiguity tests restricted to lexical categories but we will start the analysis on a higher level. Here, form tests and meaning tests must be simultaneously applied from the very beginning, also for identifying major lexical categories, that is, hypothetical results indicating ambiguity must be successively integrated in the analysis of major lexical categories. In a language with established major lexical categories, we can split the job of identifying 'real' homonyms in a lexical category, and the job of identifying 'pseudo' homonyms, which were classified as homonyms only due to their property of cross-cutting established lexical categories. These 'pseudo' homonyms may indicate a 'wrong' categorization.

¹⁵ I do not group strongly ambiguous lexical units automatically in different lexemes for two reasons. First, I assume that more than two hierarchical levels (different lexemes vs. different senses of a lexeme) are possible (cf. Behrens 1995). Second, I reject the lexicographic and linguistic practice, according to which members of a systematic ambiguity are automatically assigned to different lexemes simply because they belong to different major lexical categories.

¹⁶ The three senses or lexical units have a common etymological source. Typically, the monolingual dictionary DUDEN groups them together under one headword, whereas the bilingual dictionary LANGENSCHIEDT establishes three different headwords, which is a typical method for homonyms. Both DUDEN and LANGENSCHIEDT mark the second lexical unit with 'no plural'. The first lexical unit is 'rarely used in plural' for DUDEN and 'never used in plural' ('no pl') for LANGENSCHIEDT. The third lexical unit has a plural form.

lexical subregularities are indirectly relevant, such as the lexical categorial ambiguity found in the lexical field of MATERIAL (*silk, gold*) or in the field of HUMAN PROPENSITIES (*adult*). It is the interplay of different alternations that is not very well understood. There are alternation chains such that one member of an alternation simultaneously acts as a member of another alternation (cf. the chain *open* (ADJ) - *open* (V_{INTR}) - *open* (V_{TR})). Moreover, alternations can be neutralized in certain lexical and/or grammatical contexts. This neutralization of contexts certainly has a bearing on the classificatory effects a grammatical category imposes on the lexicon, as long as members of the paradigm of this category themselves constitute neutral contexts. In summary, the study of categorization cannot be separated from the study of systematic ambiguities.

2.2.3 Interrelations Between Paradigmatic Uniformity and Lexical Sensitivity

We will now turn to the interrelation of paradigmatic uniformity and lexical sensitivity of grammatical categories. The four logical possibilities are shown in Table 1:

Paradigmatic Uniformity	Lexical Sensitivity
+	-
+	+
-	+
-	-

Table 1

'+' indicates existence or strong value, '-' indicates non-existence or weak value.

We encounter uniform (grammatical) paradigms or subparadigms without neutralization of the relevant distinctions and, in addition, almost totally independent from the lexical organization. Subparadigms organized along the dimension of PERSON often show both properties (cf. first row in Table 1), as long as they are not fused with other categories such as TENSE, ASPECT, etc.¹⁷ Formal variations can, of course, violate the biuniqueness principle here, too, but these do not have semantic correlates and thus establish purely formal classes, i.e., the category of PERSON does not show semantic variation with respect to the interpretation of its values depending on lexical semantics.

Other grammatical categories are lexically sensitive, can nevertheless be organized in a uniform paradigm (cf. second row in Table 1). In contrast to PERSON, ASPECT, for instance, is generally not independent from the entire lexical organization.¹⁸ That is, ASPECT markers, whether or not expressed as free forms, may be organized in a uniform and maximally contrastive way (as in Modern Greek or Spanish), but the non-applicability and the interpretation of certain ASPECT forms is dependent on lexical-semantic classes (for instance, resultativity vs. ingressivity if a 'perfect' value is applied; cf. Comrie 1976; Sasse 1991).

¹⁷ This is not exactly the same as saying that PERSON is a 'very good inflectional' category and is therefore placed, in universal hierarchies of grammatical categories, on that end or near that end of the hierarchy which is associated with inflection (cf. Bybee 1985). Clitic pronouns or the combination of pronouns and affixes may also lead to a uniform and maximally contrastive paradigm.

¹⁸ There is, perhaps, a qualitative gap between PERSON and other grammatical categories. It is significant that languages having no stem-modifying markers for PERSON but for ASPECT, etc., give rise to notorious difficulties in evaluating these markers as inflection or derivation (cf. Tagalog).

Aspectual differences via AKTIONSART, as found in Hungarian, exemplify the third row in Table 1. The main difference between ASPECT and AKTIONSART does not reside, as often assumed, in the fact that the former is a purely grammatical and the latter a purely lexical category. Both are lexically and grammatically sensitive. However, the paradigmatic organization of ASPECT allows a more clear-cut distinction between contrastive grammatical values than that of AKTIONSART.¹⁹

In English, the MASS/COUNT distinction seems to lack the necessary criteria of an ideal paradigm and is, as a grammatical category, clearly lexically sensitive. My hypothesis is that the MASS/COUNT distinction represents a genuine mapping problem. It is, however, a problem of mapping between the general grammatical organization and the general lexical organization of a language, rather than a problem of mapping only between single units on the lexical level and single grammatical values on the sentence level. If neither grammatical categories are disjunctively defined nor corresponding lexical categories can be identified by necessary and sufficient criteria, specific strategies are required to adjust the mapping between lexicon and grammar. Linguists have developed a number of standard strategies for this purpose.

2.3 Linguistic Strategies for Mapping Between Lexicon and Grammar

I will illustrate these strategies by an excursus on a problem very closely related to the MASS/COUNT issue, namely, that of the adjective-noun distinction.

2.3.1 Adjectives and Nouns: A Brief Case Study

It is well known that from a notional perspective, it is very difficult to distinguish between 'properties' and 'objects'. To use Miller's words:

"...if an object is thought of as a bundle of properties, it is impossible to distinguish between an object and its properties. The properties are the object."

(Miller 1985: 207).

In contrast to philosophers and logicians, however, linguists have commonly taken for granted that the part of speech distinction between adjectives and nouns provides a perfect

¹⁹ One can consider the productive use of prefixes, which are often said to be AKTIONSART markers in some languages (for instance, Hungarian) as overt markers of the same semantic alternations as those attested in languages having a classical ASPECT distinction on the sentence level (Modern Greek, Spanish).

linguistic indication for a distinction between objects/things and properties (cf. Lyons 1977: 447). The fact that a categorial difference between adjectives and nouns is attested for many languages has always been interpreted as evidence for the existence of a corresponding ontological difference. Linguistic criticism of predicate logic, for instance, has notoriously concentrated on the neglect of such crucial factors as the copula or the part-of-speech distinction (cf. Jackendoff 1983/1988: 60ff). The linguistic tradition has, of course, not denied that languages show differences in subcategorization; they may, for instance, group stative verbs and 'adjectives' together, or 'nouns' (as names of universal terms) and 'adjectives' (as names of general terms), etc. That is, we are told either that 'property concepts' as real candidates for 'adjectives' may appear cross-linguistically as subclasses of 'nouns' or 'verbs' (cf. Thompson 1989: 247), or that 'nouns' or 'verbs' build subcategories of a higher-order category together with 'adjectives' (cf. Lyons 1977). For English, the possibility of considering 'adjectives' and 'nouns' as subcategories of a higher order 'nominal', in a way analogous to the Latin model, is commonly rejected (cf. Thompson 1989).

This is a bit surprising since there is a great deal of syntactic overlapping between these two classes. First, adjectives can be productively used in NPs without an additional overt nominal head referring to "groups of people who share the same characteristic or quality" (Collins COBUILD English Grammar 1990: 21).²⁰ Used as subjects, such phrases show a plural agreement; however, the adjective itself is never marked with an overt plural affix:

- (4) a. *The rich have benefited much more than the poor.*
 b. *He pities the poor.* / **He pities the poors.*

Second, adjectives and nouns share the prenominal distribution. That is, generally, both can modify other nouns in this position without necessarily altering the respective lexical form (i.e., without derivation). Neither of the further indicative features for adjectives, namely (a) that they can be modified by *very*; and (b) that they can build comparative forms, can be seen as a necessary criterion for adjectives (cf. non-gradable adjectives). Moreover, if *very* is used as a distinguishing criterion, it must be treated as ambiguous, since it can also modify nouns (*the very question*).

In Hungarian, the categorial (non-)distinction between adjectives and nouns is, in many respects, similar to that in English. However, here, the syntactic overlapping is more extended and more significant. Lexical units, classified as adjectives and nouns, equally appear as uninflected noun modifiers. In addition, adjectives and nouns also generally overlap in

²⁰ For the COBUILD GRAMMAR, these adjectives are headwords and should be chosen in the mentioned case rather than nouns as headwords.

predicative position since there is no such class of nouns which would require an article in this position (cf. (5)) and the English translations with and without the article). Furthermore, not even the comparative form is excluded for some of those lexical forms which are commonly classified as nouns.²¹ Example (6) demonstrates how the lexicalized metaphor *szamár* ('donkey' > 'a stupid person') is used in (a) attributive position, (b) predicative position and (c) as a comparative form.

(5) *Ez autó / arany / szerelem / kérdés.*

This car / gold / love / question

'This is a car / gold / love / a question.'

(6) a. *Ez a szamár Péter megint lekéste a vonatot.*

This donkey P. again missed the train

'This dummy, Péter, missed the train again.'

b. *Szamár vagy.*

donkey are-you

'You're a dummy.'

c. *Péter még annál is szamarabb.*

P. still at this also more donkey

'Péter is even more a dummy (than I/you... thought).'

Searching dictionaries through for both languages shows that lexicographers favor a double classification (adjective + noun) in very similar semantic fields. These are, and this is not very surprising, exactly those which are likely not to be assigned to the subcategory 'adjective' in languages lacking a major lexical category 'adjective'. I will concentrate on two significant groups:

²¹ For English, one might argue that as long as adjectives and nouns are productively used in the attributive position, different stress patterns indicate the category difference correlating with those classes which were established on the basis of the other distinguishing criteria. The modified noun receives stress only if the modifier can be regarded as a member of the class of adjectives rather than nouns. That is, adjectives would not automatically generate compounds according to this argument. However, this difference between adjectives and nouns can be neutralized in lexicalized compounds (*blackbird*), as well as when there is a contrastive stress on the modifier (*black bird*). In Hungarian, it is even more difficult to find different stress patterns between modifying adjectives and nouns; there is a clear stress difference only between lexicalized and non-lexicalized constructions. The stress patterns of non-lexicalized constructions probably also depend on factors such as the complexity and length of the modified unit. In any case, orthography is a bad indicator for real stress behavior in both languages.

- (a) AGES, NATIONALITIES and other HUMAN PROPENSITIES
- (b) MATERIALS

In English, typical examples for the first group are *adult* and *American*. According to the common lexicographic practice (cf. COLLINS), both are classified as adjectives and nouns and paraphrased with the following standard pattern:

- (7) a. having the **PROPERTY X** (adjective) vs.
a person who has **PROPERTY X** (noun)
- b. relating to the **NAME OF THE STATE X** (adjective) vs.
a native or citizen of the **NAME OF THE STATE X** (noun).

According to the common linguistic practice, the two categorially different lexical units are distinguished in the predicative position by the presence or absence of the indefinite article:

- (8) *Peter is American* (adjective) / *an American* (noun).

Two corresponding examples in Hungarian are *őreg* ('old') and *amerikai* ('American'). Since, in Hungarian, all nouns can appear without an indefinite article in the predicative position, there is no analogy in the way in which one could analyze a predicatively-used expression as noun or adjective:

- (9) *Péter amerikai*.²²
P. American
P. is American / an American.

Both languages show roughly the same lexical subregularities in the semantic field of MATERIALS. The most important semi-productive alternation patterns, which are listed in English and Hungarian dictionaries in addition to the (primary) MATERIAL or SUBSTANCE sense as further senses or 'uses', can be seen in (10) (the paraphrasing patterns underlying (10) are almost identical in both languages):

- (10)a. the **COLOR** of **X**; **ARTIFACTS** (largely) consisting of **X**
- b. consisting of **X**; having the **COLOR** of **X**; covered with **X**

²² *amerikai* has an overt adjective marker (-i), but it is, in the same way, double classified in dictionaries like all other expressions for NATIONALITIES without an adjective suffix (*német* ('German'), *japán* ('Japanese')).

The senses under (10) b. are commonly marked either as adjectives or as 'modifying' (COLLINS)/'attributive' (MNYÉSZ) uses of the 'noun' (i.e., as new senses, but not necessarily assigned to one of the noun senses listed in the same dictionary entry). These senses together constitute a systematic pattern, which can also be seen with overtly derived adjectives, and it is largely a matter of conventionalization whether the derived or the underived one is used as modifier (*gold carpet, golden disc*). Often both are possible, as in *metal/metallic paint*. (11) is a randomly selected list from the dictionaries COLLINS and MNYÉSZ, indicating the classification as noun (N), adjective (ADJ), noun used as 'modifier' (MOD)/'attribute' (ATTR):

(11)

<i>gold</i>	N; ADJ: <i>gold carpet</i> ; MOD: <i>gold mine</i> (cf. <i>golden</i>)
<i>silver</i>	N; ADJ: <i>silver hair</i> ; MOD: <i>silver coin</i> (cf. <i>silvery</i>)
<i>sand</i>	N; but cf. <i>sand castle, sand bank, sand-beach</i> (cf. <i>sandy</i>)
<i>silk</i>	N; MOD: <i>silk dress</i> (cf. <i>silky</i>)
<i>metal</i>	N; ADJ: without examples (cf. <i>metallic</i>)
<i>arany</i> ('gold')	N; ATTR: <i>arany ékszer</i> ('gold jewel'), <i>arany könyv</i> (golden book), <i>arany szív</i> ('gold heart') (cf. <i>aranyos</i> ('golden'))
<i>ezüst</i> ('silver')	N; ATTR: <i>ezüst ékszer</i> ²³ ('silver jewel'), <i>ezüst fény</i> ('silver sheen'), <i>ezüst hang</i> ('silvery voice') (cf. <i>ezüstös</i> ('silvery'))
<i>homok</i> ('sand')	N; but cf. <i>homokvár</i> ('sand castle'), <i>homokpad</i> ('sand bank'), <i>homokpart</i> ('sand-beach'), <i>homoktalaj</i> ('sandy soil'), <i>homokréteg</i> ('sandy stratum') (cf. <i>homokos</i> ('sandy'))
<i>selyem</i> ('silk')	N; ATTR: <i>selyem fehérenemű</i> ('silk underwear'), <i>selyemharisnya</i> ('silk stockings') (cf. <i>selymes</i> ('silky'))
<i>fém</i> ('metal')	N; but cf. <i>fém ékszer</i> ('metal jewel'), <i>fémtükrő</i> ('metal/metallic mirror'), <i>fémfény</i> ('metallic lustre') (cf. <i>fémes</i> ('metallic'))

It is not my intention here to address the usual inconsistencies found in dictionaries. Nevertheless, we have to ask the question: Why, in these cases, are prenominal uses recorded as 'modifying' or 'attributive' ones? This is a feature of all nouns, most of them, however, are not specified in this way. And, what is the difference between a 'modifying use' and an 'adjective category', as found in COLLINS? All Hungarian and English entries tested point to a certain lexicographic ratio: The decisive factor for the specification as 'attributive'²⁴ in

²³ Hungarian orthography is not indicative of the category, as mentioned already in footnote 21. See, for instance: *arany ékszer* ('gold jewel') vs. *aranylánc* ('gold chain') vs. *arany nyaklánc* ('gold neck-chain').

²⁴ Generally, bilingual dictionaries are less cautious and use 'adjective' instead of 'attributive' (cf. HALÁSZ, ORSZÁGH).

MNYÉSZ is not the semantic interpretation of the lexical form in question itself, but, the sense relations between this lexical form and its morphologically derived counterpart. As long as the derived counterpart has a synonymous sense and can be substituted as an attributive element, a strong tendency is observed for marking all prenominal uses as 'attributive'. This is a good example for a specific sort of semantic motivation for assigning syntactic categories, namely, a lexically-based semantic motivation taking into account paradigmatic issues. The assignment of the lexical category 'adjectives' in COLLINS is also largely based on semantic considerations in the paradigmatic dimension. All those senses of a prenominally-used lexical form, which are synonymous to senses of a derived lexical form, tend to be coded as 'adjective', with the exception of those which can be paraphrased as 'consisting of x'. It is obvious that these adjectives would represent defective members of the lexical category 'adjective'.

2.3.2 Lack of Necessary Criteria for Lexical Categories

The lack of necessary criteria for lexical categories simultaneously considered as syntactic categories (cf. Bloomfield's large form-classes) is commonly seen as the defectivity of the category members. This is generally represented (i.e., in various lexicon types) either by single restrictive features ('no attributive use', 'no comparative form' for adjectives, 'no plural' for nouns, etc.), or, by corresponding subcategories ('adjectives used only attributively', 'adjectives used only predicatively', etc.).²⁵ The subcategory solution is favored mainly in the case of major lexical categories (i.e., adjectives). However, I have never heard of a proposal for subcategorizing MASS and COUNT nouns. Allan (1980), for instance, proposes a fine-grained subcategorization of the entire domain rather than subcategorizing the two subdomains of MASS and COUNT nouns.

Lehrer (1987: 96) notes that one may doubt the legitimacy of the class membership of an item which lacks more than half of the properties maximally characterizing a class (i.e., those which are "found in the prototypical members"). What matters, however, is **which** criteria are or are not satisfied, rather than **how many** are satisfied or not. If the categories, defined from the perspective of the grammar (syntax), are, in principle, allowed to have shared

²⁵ For German, the following four criteria are commonly employed in defining subcategories: 'can be used attributively', 'can be inflected', 'can be used predicatively', 'gradable (can build comparative forms)' (cf. CISLEX München). Since the criteria are (empirically) not entirely independent, this results in five feature configurations (subcategories): +ATTR/+INFL/+PRED/-GRAD (*ledig* ('unmarried')), +ATTR/+INFL/-PRED/-GRAD (*väterlich* ('paternal')), +ATTR/-INFL/+PRED/-GRAD (*super* ('super')), +ATTR/-INFL/-PRED/-GRAD (*Schweizer* ('Swiss')), -ATTR/-INFL/+PRED/-GRAD (*entzwei* ('in two'))).

distributions, criteria related to these distributions are always problematic. In other words, the category membership of an item, which has only such properties which are shared by the members of other categories, can be seen as more critical as the category membership of an item which has only a few, perhaps peripheral but exclusive, properties of the relevant category. For German adjectives and adverbs, the predicative position represents an overlapping distribution established by the grammar. The same was shown for Hungarian nouns and adjectives with respect to the attributive and predicative position, and for English nouns and adjectives, with respect to the attributive position in lexicalized compounds, and to the predicative position in certain semantic fields. Actually, there are proposals for prohibiting defective members which are restricted to a distribution shared with other categories, for instance, 'only predicative adjectives' in German (cf. Bergenholtz 1983: 24).

In the case of the MASS/COUNT distinction, another problem surfaces. Here, an autonomous lexical feature 'MASS/COUNT' is sometimes assumed, which is represented independently from information on defectivity, i.e., from 'idiosyncratic' restrictions of those noun features (NUMBER behavior, determiner behavior, etc.), which jointly define the actual lexical classes relevant for the syntactic MASS/COUNT behavior.

2.3.3 Lack of Sufficient Criteria for Lexical Categories

As mentioned, the lack of sufficient criteria can be a matter of general grammar design which results in an a priori overlapping distribution of categories. In other cases, members of a category appear, perhaps satisfying all definitory criteria, in a distribution which is exclusively reserved for a different category. The productive use of English adjectives in NPs without an overt noun as head (*the poor, the rich*; cf. (4)) is a good example of such a constellation. In (12), 5 typical strategies are listed, which have been developed in the history of linguistics to deal with categorial overlapping and categorial defectivity.

(12)

multiple categories

mixed categories

dynamic category change

ellipsis

change of grammar

2.3.3.1 Strategy of Multiple Categories

This is the simplest strategy, and which has the worst reputation among linguists. According to this strategy, all lexical forms which can be associated with distributions canonically belonging to more than one (syntactic/grammatical) category are represented with multiple category labels in a static lexicon. That is, *poor* and *rich* are represented as adjectives and (defective) nouns; *adult* and *American* in English, *öreg* ('old'), *amerikai* ('American') and *szamárr* ('donkey' > 'stupid') in Hungarian, and probably all expressions in the semantic field of MATERIALS are likewise double-categorized. The most popular argument against this strategy states that it would 'unnecessarily' almost double the number of nouns in the lexicon (Allan 1980; Dowty 1979). If this argument is grounded on a vague linguistic principle determining the design of linguistic models, such as 'elegance', it is not very relevant, at least not for issues of the mental lexicon and of computational lexicon models. However, there are actually some more serious arguments against this strategy.

First, it basically neutralizes the important differences between different ambiguity types: between homonymy, lexicalized polysemy and different sorts of more or less productive alternations (thus, also between the *poor* case and the *adult* case). A second problem is pointed out by Allan (1980). He remarks, in connection with the MASS/COUNT distinction, that the strategy of lexical double categories fails "to show that some nouns are more countable than others" (1980: 546). Putting it more generally, the strategy of multiple categories fails to capture subclasses of the multiple-categorized class differing with respect to the range of their overlapping and defectivity. One might argue that the lexical specification of defectivity for every category of every entry contains implicit information about such subclasses which, in turn, can probably be interpreted as a reflex of ambiguity types. (The *poor* case is associated with restricted NUMBER behavior, for example.) This is probably true in the domain of productive ambiguity patterns. However, this does not work if we intermix homonymy with productive ambiguity since instances of the former can show the same defectivity pattern as instances of the latter.

The third argument against multiple categorizing is a rather difficult one. Let us assume we allow double categories (noun and adjective) to be assigned to all those lexical forms in the semantic field of MATERIALS having at least one (lexically established) sense, which can be paraphrased as 'material' or 'substance', i.e., not only to *gold*, *silver*, *silk*, but also to *sand*, *metal*. Without further differentiation, we could not decide in this case, as in the case of the predicatively-used expressions for human properties in Hungarian (cf. (5)) and in contrast to English *adult* and *American*, whether the predicatively-used form 'goes back' to an adjective or a noun (i.e., *gold* in *This is gold*). If we confine the category 'adjective' to those cases in which the attributively-used lexical form has a sense other than 'consisting of the MATERIAL X', we fail to capture the systematic sense relations exploited in the attributive position and paralleled by derived adjectives (cf. (13)). Sometimes we are told that attributive nouns

resemble adjectives, but it can be shown by corresponding prepositional phrases that they are really 'nouns' rather than adjectives (cf. Quirk et al. 1972: 240-241):

(13) *a stone wall vs. a wall made of stone*

However, this argument is grounded on the following highly problematic assumption: If two constructions can be seen as paraphrases of one another, then the corresponding word-forms occurring in different positions in these constructions must be based on a single lexical unit with a single category.²⁶ In a stronger version, this assumption would imply that even noun-verb alternations ('conversion') are to be considered as instantiations of a single category, as long as we are able to build sentences with them which stand in a paraphrase relation.²⁷ In my opinion, the approach of syntacto-semantic alternations (cf. Levin 1993) calls exactly this assumption into question, focusing on regular correlations between semantic interpretation and syntactic/grammatical environment. The latter may, but need not, be associated with different lexical categories or subcategories.

2.3.3.2 Strategy of Mixed Categories

²⁶ As Peter Lutzeier has pointed out to me (p.c.) the crucial question is: What establishes categorial identity? Paraphrasing may generally change the category of corresponding lexical material. This is obvious when category change is indicated by overt morphological markers. What shall we do if categorial identity is, at first, established by syntactic environment and certain environments are canonically associated with more than one categories? My contention is that the very fact that a lexical item occurs as a categorially unambiguous form in a paraphrased sentence does not provide a sufficient criterion for the categorial status of this lexical item in an ambiguous environment in the original sentence. The argument of paraphrasing only works if we resort to certain metalinguistic principles such as: If you can establish either one single category or two different categories for a lexical form, you should favor the one-category solution.

²⁷ Cruse (1986: 74ff.) discusses exactly this problem. He asks whether or not occurrences of *open*, used attributively, predicatively, as an intransitive verb, and as a transitive verb, are lexically distinct, that is, whether or not they constitute different lexical units. He points out that occurrence in syntactically different environments is not a sufficient criterion (the attributive and predicative position are syntactically different). He votes for regular correlations between semantic and grammatical properties as a better criterion and tries to show that, according to this criterion, the transitive and intransitive uses are lexically distinct, but that the attributive and predicative uses are not. However, Cruse's own examples illustrating this point are not very convincing. At least, in the semantic field of SUBSTANCES/MATERIALS, we can observe a fairly regular correlation between semantic and grammatical properties. At any rate, the approach of syntacto-semantic alternations, as a general approach, does not make any commitments with respect to static or dynamic representation.

According to this strategy, all lexical forms, which can be associated with distributions canonically belonging to more than one category, are represented with a single category label in a static lexicon indicating the mixed behavior, for instance, MASS/COUNT for so-called 'dual nouns'. It is unusual to establish mixed categories for major lexical categories. That is, a category label like 'ADJ-N' is not used - as an alternative device for the double category solution - for representing lexical units like *adult*, *American* or *poor*.

We can consider the difference between the multiple category strategy and the mixed category strategy as a cosmetic one. In this case, the critique mentioned for the former also holds true for the latter (There are different subclasses with 'mixed behavior'). There is, however, another possibility for conceiving mixed categories. Wierzbicka (1991) gives a hint to a different interpretation of mixed categories in the MASS/COUNT domain. She advocates a fine-grained cross-classification of nouns according to their entire grammatical behavior, i.e., their compatibility with classifiers, quantifiers, determiners and NUMBER marking. The various classes resulting from the cross-classification - which constitute lexical subclasses of so-called 'dual nouns' and are, to a certain degree, comparable to the classes on Allan's countability hierarchy - are interpreted by her as extensional classes of the lexicon rather than from the perspective of a sentence-level MASS/COUNT distinction. However, this also presupposes a certain amount of generalization, especially with respect to dialectal and historical variation. A single lexical unit - *scale*, for instance - can belong to two or three different lexical classes, preferred in different dialects, generations or social groups. Groups of speakers or single speakers, however, can also mix properties of the alternative classes. In practice, it is not always easy to separate this type of mixed category from mixed categories concerning the lexicon-grammar mapping. Nevertheless, a conceptual distinction is necessary.

2.3.3.3 Strategy of Dynamic Category Change

According to this strategy, all lexical forms, which can be associated with distributions canonically belonging to more than one (syntactic/grammatical) category, are represented with a single category label.

This can be (a) identical to one of the (syntactic/grammatical) categories. In this case, one category is considered as the 'basic' or 'primary' category for the lexical forms in question. Several reasons can be found for selecting a category as 'basic'. The most obvious reason is the (relative) defectivity of one category. Another reason may be that certain senses associated with certain categories are generally considered as 'basic' on philosophical grounds or by the linguistic theory in question. For instance, referential uses are seen as more basic

than predicative uses; general term uses are seen as more basic than universal term uses (cf. Miller 1985: 216); spatio-temporally concrete uses are seen as more basic than abstract ones, etc. Lexical semanticians have also proposed some tests similar to ambiguity tests for identifying 'primary' vs. 'secondary' lexical units (cf. Cruse 1986). Taking our examples, *poor* could be seen as 'basically' adjectival due to the noun defectivity. For the Hungarian nouns *öreg* and *amerikai*, however, there is no way of selecting one category as more basic solely on language-specific grounds.

The category label is sometimes (b) a sort of hidden mixed category, interpreted as an unspecified initial value with respect to certain syntactic/grammatical categories. Category values are conceived of as basically changeable according to actual contexts in the sentence. That is, if an initial value is assumed, this will be left unchanged in certain contexts and will be (successively) converted in certain other contexts while the sentence is compositionally build up. If no initial value is assumed, this will be assigned in the first specified context and then successively changed. The treatment of ACTIVITY and ACCOMPLISHMENT verbs provides a good example. The relevant class of verbs could be, for instance, represented either as 'basically' ACTIVITY verbs, or as a class *x* whose members will be specified as ACTIVITY or ACCOMPLISHMENT in actual sentences (cf. strategy of mixed categories).

It is almost superfluous to mention that the discussed problems of lexical ('idiosyncratic'; but cf. footnote 13) subclasses not matching with syntactic classes also exist here.

2.3.3.4 Strategy of Ellipsis

The applicability of this strategy is more restricted than that of the other strategies mentioned above. It only comes into consideration if the category difference depends on the presence vs. absence of an element in two alternative constructions which can be seen as paraphrases of each other. Consequently, the strategy of ellipsis is a possible option for *poor*, since *the poor* can be paraphrased as *the poor men*. The same is valid for *milk* in *three milks* (vs. *three glasses/liters/kinds, etc., of milk*). Since the discussion of the extremely controversial issue of ellipsis is outside the scope of the present paper, I confine myself to the claiming that ellipsis is generally a matter of language-specific properties. Consequently, cases in which the constructions of comparison do not constitute a natural-language expression in the language in question will be ruled out as dubious candidates for the strategy of ellipsis. This concerns, for instance, an analysis which treats all nouns as basic MASS nouns in English and German and assumes an obligatory 'deletion' of classifiers, not only in cases such as *three milks*, but also in cases such as *three horses* (i.e., *three horses* < *three heads of horse(s)*) (cf. Sharvy 1978).

2.3.3.5 Strategy of Grammar Change

What I mean by grammar change is a more or less radical reorganizing of the established mapping between categories of lexicon and grammar. For the lexicon, we can imagine a new hierarchical organization into major and minor lexical categories (for instance, a categorization of adjectives and nouns as subcategories of a higher order category 'nominal'). Accordingly, one could establish a syntactic (phrasal) category which is generally permitted to take such a higher order category as head, but is restricted, in special cases, to subcategories. A still more radical change would be a top division between two types of phrasal categories corresponding to our good old NP, namely, between referentially-used and predicatively/attributively-used NPs.²⁸ Finally, the very legitimacy of a default mapping between lexicon and grammar could be called in question. This is exactly what occurs in the discussion of the MASS/COUNT distinction.

3 MASS and COUNT

3.1 English and Its Influence on the MASS/COUNT Theory

3.1.1 Alternative Approaches

Pelletier and Schubert (1989) present a very detailed discussion of alternative interpretations of the MASS/COUNT distinction which depend on:

- (a) which level is relevant for the distinction (lexical vs. phrasal) and
- (b) which entities are subject to this distinction ("expressions", "senses of expressions", or "occurrences of expressions").

²⁸ I do not want to suggest that such changes would be adequate for English, but I would emphasize that such considerations should also be possible for English. We have a sort of double standard for European and non-European languages. For non-European languages, considering such options is normal scientific routine. Unfortunately, decisions based only on 'semantics' (i.e., on translations from or into a mediator language) are also much more readily accepted. In European languages, on the other hand, we expect strong formal and semantic predictability, but leave the traditional design of mapping between lexicon and grammar basically unchanged.

The authors associate "expressions" ("expression approach") and "senses of expressions" ("sense approach") with the lexical level and "occurrences of expression" ("occurrence approach") with the phrasal level or phrasal scope (cf. (14) a./b. vs. (14) c./d.). In their view, the main difference between the two lexical approaches is related to a third parameter according to which MASS and COUNT are to be seen as 'syntactic' vs. 'semantic' features. That is, a "sense approach" which assigns syntactic classes or distributional restrictions to senses is said to not be very different from an "expression approach". The third parameter is actually seen as a cross-classifying one, also dividing non-lexical approaches. In one case, it is the noun itself, which is classified in a certain context as MASS vs. COUNT ("syntactic occurrence approach") (cf. (14) c.); in the other case, it is the entire phrase which bears the semantic values MASS vs. COUNT ("semantic occurrence approach") (cf. (14) d.). The authors (1989: 339-340) point out that it is not obvious why the noun itself, rather than the phrase, in a "syntactic occurrence approach", is classified at all since the syntactic clues which are used to classify it as MASS or COUNT "are exactly the ones for which the information (+count / +mass) might be syntactically useful". I would say that it is not the lexical unit 'noun' that is classified here. Rather, it is the grammatical form based on the corresponding lexical unit which is determined, in the same way as a verb form, for instance, is determined as 3rd person singular, which is a shorthand expression for saying that the verb form has the morphological exponent of 3rd person singular, and contributes to a phrase having the syntactic value 3rd person singular. In other words, to say that a noun is MASS seems to be, in this approach, a shorter version of 'the phrase having this noun as head is syntactically a MASS phrase'.

- (14) a. level: lexical - relevant entities: expressions / senses - orientation: syntactic
 b. level: lexical - relevant entities: senses - orientation: semantic
 c. level: phrasal - relevant entities: occurrences of expressions - orientation: syntactic
 d. level: phrasal - relevant entities: occurrences of expressions - orientation: semantic

Obviously, the term 'sense', connected with the "lexical sense approach", is used in Pelletier/Schubert (1989) in a rather different way from the use of 'word sense' in lexical semantics. Here, it is assumed that the grammatical behavior or the ability of being used in certain syntactic contexts is part of the lexical meaning and is reflected in all its variations, i.e., 'word senses' (cf. Morrow 1986). Therefore, even for a lexical-semantic treatment, the following requirement is too rigid: "whether a sense is MASS or COUNT is something to be discovered after examining the semantic representation of the expression in question" (cf. Pelletier/Schubert 1989: 335), rather than after an examination of the syntactic distribution. Consequently, we come back to the main question: Is the MASS/COUNT distinction a lexical matter at all? If so, then the semantics can also be discovered through 'syntax', in one sense of this term.

In less detailed works than that of Pelletier/Schubert (1989), the difference between alternative approaches is reduced to a difference between 'syntactic' vs. 'semantic' ones. Sometimes, a distinction is made between three possible approaches, namely, between the 'ontological', the 'semantic' and the 'contextual' or 'syntactic' approach (cf. Gathercole 1985). Since these approaches are often considered as incompatible with one another (McPherson 1991; Soja et al. 1991; Levy 1988), and since the term 'syntactic' is notoriously used in an ambiguous way, some remarks are in order here.

First, 'syntactic' is often used synonymously with 'distributional' (cf. Gordon 1988; see generally all psycholinguistic articles cited in the references). Second, 'syntactic' is explicitly contrasted with 'semantic'. That is, 'syntactic' is used in a way which goes back to a Post-Bloomfieldian concept of language (cf. 2.2.2.1.2), which states that 'formal' or 'distributional' analysis can only take place in terms of (semantically) uninterpreted symbols. And third, 'syntactic' is used with respect to the level of actual utterances in opposition to lexicon, is said to be (or not to be) the domain of a certain categorial distinction.

One might wonder if it makes sense to contrast and test, for instance, the 'semantic' and 'ontological' views together against the 'contextual' view as long as these are characterized in the following way: The first is said to attribute the MASS/COUNT distinction to the systematic semantic content of terms (individuation, for instance), the second to the nature of their referents, whereas the third is said to locate the MASS/COUNT distinction in the use of the terms and to have "little to do with meaning" (cf. Gathercole 1985). The contrast presented here is obviously the result of intermixing the different senses or views on 'syntax' and 'semantics'. The attribute 'contextual' focuses on the third view on 'syntax', 'little to do with meaning' on the second.

Moreover, the relevance of distributional correlates is not contradictory to the 'semantic view'. Rather, the opposite is the case. All proponents of a 'semantic view' proceed from a strong correlation between distributional factors and language-specific semantics. They do not explore meaning independently from distributional facts but, rather, base their work on distributional facts by interpreting them. In this sense, the semantic view is neutral with respect to the level issue, i.e., a semantic interpretation can be assigned to classes in the lexicon, as done by Wierzbicka (1991),²⁹ or to syntactic configurations, as done by Mufwene (1984), or to both, as done by Langacker (1987a, 1987b). In other words, to say that apparently idiosyncratic classes or morphosyntactic rules are entirely semantic in their effect,

²⁹ This is not to say that Wierzbicka would reject the phrasal relevance of the MASS/COUNT distinction. But in the cited article, she deals with the question of semantic interpretability of seemingly 'idiosyncratic' lexical subclasses.

forcing a specific semantic interpretation, is not to deny the existence of 'formal restrictions' or, what is more important, the interrelatedness of 'formal means'.

Not even the 'semantic' and 'ontological' views are incompatible. Rather, they are interested in different questions. The semantic view is interested in the investigation or representation of systematic lexical or grammatical principles triggering "the way of talking about things" (cf. Bunt 1985: 598; Gillon 1992: 598; Langacker 1987a, 1987b) and the interpretation of grammatical behavior for adult speakers/hearers in single languages. The ontological view is interested in the question as to how it comes about³⁰ that certain lexical and grammatical principles evolve which are not entirely independent of properties of extralinguistic entities. Are certain ontological differences (cumulativity, divisivity, etc.) conceptually or perceptually so salient that they will be indirectly reflected, with a high probability, in the lexical/grammatical apparatus of languages? Are such ontological differences acquired prior to the acquisition of distributional privileges and language-specific semantics? Is there any likelihood of lexically and grammatically integrating new 'content words' according to ontological differences of extralinguistic entities which constitute potential referents for them?

Thus, it comes as no surprise that studies in language acquisition show 'contradictory' results, either supporting the 'syntactic' (i.e., distributional) or 'syntax-first' hypothesis (cf. Gordon 1985, 1988; Gathercole 1985) or the 'ontological' hypothesis (cf. McPherson 1991; Soja et al. 1991). Tests for proving the 'syntactic hypothesis' address the question as to whether or not certain implicational relations between grammatical means like quantifiers and NUMBER markers exist or, alternatively, whether such implicational relations are ignored in favor of ontological cues with the result that nouns are miscategorized - compared to the categories of adults -, as seen in **a furniture*, or that components of different categories (for instance, *much* and plural) are mixed. Gordon's results show that children of a certain age (3-4 years) make use of implicational relations, that is, they have a rudimentary paradigm in the domain of quantification/determination, and Lewy's results show that this paradigm is acquired successively. This is very consistent with findings showing analogy effects based on ontological similarity in grammatically neutral contexts (cf. also Gordon 1988). So far, no convincing evidence has been found to support the assumption that children would acquire language-specific 'MASS/COUNT syntax' independently of language-specific 'MASS/COUNT semantics', in contrast to a 'parallel-learning' hypothesis, as commonly associated with Quine (1960) (cf. Soja et al. 1991). The 'ontology-first' hypothesis requires experiments with

³⁰ Newmeyer (1994) associates these two perspectives on language (How is it? vs. How does it come about that something has such and such properties?) with formal and functional linguistics, respectively. I do not believe that this estimation is entirely correct, but I agree with him that these two perspectives must be distinguished.

children without knowledge of 'MASS/COUNT syntax'. There are two notorious problems with this sort of experiments. First, the produced data are always interpreted on the basis of adult syntax, that is, largely on the basis of Standard English syntax. Second, the absence of an adultlike use of overt markers (e.g., the distinction between *much* and *many*) does not imply the absence of knowledge of 'MASS/COUNT syntax'. An important aspect of this knowledge includes knowing how compositionality works, i.e., how lexical units and classifiers are composed into phrases to be used referentially.

There seem to be no good reasons for the assumption that children acquire 'MASS/COUNT syntax' if 'syntax' is understood only in terms of operations of uninterpreted symbols (cf. the second sense of 'syntax' above). This would be in sharp contrast both to common linguistic analyses and to the linguistic behavior of adults, which always rely, to a certain extent, on lexical-semantic knowledge and presuppose a semantic interpretation of the sentence in different dimensions (for instance, as generic or not). It is therefore doubtful whether 'syntactic' approaches, in this strong sense of the term, are at all feasible.

3.1.2 Evaluation of Linguistic Argumentation

The general impression in studying the literature is that all linguists consider the same criteria, but draw different conclusions. For this reason, I have examined 19 articles (15 in English and 4 in German)³¹ for the following questions:

- Presentation of the first examples: Are they presented as prototypical examples of a general MASS/COUNT distinction or as examples which follow from the rules mentioned before? If prototypical examples are given, which ones?
- Presentation of criteria for distinguishing between MASS and COUNT nouns: Do the restrictions which are given apply to lexical units or to configurations of grammatical means? Which of the common criteria are mentioned? Which of the common criteria are deliberately excluded? Are phonologically different grammatical forms (for instance, *some*) - as potential distinguishing criteria - collapsed into one 'homophonic' form or not?
- Terminological consistency: Is a consequent distinction made between the terms 'form-class'/'lexical class' vs. 'syntactic category'/'grammatical category' and between denote/denotation vs. refer/reference? Is a consequent distinction made between 'noun' as a lexical unit and 'noun' as an inflected form, or more generally, as a grammatical form?

³¹ In the present context, the language of argumentation and the language of the illustrative examples were relevant, rather than the object language. For the 15 articles written in English, the two were identical.

- Arguments for the 'ungrammaticality' of certain 'forms': Are 'certain forms' which violate MASS/COUNT restrictions judged as 'ungrammatical'. On which assumption are statements about 'ungrammaticality' founded?
- Arguments for the 'basic' vs. 'secondary'/'extended' status of a specific sense or dimension: Are there any criteria explicitly mentioned for the assumed status of senses or dimensions?

The following articles from the fields of psycholinguistics, general linguistics and formal semantics have been considered: Gathercole (1985), Gordon (1985), Gordon (1988), Levy (1987), Shapiro et al. (1989), McPherson (1991), Soja et al. (1991), Akiyama/Wilcox (1993), Carroll (1978), Allan (1980), Birkenmaier (1980), Harweg (1987), ter Meulen (1984), Bunt (1985), Lønning (1987), Pelletier/Schubert (1989), Link (1991), Krifka (1991), Gillon (1992).

3.1.2.1 Presentation of the First Examples

In 13 articles, 'prototypical examples' were introduced, together with a set of mutual restrictions ("English common nouns are, in general, divided into two categories: COUNT nouns such as ..., and MASS nouns such as ..." (cf. McPherson 1991: 316)). The two lists (15) and (16) contain the examples given, with the number of tokens in parentheses.

(15) Prototypical examples for COUNT nouns

car (3), *book* (2), *table* (2), *apple* (2), *man* (2), *chair* (1), *shoe* (1), *shirt* (1), *house* (1), *suitcase* (1), *beach* (1), *lake* (1) *canal* (1), *grain* (1), *symphony* (1), *condition* (1); *Pferd* (1), *Apfel* (1), *Ring* (1), *Mensch* (1), *Haus* (1), *Tisch* (1), *Aufgabe* (1), *Menge* (1), *Sonate* (1)

(16) Prototypical examples for MASS nouns

sand (6), *water* (6), *milk* (1), *gold* (1), *cheese* (1), *rice* (1), *luggage* (1), *music* (1), *lightning* (1), *courage* (1), *information* (1); *Gold* (3), *Wasser* (2), *Milch* (1), *Wein* (1), *Atommüll* (1), *Schmuck* (1), *Vieh* (1), *Polizei* (1), *Solidarität* (1), *Musik* (1)

The high number of tokens for certain MASS examples (*sand*, *water*, *gold/Gold*) reflects, to some degree, the small lexical range underlying considerations about MASS nouns in the literature. Interestingly, the two lists also contain examples which are very popular examples for dual uses: *apple* (2)/*Apfel* (1) (here COUNT) and *cheese* (1) (here MASS). Similarly, the MASS characterization of *Musik* ('music') is only valid under the standard criteria for MASS nouns, as long as one of the senses is estimated as primary and the other as extended (cf. *Das ist eine tolle Musik.* ('This is great music.')).

3.1.2.2 Presentation of Criteria for Distinguishing Between MASS and COUNT Nouns

As expected, two opposite ways of presenting criteria could be found. First, criteria are presented as defining two lexical classes, for instance, 'a MASS noun can occur with (in) x/take x/may be modified by x, etc.' Typically, such statements differ linguistically in using adverbs like *normally*, *only*, etc., indicating whether or not the criterion in question is to be seen as a necessary criterion. Second, criteria are presented from the point of view of the grammar as context frames, for instance 'preceded by x a noun will be a MASS noun/a MASS term' or 'x modifies MASS nouns', etc. The criteria mentioned in the evaluated articles are shown in Table 2 and Table 3 (the number in parentheses stands for the occurrence of a criterion). The criteria are organized in five groups and distinguished according to their application in a lexicon-oriented vs. grammar-oriented approach.

Table 2 + Table 3

Typically, grammar-oriented approaches make statements about configurations of determiners/quantifiers and NUMBER specification, rather than about restrictions of NUMBER values (SG vs. PL). That is, 'plural' as an independent criterion is mainly used only in lexicon-oriented approaches. My sample contains only one criterial mention of 'plural', where 'plural' alone is said to select COUNT nouns in an unspecified phrasal environment (cf. Allan's (1980: 542) second rule). There are cases in grammar-oriented approaches where a determiner/quantifier without NUMBER restriction is characterized as a sufficient criterion for specifying nouns as MASS vs. COUNT (see Table 2 and Table 3, where the mention of NUMBER restriction is explicitly indicated), but the reverse is almost non-existent, i.e., lexical classes are almost never restricted by a determiner/quantifier and a NUMBER value together (cf. however Carroll 1978). The fact that determiners/quantifiers and NUMBER specification are used together for context definitions points to the concept of macroparadigm as explained in 2.2.1.

For a paradigm-based representation, it is necessary to take all distinct features of grammatical means which indicate distinct values in the paradigm into account, for instance, the phonological difference between stressed and unstressed *some*, the difference between *a little* as a quantifier and as a concatenation of a determiner and an adjective, etc.³² According to the usual representation of paradigms, exponents of paradigmatic values are not explicitly

³² This is not necessary in a lexicon-oriented approach as long as further criteria are available, which make exactly the same distinction between lexical units. In this case, regarding a grammatical marker as neutral means is the same as regarding it as superfluous. Note, however, that the phonological difference is also ignored in some grammar-oriented approaches (cf. Gillon 1992).

specified as lacking all other possible values. For instance, it is not explicitly stated that an affix specified as the exponent of 3rd SG is not allowed to be used for 1st SG. Negative restrictive statements, such as 'the indefinite article cannot be used with COUNT nouns' in grammar-oriented approaches, are actually the reversal of restrictive statements usually found in lexicon-oriented approaches (cf. Gillon's (1992) shift of perspective to lexicon in the case of the indefinite article). Such exclusive specification is marked in Table 2 and 3 with '-'. To say that zero article specifies a noun as MASS noun or as the head of a MASS phrase is, however, simply not the same as saying that the indefinite article cannot specify MASS nouns/heads of MASS phrases (see the remarks with respect to the 'syntactic occurrence approach' above). In paradigm-oriented representations, we would also expect that amount phrases, as complex grammatical structures containing numerals, are listed, reflecting the different compositionality structure of MASS and COUNT phrases with numerals, rather than that numerals are excluded for MASS nouns. The presentation of criteria is paradigmatically structured in the sense understood here, for instance, in Bunt (1985), who explicitly denies the possibility of lexical classification of nouns as MASS vs. COUNT and who also excludes the indefinite article as a formal criterion.³³

I will recapitulate the differences just described on the basis of the examples (17) - (21).

(17) *The water is very cold.*

- (18) a. *The car is very expensive.*
 b. *Car is the best mode of transport.*
 c. *A car is a vehicle with four wheels.*

- (19) a. *This is a very good marriage.*
 b. *I have a great admiration for Bette.*

- (20) a. *Kim put an apple into the salad.*
 b. *Kim put apple into the salad.*
 c. *Don't put so much apple into the salad.*

- (21) *Wir essen sehr viel Kartoffeln.*
 weeat very much potato-PL
 'We eat very many potatoes.'

³³ There is a not very surprising correlation between research field and the lexicon-oriented vs. grammar-oriented presentation. Psycholinguists favor lexicon-oriented presentation, formal-semanticists grammar-oriented presentation and general linguists are not unwilling to change the perspective.

Based on the criteria usually proposed, the grammar-oriented approach which denies lexical preferences or lexical built-in modes would have to assign the phrase *The water* in (17) both to the MASS and to the COUNT category, regardless of whether it is used to refer to the water in the sea or to a bottle of water, since the relevant grammatical context is a neutral one. This conclusion has actually been made explicit, for instance, by Lønning (1987: 1-2):

"However, it is not possible to distinguish between these two definite descriptions [*the red car* and *the water that John drank*, L.B.] on purely syntactic grounds. But that is as it should be since a definite description like *the water* may have two different readings, a "count" reading or a "mass" reading, even though the larger context normally removes the ambiguity."

From this, it follows that also *the car* in (18) a. would be assigned the values MASS and COUNT (Note that the linguistic context does not 'disambiguate' in (18) a.). If the absence of a determiner is chosen as a criterion, the generically-used *Car* in (18) b. would be a MASS phrase, just as *apple* in (20) b. or *much apple* in (20) c. And the other way round, if the indefinite article is seen as an indicative criterion for the COUNT value, *A car* in (18) c. would be in the same way a COUNT phrase as *a very good marriage* and *a great admiration* in (19) and *an apple* in (20) a. The German phrase *viel Kartoffeln*, which contains a MASS quantifier similar to *much* and a word-form in the plural (cf. (21)), could be analyzed in three different ways, depending on the status of plural as a criterion. If all determiners/quantifiers are restricted by NUMBER values, it should turn out to be a 'wrong form' or should receive a new value. If determiners/quantifiers are considered as sufficient criteria without NUMBER specification, it would be a perfect MASS phrase. And if determiners/quantifiers and NUMBER are considered independently it would also be a 'wrong form', or, alternatively, a member of a lexical class with a special MASS/COUNT behavior. On the other hand, lexicon-oriented approaches, in all cases except (17), (18) a., (18) c., (20) a., should go back to one of the strategies mentioned in 2.3 (i.e., strategy of multiple categories, mixed categories, etc.).

3.1.2.3 Nouns as Lexical Units vs. Nouns as Grammatical Forms

A consistent terminological distinction between nouns as units of the lexicon and nouns as inflected forms, or more generally, as grammatical forms, is almost never made in the literature on the MASS/COUNT distinction. Just like the term 'word', 'noun' is used ambiguously in a systematic fashion. The term 'noun' is either related to an element of the lexicon (lexeme, morpheme, lexical unit or whatever) or to a grammatical form (often used as a synonym of 'term'). We can even find quotations where 'noun' is explicitly used as *pars pro toto* for the whole noun phrase (Note the italics!):

"For the purpose imagine *a chicken* vs. *chicken*. The former, a count noun, denotes a bird and the latter, a mass noun, is meat."

(Akiyama/Wilcox 1993: 420)

This sort of mismatch is very problematic when the MASS/COUNT distinction cross-cuts the distinction between inflection and derivation. This is the case if the use of a word-form is conventionalized in a specific context, resulting in a new lexical form with a defective paradigm. Systematic lexical alternations (for instance, between a distributive and a collective sense or between a TYPE and a TOKEN sense (cf. 3.1.3) correlating with different grammatical forms may have this effect. When reading the following quotation (Gillon 1992: 613, footnote 11), we do not really understand what is ambiguous and on which level:

"Clearly, some plural mass noun have singular counterparts; but these counterparts have a sense different from that of the plural. Thus, there are words "brain" and "brains". But "brains" is ambiguous between the plural of the singular count noun "brain" and the plural mass noun, an ambiguity found in the following sentence: (1) These animals have brains."

The expression "plural mass noun" could stand for a grammatical form (word-form) actually used in a PLURAL/MASS context or for a lexical unit, which is restricted to a PLURAL/MASS context. The lexical-unit interpretation is likely in the case of "singular count noun", since only this can have a plural ("the plural of the singular count noun" (?)). If so, however, the second mention of "plural mass noun" can only be understood as a grammatical form, i.e., as a plural form of a lexical unit, which is restricted to plural uses since ambiguity is only possible between lexical units or between grammatical forms but not between a lexical unit and a grammatical form.

3.1.2.4 'Form-Classes/Lexical Classes' vs. 'Syntactic/Grammatical Categories', Denote/Denotation vs. Refer/Reference

If the terms 'form-class' or 'lexical class' are used at all, they are generally not clearly distinguished from the terms 'syntactic category' or 'grammatical category'. This corresponds to Bloomfield's understanding of 'large form-classes' building 'grammatical categories'. As argued in 2.2.2.1.1 with respect to Bloomfield's approach, the transition from overlapping form-classes to grammatical patterns is not entirely clear. Sometimes, the idiosyncratic property of lexical classes with different behavior is foregrounded, and sometimes, the systematicity of a grammatical pattern, i.e., the MASS/COUNT distinction, is seen as a

grammatical category of individuation along the same line as the grammatical category of definiteness.

Strongly connected with the lack of a consequential terminological distinction between nouns as lexical units and nouns as grammatical forms is the fact that the terms 'denote'/'denotation' and 'refer'/'reference' are very often used interchangeably. That is, a terminological difference which reflects the following facts, though often required (for instance, by Lyons (1977; 1991)), is not drawn. Elements of the lexicon have a denotational range. They also have a referential potential insofar as it can be specified in which way they can be used referentially and non-referentially when occurring in actual utterances in a grammatical form and combined with other grammatical elements. But as such, elements of the lexicon have no referents. Only phrasal expressions may be said to have referents. However, the difference found in *a piece of footwear* vs. *a shoe* or in *two glasses of milk* vs. *two milks* (in one reading) does not lie in the reference, but, rather, is a difference in compositionality between lexical units and phrases.

The terms 'denote'/'denotation' are used in relation (a) to class-meanings, (b) to the meaning of a member of a class, and (c) to the ontological properties of referential phrases. The same holds true for the terms 'refer'/'reference'.³⁴ One might argue that 'refer'/'reference', in connection with lexical elements, is to be understood as referential potential. The problem, however, lies exactly here, as long as predominantly referential uses are equated with lexical class meaning and, accordingly, with the lexical meaning of single members of that class as well. Take, for instance, *adult* and *American* (cf. 2.3.1). The predominant use of English nouns in referential expressions coins their lexical class-meaning, which is supposedly inherited to single nouns. Nevertheless, it is doubtful whether *adult* classified as noun differs with respect to denotation from *adult* classified as adjective.

³⁴ It should be emphasized that this terminological indifference has nothing to do with the question of whether an author is committed to an 'ontological' or 'extralinguistic' explanation or to an 'internal' explanation in terms of language-specific syntactic/semantic properties. Gordon, for instance, advocates, very convincingly, the second alternative. Nevertheless, he uses 'denote'/'denotation' in all senses mentioned, and interchangeably, with 'refer'/'reference' and 'name'/'naming':

"Mass nouns, on the other hand, do not denote individuals when quantified (e.g. *some water*, *much sand*) (cf. 1985: 211)"

"The most notable exceptions are superordinate terms such as *furniture*, *silverware* and *jewelry* which denote classes of discrete objects and yet are mass nouns (cf. **a furniture*, **many silverwares*). Similarly, nouns such as *liquid*, *solution* and *substance* all denote non-discrete substances but they are (or can be) count nouns (cf. *a substance*, *many liquids*). (cf. 1988: 110)."

3.1.2.5 Arguments for the 'Ungrammaticality' of Certain 'Forms'

Though all authors deal with the fact that a number of lexical units can be used equally well in a MASS and COUNT context, we are again and again confronted with the following types of arguments:

- (a) MASS nouns cannot be used in plural, without an indefinite article, with numerals, etc. (cf. the criteria in Table 2 and Table 3). *Sand, water, etc.*, are MASS nouns, hence **a sand, *several sands, *two waters* are ungrammatical (cf. Shapiro et al. 1989; Krifka 1991: 400).
- (b) MASS nouns do not denote individuals, are not individuated, are not used for discretely perceived objects, etc. Consequently they may not be pluralized, individuated and counted, etc., as the ungrammaticality of **a sand, *several sands, *two waters* show. This can be expressed in distributional terms as follows: ... (cf. Gordon 1985: 211)

It is well-known, of course, that there are 'exceptions'. A number of lexical forms which satisfy MASS criteria, for instance, the ability to co-occur with the quantifier *much*, also occur with the indefinite article. A number of lexical forms which satisfy COUNT criteria, for instance, the ability to co-occur with the quantifier *several*, also nonetheless occur in bare phrases.

There are two basic problems with the above arguments.

First, the standard distributional criteria have been primarily established with respect to the dimension of spatio-temporally concrete situations, i.e., to situations which correspond directly to experience. Generic uses have been taken into account only to the extent that the compositional regularities for building phrasal expressions harmonize with those in the non-generic domain.³⁵ Uses in which expressions refer to SORTS as a taxonomy property or predicate SORTS have not been taken into account in the distributional criteria (see, for instance, *Italian wines*).

³⁵ Mufwene (1984: 213-214) assumes a total parallelism of MASS/COUNT interpretation (individuation and non-individuation in his terminology) between generic and non-generic uses. This should explain why, for example, a bare phrase with a COUNT noun like *lion* is not permitted as a generic subject (**Lion is a carnivore.*). A bare phrase would imply non-individuation and this is not allowed for generically-used COUNT nouns, just as for non-generically ones. *Car*, in (18) b., is a counterexample and the individuation hypothesis of generics is not without doubt. And Mufwene himself presents some counterexamples without explaining them.

The semantic characterization in argument (b) 'they may not be pluralized, individuated and counted' addresses exactly this 'primary dimension', assuming that there are exact morphosyntactic correlates to 'pluralization', 'individuation' and 'counting'. If one takes individuation and counting as semantic specification carried out on the phrasal level, the restriction to COUNT nouns does not hold.³⁶ In addition, if pluralization is semantically understood as distributive interpretation of multiplicity as opposed to a collective interpretation, the biunique association with the morpheme '-s' is also incorrect (cf. lexical forms restricted to plural and connected with MASS quantifiers, i.e., 'plural MASS nouns').

For a lexical classification, however, the following are the relevant factors:

- Is plural morphologically possible or not?
Is the morphological plural form conventionalized as the only inflectional form?
- Are, for an individuated interpretation as SORT, PORTION, SINGLE PIECE or INSTANCE, measure terms or classifiers obligatory or not, unless neutral determiners in singular are used (for instance, *the*)?
In other words: Is there a synonymy relation between constructions such as *two milks* and *two glasses of milk* or **an information* and *a piece of information*? Whereas *two milks* is possibly used instead of *two glasses of milk* in certain contexts, one cannot say **an information* instead of *a piece of information*.

Countability preferences of the kind proposed by Allan (1980) account, among other things, for such lexical differences (cf. the differences between *mankind*, *admiration*, *equipment*).

The second problem concerns the fact that the standard distributional criteria are too rough. Different syntactic and referential uses are only implicitly, rather than explicitly, taken into account in the set of standard distributional criteria. The statement that a bare phrase for the lexical form *car* is ungrammatical would be partly true and partly false, depending on where and how *car* is used. The fact that *car* may constitute the head of a bare phrase if used as a generic subject (as in (18) b.), or as a generic object, does not imply that it can be used without an article in the predicative position:

(22) **This is car.* / **This is beautiful car.*

³⁶ Sometimes, it is argued that counting with respect to amorphous, non-discrete, non-solid, etc., substances cognitively differs from counting of discrete entities (cf. Krifka 1991). But this is not the issue here. This would probably 'explain' why languages make a distinction between lexical forms used obligatorily with classifiers in certain contexts and lexical forms used without classifier in the same contexts, that is, why English appears to us more natural than, for instance, Vietnamese.

For good COUNT candidates, it is the generic use that is implicitly or deliberately disregarded in distributional analyses (cf. Allan 1980) because it would disturb some regularities in the non-generic domain, similarly to the treatment of the SORT use of MASS candidates.

All of this indicates that strings, judged as ungrammatical, are of the special type characterized by Householder (1973: 369) for sentences:

"A special oddity is the asterisked sentence which is said to be ungrammatical (or unacceptable) only in a particular sense, while quite O.K., in other senses. Here (as sometimes elsewhere) we are judging a structure-meaning pair, instead of just a structure."

From this follows that a systematic analysis should take into account from the outset (a) the entire set of semantic interpretations which are allowed in a certain distributional configuration and (b) the entire set of distributional configurations which are allowed for a certain semantic interpretation.

There is a long tradition in linguistics to link distributional techniques to 'meaning shift'/'meaning change'. That is, if a certain distributional test does not show the expected ungrammaticality of a construction, it is assumed that a 'meaning shift' is involved.³⁷ There is a subtle difference between this practice and the requirement to take all possible semantic interpretations in a distribution into account: Only the former, but not the latter, is forced to choose one meaning as 'primary' or 'basic' which can then be said to be 'shifted' (cf. also the remarks on the strategy of dynamic category change in 2.3.3.3).

Take, for instance, *Italian wines*. It is often said that in such phrases 'the meaning has been changed'. Strictly speaking, this implies that there is one lexical meaning which is selected as default sense and then changed in a specific grammatical context.³⁸ For, if more than one conventionalized senses of equal status or systematic sense alternations are assumed, that

³⁷ See Bußmann 1990: 222.

³⁸ There is another possibility of how 'meaning change' can be interpreted. Instead of the change of the lexical meaning, the non-occurring phrasal meaning which is expected by the grammatical context is assumed to 'change'. The following quotation is probably to be understood in this sense:

"Außerdem haben nur Dingwörter einen Plural, nicht aber Stoffbezeichnungen. Wenn letztere in der Form des Plural erscheinen, ändert sich ihre Bedeutung. In Sätzen wie *wir hatten zwei Bier und vier Brote* sind "die Stoffnamen zu Dingnamen geworden, denn wir meinen eigentlich 'zwei Glas Bier, vier Stück Brot'" (cited after A.V. Isaenko, *Die Russische Sprache der Gegenwart. Teil I. Formenlehre*. München: Hueber, 1968, p.75. Birkenmaier 1980: 32)"

sense which fits in the respective grammatical context would be selected. According to the 'meaning-change' view, *wine* is supposed not to be subject to a lexically established (systematic) 'UNIT-SORT'-ambiguity ("unit-type ambiguity" in the terminology of Cruse (1986)). As a consequence of this, SORT interpretations in singular contexts and SORT interpretations of COUNT candidates should also be treated as instances of 'meaning change', as shown in (23) a. and b.

- (23) a. *This wine is really excellent.*
 (> 'meaning change' in the SORT interpretation)
 b. *My sister has the skirt Sue is wearing now.*³⁹

But this solution is counterintuitive. Even more problematic is the practice of supposing 'meaning change' in the case of generic uses if a test fails to show the expected ungrammaticality of a construction (cf. (18) b.). Here, there is a consensus that the ability to be used both generically and non-generically is a general inherent property of the entire lexical category of nouns, rather than that one of the uses (the non-generic one) is lexically preestablished whereas the other (the generic one) applies only on the phrasal level.

3.1.2.6 Arguments for the 'Basic' vs. 'Secondary'/'Extended' Status of a Specific Sense or Dimension

One of the conclusions drawn from the frequent dual uses of English lexical units in MASS and COUNT contexts is, as mentioned above, that the MASS/COUNT distinction has nothing to do with the lexicon. I have argued that this conclusion is unjustified insofar as it suggests that there is no lexical built-in mode at all which would determine distributional restrictions and the range of interpretations. Nevertheless, the very existence of dual uses and 'counterexamples' forces us to look for explicit criteria determining the primacy of certain senses or dimensions.

Explicit arguments for judging one sense or one corresponding dimension as 'basic' are relatively rare in the literature. In some cases, the decision is guided by the general philosophy of a model, as, for instance, in Cognitive Grammar. Langacker (1987a: 67) explicitly points out that the distinction made in Cognitive Grammar between 'basic domain' such as the domain of three-dimensional space and 'non-basic domains' such as the domain of quality space is directly relevant to the derived status of the COUNT category of *a fine wine* on the phrase level. The decision does not seem to be so easy for dually-used lexical forms such

³⁹ Cf. Cruse 1986: 63.

as *brick, stone, diamond, cake* which are all assigned dual category status on the lexical level by Langacker (cf. 1987a: 66). No clear statement is made about 'abstract nouns'. On the one hand, nominalizations are said to inherit, as usually assumed (cf. Mourelatos 1981), the aspectual properties of the corresponding verbs, ACHIEVEMENT verbs resulting in COUNT nouns and STATE verbs resulting in MASS nouns. On the other hand, we are told that there are "other patterns of nominalizations" by which STATE verbs may result in (lexical) COUNT nouns, for instance, in the case of *hope* or *belief*.

Usually, the assumptions which guide the classification of a lexical form with dual behavior as 'real flexible noun', 'non-prototypical MASS noun', and 'non-prototypical COUNT noun' are ontologically founded and not always made explicit (cf. Gathercole 1985: 402-403; Levy 1988: 185-186).⁴⁰ Different sorts of background knowledge jointly contribute to a decision about classification here: Knowledge about the characteristic shape of the entities, to which noun phrases (containing the noun in question as head) may, among other things, refer; knowledge about further properties of these entities which are partly identical to the standard properties in logical semantics (for instance, divisibility); and, lastly, knowledge about differences between ontological categories, i.e., between natural kinds and human artifacts, etc.

I have made a list of the most common types of examples (cf. (24) - (28)) cited as evidence for the non-lexical character of the MASS/COUNT distinction and/or for the unitary semantic principle carried out on the phrasal level.

(24)

- *beer, wine, coffee*
 - a. *some beer, a good beer, three beers* (cf. (2), (23))

(25)

- *lamb* ('animal' + 'the meat of a lamb USED AS FOOD'), *fish* ('animal' + 'the flesh of fish USED AS FOOD'), *chicken* ('animal' + 'the meat of a chicken USED AS FOOD'), *dog* ('animal' + 'the meat of a dog USED AS FOOD')
- < ANIMAL + MEAT OF ANIMAL USED AS FOOD
- a. *My favorite meat is lamb.*
 - b. *This is a lamb.* (i.e., 'an animal') / *This is lamb.* (i.e., 'lamb meat')
 - c. *That man eats dog.* (i.e., 'dog meat')⁴¹

⁴⁰ Gathercole (1985) has classified *bread, money, chalk, ice* as 'non-prototypical MASS nouns', *rock, noodle, crayon, stick* as 'non-prototypical COUNT nouns', and *cake, glass, hamburger, fire* as 'flexible nouns'.

⁴¹ Cf. McPherson 1991: 317.

d. *This is a dog.* (i.e., 'an animal') / ?*This is dog.* (i.e., 'dog meat')

- *cake, cheese, bread, ice, apple, salad*
e. *some cake, a cake, three cakes* (cf. (20))
- *glass* ('something **MADE OF** glass' ('drinking vessel', etc.), *iron* ('implement **MADE OF** iron'), *silver* (cf. 'something **MADE OF** silver' ('coin', 'household articles', etc.) (cf. (10))
< **SUBSTANCE/MATERIAL + ARTIFACTS MADE OF SUBSTANCE/MATERIAL**
f. *The window is made of glass rather than of plastic.*
g. *We have only five glasses.*
- *stone, crystal, brick, diamond, paper*
< **SUBSTANCE/MATERIAL + PIECE OF SUBSTANCE/MATERIAL AS NATURAL/USUAL OCCURRENCE**
h. *The artist works with stone.*
i. *We have found very beautiful stones.*

(26)

- *beauty* ('property' + 'a girl or woman having this property')
< **PROPERTY + PERSON HAVING PROPERTY**
a. *There is much beauty in Montréal.*⁴²
b. *There are many beauties in Montréal.* (i.e., 'many beautiful women')
- *experience(s), pain(s), effort(s), time(s)*
c. *I feel pain.*
d. *I have a pain in my stomach.*
e. *She takes pains in cooking a good meal.*
f. *Past experience has shown that life is difficult.*
g. *I had a strange experience.*
h. *Corruption has occurred in many business experiences.*
- *theory, grammar, meaning*
i. *John likes theory.*⁴³
j. *Do you like the theories presented in the workshop?*

(27)

⁴² Cf. McPherson 1991: 317.

⁴³ Cf. Mufwene 1984: 202.

- *school, church, jail*
 < SOCIAL INSTITUTION + BUILDING FOR SOCIAL INSTITUTION, ETC.
 a. *Mary goes to school.*
 b. *Mary is going to the school* (i.e., to a specific school).
 c. *Harry is in church / jail.*⁴⁴
 d. *Harry is in the church / jail* (i.e., in a specific church / jail).

- *guitar, piano, train, bus, plane, hand*
 e. *My friend plays guitar.*
 f. *My friend has bought a new guitar.*
 g. *My friend went by train / bus / plane to Chicago.*
 h. *My friend took the train / bus / plane.*

- *cat, worm*
 i. *This room smells of cat.*⁴⁵
 j. *He is collecting worm for his fishing.*

(28)

- *car*
 a. *That '59 Chevy is a lot of car.*⁴⁶ (i.e., 'a lot of car-matter')
 b. *That '59 Chevy is a big car.*

It is worth noting that the MASS/COUNT alternation goes hand in hand in some of the cases with systematic sense relations which, in turn, can be seen (and have actually been seen) as independent instances of systematic polysemy or metonymy: ANIMAL + ANIMAL USED AS FOOD (cf. *lamb*, etc., in (25)), SUBSTANCE/MATERIAL + ARTIFACTS MADE OF SUBSTANCE/MATERIAL (cf. *glass*, etc., in (25)), PROPERTY + PERSON HAVING PROPERTY (cf. *beauty*, etc., in (26)), SOCIAL INSTITUTION + BUILDING FOR SOCIAL INSTITUTION (cf. *school*, etc., in (27)).

These systematic sense alternations are clearly distinct from the MASS/COUNT alternation in that they can be definitely observed in neutral MASS/COUNT contexts as well (see, for instance, the following examples: *The university has decided that...* vs. *The university is in Guildford.*). For this reason, it is the lexical context that is commonly assumed to 'disambiguate' or select the appropriate senses, rather than the grammatical context (cf. Bierwisch 1983). Typically,

⁴⁴ Cf. Mufwene 1984: 202.

⁴⁵ Cf. Brown/Miller 1980: 242.

⁴⁶ Cf. McPherson 1991: 317.

such sense alternations provide a productive device for extending vocabulary and are also found in languages which lack a grammatical paradigm for MASS/COUNT distinction comparable to the English paradigm (in Hungarian, for instance). Nevertheless, the MASS/COUNT syntax has a side effect in English. Some systematic metonymies correlate not only with distinct lexical environments but also with MASS or COUNT contexts and thus, are syntacto-semantic alternations in the same way as the transitivity alternations in the verbal domain are (cf. Levin 1993). See examples (27) a. and b.

In the case of the sense alternation between SUBSTANCE/MATERIAL + PIECE OF SUBSTANCE/MATERIAL, it is less clear whether we have to deal with two separate alternations or, whether simply the difference between MASS and COUNT environment in combination with the syntactic difference between referential and attributive positions is exploited for this sense alternation. The latter seems to be the case in the second 'FOOD field' under (25) (*cake, cheese, bread, etc.*). Accordingly, it is impossible to decide on a non-linguistic basis, whether *bread*, for instance, is a 'non-prototypical COUNT noun' or a 'non-prototypical MASS noun', or neither of the two. All examples under (25), and only those, exhibit so-called 'grinder effects'⁴⁷, giving the impression that the phrasal interpretation would shift between the perception of an amorphous, 'grinded' substance and a discrete object.

Similarly, it is extremely difficult to find sound language-independent criteria for the primacy of one of the two following interpretations in the domain of 'abstract nouns': (a) as a universal term vs. (b) as a (particular) instance of the same concept (cf. the groups *experience, pain, etc.*, and *theory, grammar, etc.* ((26) c. vs. d.; f. vs. g.; i. vs. j.). There are significant differences among lexical units in this domain with respect to the tendency towards conventionalization. This is considerably stronger in the field of mental SENSATION or physical SENSATION such as *experience, pain* than among lexical forms like *theory* and *grammar*. In English, and also in German, there is a further tendency to lexicalize the

⁴⁷ The metaphor of the 'Universal Grinder', first proposed by Pelletier (1975), has become very popular in linguistics. This is a hypothetical machine that may chop any object into a homogeneous mass, that is, it turns an object 'apple' into a 'mass of apple', parallel to the linguistic change from *an apple* to *apple*. This event can be thought in the reverse direction as well, i.e., as a bidirectional event. In psycholinguistic research, the question as to in which semantic fields the MASS/COUNT syntax evokes 'grinder effects' has been investigated as well. Akiyama/Wilcox (1993) have studied adults and children with respect to whether or not they use phrases based on the same lexical form if they refer to objects cut into pieces and to corresponding intact objects. Their results indicate that adults, and children as well, though to a lesser extent, are likely to use the same lexical form ('name') in 'FOOD contexts' much more than in other contexts.

INSTANCE sense, resulting in a new lexical unit restricted to plural and often to specific linguistic environments (i.e., idiomatic expressions), or to sublanguages (cf. (26) e. and g.).⁴⁸

We do not find any 'grinder effect' in the second and third group in (27) (*guitar, train, and cat, worm*), where the MASS/COUNT distinction is said to be carried out by the presence or absence of an article. The respective bare phrases occur partly (i.e., those in the second group) as the variable components of so-called 'lexical phrases', i.e., conventionalized phrases consisting of a lexically fixed part (verbal head, preposition, etc.) and a lexically empty part, which can only be filled by the members (co-hyponyms, synonyms, etc.) of a specific semantic field: *play* MUSICAL INSTRUMENT/SPORT, *be in* or *at/go to* SOCIAL INSTITUTION, *go by* VEHICLE, ETC.).⁴⁹ Typically, some of these lexical phrases correspond to lexical forms morphologically derived from nouns in other languages. Hungarian, for instance, systematically uses derived verbs being translation-equivalent to lexical phrases such as *play* MUSICAL INSTRUMENT/SPORT (*zongorázik* ('play piano'), *teniszezik* ('play tennis')).

The third group in (27) contains no lexicalized examples at all. On the contrary, (27) j. and k. belong to the few examples of dual uses cited in the literature which can be assessed as really creative uses. (27) j. differs from (18) b. (*Car is the best mode of transport*), first of all, by the fact that the attributively-used *cat* does not appear as the subject of a generic statement. The question as to whether such attributive or non-specific uses have to be seen as a sort of generic use, will be discussed later and considered particularly with respect to Hungarian data. It has been often observed, by Hopper/Thompson (1984), for instance, that there is a cross-linguistic tendency for associating features of 'non-individuation', including non-specific uses⁵⁰, with peripheral CASE marking, i.e., with locative, dative or instrumental, rather than with a direct object which signals high transitivity in the interpretation of

⁴⁸ The 'lexicalization history' of single lexical forms is, of course, much more complicated as schematically described here. Neither this fact nor the difference among groups of lexical units, referred to above, can be commented on in a detailed fashion here. Nevertheless, two points should be mentioned: First, the sense difference considered here is not identical with the well-known difference between 'event interpretation' and 'result interpretation' shared by lexical forms derived from corresponding (ACTIVITY/ACCOMPLISHMENT) verbs. Second, I do not believe that there is a language-independent way of giving priority to one of the considered senses. Searle (1969), for instance, has claimed on philosophical grounds that lexical forms of general terms like *kind* have, universally, a primary status as opposed to universals like *kindness*, which is, according to Searle, naturally reflected in the morphology of English. He has raised the hypothesis that languages could lack expressions for universal terms, but they should have expressions for general terms since only these could be used in predicative expressions. As shown by Miller (1985: 216), this hypothesis is clearly wrong.

⁴⁹ Pawley (1985: 114) used the term "lexicalized phrase" for this sort of 'multi-words'.

⁵⁰ "Non-referential" uses in the terminology of Hopper/Thompson 1984.

Hopper/Thompson and is generally placed high on a CASE hierarchy. An attributive phrase like *of cat* in (27) is even more likely to be associated with features of 'non-individuation' than peripheral CASES. By generalizing the determiner system for specific uses to non-specific uses in the domain of 'non-abstract' nouns which are commonly called COUNT nouns, English opted for a typological configuration which lends force to a grammatical MASS/COUNT distinction. This, in turn, contributes to the 'exceptional character' of a bare phrase with a singular noun-form (*of cat* or *worm*), in non-specific use, when not supported by a systematic lexical alternation or an idiomatic pattern.

(28) a. can also be regarded as an "extraordinary way of expressing a familiar idea" (cf. Pawley 1985: 113). Here, this effect is produced by the choice of a specific quantifier in a context where this quantifier is commonly not used as an alternative expression for an adjective of size. That is, for 'non-abstract nouns', a MASS quantifier such as *a lot of* or *much* is not used synonymously in a systematic fashion with a combination of an indefinite article and an adjective like *big* or *great*, as is the case with 'abstract nouns' (cf. *much admiration* / *a great admiration*; cf. (19) b.).

Those who advocate a semantic principle behind two distinct syntactic configurations called MASS and COUNT (cf. Mufwene 1984; Langacker 1987a, 1987b; Gordon 1985; McPherson 1991) probably identify all these dual uses as manifestations of the same alternation. That is, they assume that the differing syntactic distributions correspond with differing semantic perspectives which can all be subsumed under one superordinate semantic principle labeled 'individuation', 'boundedness' (cf. Langacker 1987a, 1987b), etc. Thus, changing the perspective means introduction or removal of individuation. One might wonder whether or not the 'grinder effects', i.e., effects causing the imagination of scattered, amorphous quantities, observed in some contexts, are a necessary side effect of such a general principle. Discussing the MASS/COUNT distinction in Vietnamese, Cao Xuân Hao (1988: 43) writes the following:

"There is however no natural count/mass distinction, because there is no objective, ready for use distinction between form and substance, objectness and quality or matter and massness."

The question is how these two contrasts ((a) FORM and SUBSTANCE, (b) OBJECTNESS and QUALITY) are related to each other. Are they simply variants serving the same principle of individuation and deindividuation? Does, for instance, focusing on QUALITY necessarily imply that the object, which is conceived as a bundle of properties, will be stripped of its characteristic formal properties? Or, are (a) FORM and SUBSTANCE, (b) OBJECTNESS and QUALITY located in two independent dimensions with the result that they are implicationally connected in some languages but not in others?

3.1.3 Five Dimensions in the MASS/COUNT Domain

Before proceeding to the discussion of German and Hungarian data, it is therefore necessary to disentangle the different dimensions on which the MASS/COUNT distinction can be dealt with. On the basis of empirical experience, at least the following five dimensions should be distinguished:

	DIMENSION	VALUES		
1.	referentiality	REFERENTIAL	PREDICATIVE	ATTRIBUTIVE
2.	objectness	OBJECTNESS		QUALITY
3.	spatio-temporal concreteness	TOKEN / INSTANCE		TYPE
4.	taxonomy	SORT		UNIT
5.	form	FORM / SHAPE		SUBSTANCE

Table 4

In the **dimension of referentiality**, it will be specified how lexical units of the category noun (be it a major lexical category or a subcategory) are semantically interpreted according to the position in the sentence. For this purpose, I will employ the traditional notions 'referential use', 'predicative use' and 'attributive use' which will be interpreted as abstract 'grammatical primitives' of meaning-form association (cf. Miller 1985). 'Referential use' is understood here in the narrow interpretation of the term which excludes, in English, nouns and adjectives after the copula (i.e., in ascriptive sentences, but not in identity-stating sentences), nouns as modifiers in compounds, and all other cases where a non-specific and non-generic 'reference' (i.e., a 'non-reference') is made (cf. Quine 1960; Lyons 1977).⁵¹

The distinction between OBJECTNESS and QUALITY will be made in the **dimension of objectness**. That is, it will be specified here whether, by using a lexical unit in the sentence (in a referential position or in a non-referential position), the properties of entities are foregrounded without any commitment to the individuality/objectness of their bearers, or whether the individuals themselves, which are associated with specific properties, are foregrounded. In 2.3.1, it was demonstrated that lexical forms can be unbiased with respect to this difference, which often leads to a double categorization of such forms as both adjectives and nouns in English, due to the poor morphological/syntactic distinctness of these categories.

⁵¹ I do not, in principle, reject the application of the term 'reference' to all appearances of 'content words'. This, however, presupposes a quite different interpretation of 'reference', whose adequacy is not our concern here.

The distinction between OBJECTNESS and QUALITY grasps an important aspect of the feature 'individuation' as used in the MASS/COUNT discussion. Individuation is, however, also commonly connected to the issue of genericity.

It is the **dimension of spatio-temporal concreteness** where the issue of genericity will be primarily addressed. I will not be concerned here with the highly complicated and controversial question of what genericity and generically-referring phrases are and what types of different generic statements should be distinguished (cf. Lyons 1977: 193ff.). Two related aspects of the concept of genericity, however, have to be pointed out. Traditionally, generic noun phrases are exemplified by sentences in which a generic assertion is made about the subject. Linguistic factors such as TENSE or ASPECT can be interpreted in a relatively easy way in these cases as cues of timeless and contextless propositions which 'say something about the class', rather than about specific or hypothetical non-specific members of the class. There is a second type of rather unproblematic generic expressions, namely, generic objects occurring with verbs such as *like*. Discussion about genericity in the literature (about non-essential generic propositions, about the difference between 'definite generics' and 'indefinite generics', etc.) is chiefly confined to these two classes. By comparison, the generic or non-generic status of noun phrases which occur in generic sentences in positions other than subject or object (for instance, *in the jungle* in *The lion lives in the jungle*), or which occur in context-sensitive and time-sensitive sentences but do not refer to specific individuals (for instance, *train* in *I go by train*), is rather unclear. Sometimes, the term 'generic' is used in a broader sense, resulting in an implicit or explicit equation of 'generic reference' with 'type reference' as opposed to 'token reference'.⁵² It is the idea behind the distinction between "type" and "token" in Jackendoff (1983/1988: 95ff) and between "type" and "instance" in Langacker (1991: 56ff.) which I would like to adopt here. Despite apparent differences, the main idea of both approaches is that lexical units are systematically related to abstract concepts and to instantiations of these concepts which correspond directly to experience and have a particular spatio-temporal location in the domain of instantiation.⁵³ In the third dimension, then, a specification will be

⁵² See, for instance, Zelinsky-Wibbelt (1991), who uses both terms ("type reference" and "generic reference") interchangeably.

⁵³ Jackendoff considers all TOKEN-phrases as referring phrases and all TYPE-phrases not directly corresponding to experience as non-referring phrases as stated under his "Referentiality Principle II" (cf. 1983/1988: 94). The consequence of this is that Jackendoff, too, considers those noun phrases which are often regarded as non-referential, namely predicatively-used noun phrases in 'ascriptive sentences' and noun phrases occurring under non-factual modality, as non-referential (i.e., having the TYPE-value). Noun phrases occurring in the subject position in generic sentences, however, are also seen as non-referential, since they are considered as TYPES. This concept of reference, situated on the third dimension, cross-classifies with a concept of reference which is a matter of grammatical organization rather than of perception or extralinguistic existence, and is situated on the

made with reference to spatio-temporal concreteness/abstractness, rather than to genericity/non-genericity in the classical sense.

The **dimension of taxonomy** concerns the difference between SORTS and UNITS. This means that a differentiation will be made between TYPES as unique concepts and SORTS as relational concepts which explicitly reflects the taxonomy relations between TYPES and SUBTYPES on a TYPE hierarchy. This results in a significant linguistic difference between TYPE uses and SORT uses. Since the latter are inherently contrasted to other SORTS on the same hierarchical level, noun phrases referring to SORTS typically occur with markers of multiplicity (i.e., 'plural' markers) and demonstratives. Note that a simultaneous reference to SORTS and to single INSTANCES can be made as well when the INSTANCES are seen as SORTS (*This wine is from California.*).

The **dimension of form** will be relevant only within the spatio-temporally concrete domain. It concerns the difference between conceiving/perceiving an INSTANCE as having a particular FORM and conceiving/perceiving it as SUBSTANCE without FORM.

Now, I will briefly demonstrate the most important interrelations between these dimensions by well-known examples with *gold* under (29) -(34). 'Classical generic' uses of *gold*, i.e., in the subject or object positions (cf. (29), (34)), correspond to the configuration of the following values: REFERENTIAL, QUALITY, TYPE. They will be distinguished from nominal predicates, as found in (30), only in the dimension of referentiality (see Krifka (1991) with respect to differences between 'general term approach', 'singular term approach' and 'dual approach'). The noun phrase *The gold* in (32) will receive the value OBJECTNESS, just as all other specifically-used occurrences of *gold*, such as *30 g of gold*. The difference between specific and non-specific uses (cf. also the terms 'wide scope' and 'narrow scope'), as commonly associated with the indefinite article in English and suggested also for uses like that in (31), will be represented in the dimension of objectness. That is, *Gold* in (31) will receive the value QUALITY in the dimension of objectness in contrast to *The gold* in (32), and the value ATTRIBUTIVE in the dimension of referentiality in contrast to *Gold* in (29)⁵⁴.

first dimension (see footnote 51). It is this dimension, in my opinion, which establishes a distinction between generic subjects as referring to 'kinds' or to names of 'kinds' and corresponding predicates as non-referring at all. Langacker, in contrast, explicitly permits a cross-classification between the reference/non-reference distinction, identified by structural position, modal operators, etc., and the TYPE/INSTANCE distinction. *A job* in *He desperately needed a job* is an example of what he would call a NON-REFERENTIAL INSTANCE use (cf. 1991: 56).

⁵⁴ In English and also in German, the assumed difference between REFERENTIAL and ATTRIBUTIVE use, as found in sentence pairs like (29) and (31), is phonologically expressed; other languages have morphological means for this difference (Finnish) or exploit word-order differences (Russian).

Furthermore, *Gold* in (31) will be specified in the dimension of spatio-temporal concreteness as an INSTANCE which can be spatio-temporally located, rather than as a TYPE like *Gold* in (29). The usefulness of a distinction in the dimension of spatio-temporal concreteness can be demonstrated with the German examples (33) and (34). These show that different pronouns are required when standing in anaphoric relation (a) to a bare phrase in TOKEN use (i.e., *es* as in (33)), and (b) to a bare phrase in a TYPE use (i.e., *welche(r/s)* as in (34)).

(29) *Gold has the periodic number 79.*

(30) *This powder is gold.*

(31) *Gold was lying in the safe.*

(32) *The gold will melt soon.*

(33) *Anna hat **Gold**_i verloren und Otto hat *es*_i / **welches*_i wiedergefunden.*⁵⁵

A. AUX gold lost and O. AUX it found again

'Anna has lost gold and Otto has found it again.'

(34) *Weil [since] Anna **Gold**_i liebt [likes], hat sie [she] **es*_i / *welches*_i gekauft [bought].*

'Since Anna likes gold she bought some.'

3.2 German

It is commonplace in linguistics that (a) European languages and numeral classifier languages basically differ with respect to the MASS/COUNT domain and (b) closely related languages such as English and German may also differ significantly in the ways they categorize single lexical units, which can be seen as rough translations of each other (see, for instance, *information* in English, a standard example for a MASS noun, and *Information* in German, which is held to be a COUNT noun; cf. Mufwene (1984: 211-212) for the categorization of English-French translation pairs). However, Jespersen had already pointed out in 1925 that English and German also exhibit some more general differences in the domain of MASS/COUNT distinction rather than only on the level of single lexical units. He mentioned, for instance, that the use of 'abstract nouns' "for a single act or instance of the quality" (1925: 200) is "not so universal in English as in many other languages" which he demonstrated with German examples ("*eine unerhörte Unverschämtheit* is a piece of monstrous impudence") (cf. *ibid.*).

Nevertheless, the types of criteria used for a MASS/COUNT categorization are basically the same as those for English (cf. 3.1): Compatibility with the members of quantifier pairs exhibiting distinctive values (like *much* and *many*), compatibility with numerals, amount

⁵⁵ See Krifka (1991: 404).

phrases, indefinite article, and plural. The criteria are also used, in principle, in the same way as described for English in 3.1.2.2, with corresponding differences between lexicon-oriented vs. grammar-oriented approaches. However, we rarely encounter remarks to the effect that the respective types of criteria should be treated differently due to language-specific differences.⁵⁶

The most important linguistic facts concerning the MASS/COUNT distinction in German can be summarized as follows:

- There are no suppletive forms among quantifiers, i.e., morphologically distinct forms which can be grouped together in a paradigmatic row differing only in the value MASS vs. COUNT (*much* vs. *many*), nor are there phonologically differing pairs of quantifiers (stressed *some* vs. unstressed *some*). Rather, the difference between uninflected and inflected forms is exploited for the paradigmatic contrast between MASS and COUNT on the level of grammar. This holds true for the quantifiers *viel* ('much', 'many'), *wenig* ('little', 'not much', 'few', 'a few'), including all complex forms of these (i.e., *soviel* ('so much', 'so many'), *wieviel?* ('how much?', 'how many?')). In addition to this, there are, just as in English, (a) quantifiers which are either associated with singular forms and MASS phrases or with plural forms and COUNT phrases, and (b) neutral quantifiers (*all* ('all', 'every'), *mehr* ('more')), where the configuration of the quantifier and NUMBER value together signals the grammatical value of MASS vs. COUNT.
- The morphological formation of plural is more complicated than in English. The singular and plural morphological classes constitute partly independent subparadigms, thus, the prediction of a plural form is not always possible on the basis of the singular forms alone. The consequence of this is that pragmatic gaps are often morphologically conventionalized in the vernacular. The lexical form *Milch* ('milk'), for instance, has no plural form in everyday speech, but it exhibits two alternating forms (*Milche* / *Milchen*) in the sublanguage of the dairy branch where reference to SORTS is as salient as reference to PORTIONS.
- There is only a prosodic difference, rather than a segmental one, between the numeral *ein* ('one') (stressed) and the indefinite article *ein* ('a') (unstressed). *Ein* shows ambiguity between specific and non-specific uses like the English indefinite article. Generally, the semantic implication of the absence and presence of an indefinite article in singular phrases is less clear in German than in English. We observe the same variation with

⁵⁶ Exactly such a basic comment has been made by Bäuerle (1994: 15) with respect to the domain of verbal ASPECT. In discussing Vendler's categories, he points out that these are much more problematic for German than for English, since in German, aspectual categorization is much more a matter of interpretation than of grammaticality.

predicatively-used lexical forms in the semantic field of NATIONALITIES, as in English. *Amerikaner* in (35) is, however, clearly a noun in German on morphological grounds and the use of the indefinite article cannot be associated with different lexical categories (cf. 2.3.1). There are some further semantic fields whose members display a similar formal alternation, for instance, that of OCCUPATION (cf. (35)). The question is whether or not the use of the indefinite article correlates with the distinction between QUALITY and OBJECTNESS as stated in the dimension of objectness above. For many German natives, the two constructions are without any significant semantic difference and almost totally interchangeable. In any case, we have to deal with a lexical group here where the usual interpretation in terms of the MASS/COUNT distinction is doomed to failure. The bare phrase *Lehrer* can not be interpreted as a 'MASS phrase' in the same way as, for instance, the bare phrase *Wasser* ('water') or even the bare phrase *Fisch* ('fish') in the predicate position (*Fisch* shows the same systematic alternation between ANIMAL + MEAT OF ANIMAL USED AS FOOD as its translation (cf. (25))).

- (35) *Peter ist (ein) Amerikaner // (ein) Lehrer*⁵⁷
 P. is (an) American / American // (a) teacher
 'Peter is American / an American // a teacher.'

In the following, the German data will first be discussed in selected lexical domains: (a) in the domain of DRINKS and other LIQUIDS and (b) in three subdomains of 'abstract nouns' which will be labeled 'motivation', 'crime' and 'sensation'. Secondly, on the one hand, lexical phrases and further idiomatic expressions will be investigated, as well as, on the other hand, such productively generated phrases (verbal phrases) where the nominal containing the word-form in question can be interpreted non-specifically and constitutes a unitary concept together with the verbal head. Thirdly, the existence of so-called 'grinder effects' in German will be studied.

3.2.1 DRINKS and Co.

In German, lexical forms associated with a sense denoting DRINKS can be used (a) in bare phrases or in phrases modified by a MASS quantifier (*etwas*) (cf. (36) a.) and (b) in complex phrases headed by a 'CONTAINER phrase' (i.e., *ein Glas*) (cf. (36) c.). The acceptability of (36) b., i.e., the use of the numeral *ein* without an explicit 'CONTAINER phrase', depends on factors

⁵⁷ Some important remarks with respect to the translation and presentation of the German and Hungarian examples are in order here. There are two traditional methods of presenting linguistic examples. The first is generally applied in dealing with data from European languages and provides no interlinear translation at all. The second is rooted in the typological tradition and usually gives a highly annotated interlinear translation. It became apparent that neither is satisfactory for the kind of lexical study presented here. The first is problematic for its lack of transparency, while the second contains a lot of superfluous segmental information suggesting that the glosses given for the morphological segments have a universal applicability. Since morphological segmentation is not at issue here, I have decided in favor of the following compromise. A simplified interlinear translation will be given if the sentence in question contains a high portion of lexical material not being glossed in the immediately preceding sentences. When such sentences are slightly varied in subsequent examples I will give a translation for new lexical material in brackets. The translated sequence will be underscored, indicating its scope. Grammatical information will be given selectively depending on the topic in question, for instance, information about the NUMBER value of a word-form. Unless indicated otherwise, word-forms in the noun phrases have to be interpreted as singular forms (for instance, *Amerikaner* in (35)). Alternative lexical material in the source sentence will be separated with a double slash (//). Alternative constructions and partly synonymous expressions will be separated with a single slash (/) both in the source text and in the translation. Optional elements will be put in parentheses. For indicating the acceptability of a form-meaning pair, '*' and '?!' will be used as usual. The translation, in such cases, indicates the meaning which has been expected and failed by the construction given. All nominal phrases in question will be marked with bold face.

such as the existence of a default CONTAINER or default unit of measure and the situational context (restaurant, at home, etc.). The use of *Bier* without an explicit 'CONTAINER phrase' is highly conventionalized in the restaurant context, in opposition to *Kaffee* (with a default CONTAINER) and to *Milch* (without a default unit of measure). In a store context (cf. (37)), *eine Milch* is more likely to be accepted than *ein(en) Kaffee*. Here, a different sense of *Kaffee* is selected ('coffee beans'), and asking for an unspecified amount of (unspecified) coffee is pragmatically not felicitous in German stores. Example (38) illustrates the use of numerals higher than *one* with and without an explicit 'CONTAINER phrase' in combination with the same lexical material. The distribution of acceptability is parallel to that in (36) b. Note that a singular form is used here, in contrast to English, which could be taken as formal evidence for an elliptical analysis. This, however, is not without problems, due to the agreement of the inflected numeral *ein* with the lexical unit denoting the DRINK, rather than with the 'absent' CONTAINER.⁵⁸ The pattern of using a numeral and singular is not extended to LIQUIDS other than DRINKS, for instance to soups. *Zwei Suppe* ('two soup-SG') is regularly judged as ungrammatical, whereas *zwei Suppen* ('two soup-PL') as odd or as a feature of the speech of waiters/waitresses.⁵⁹

- (36) a. *Kann ich bitte (etwas) Bier //(etwas) Kaffee //(etwas) Milch haben?*
 can I please (some) beer //(some) coffee //(some) milk have
 'Can I have some beer // some coffee // some milk, please?'
- b. *Kann ich bitte ein [a] Bier // einen Kaffee // ?eine Milch haben?*
 'Can I have a beer // a cup of coffee // a milk, please?'
- c. *Kann ich bitte ein Glas [a glass] Bier // eine Tasse [a cup] Kaffee // ein Glas Milch haben?*
 'Can I have a glass of beer // a cup of coffee // a glass of milk, please?'

- (37) a. *Ich möchte [I would like] ein Bier // *einen Kaffee // eine Milch haben.*
 'I would like to have a bottle of beer // a package of / a pound of coffee // a bottle of / a liter of milk.'

⁵⁸ Modern Greek, for instance, provides a good example for a syntactic ellipsis since it exhibits GENDER agreement between numerals and absent 'PORTION nouns', resulting in a superfluous disagreement between numeral and nouns in the 'FOOD context' (*tris* ('three-FEM-PL') *kotopulo* ('chicken-NEUT-SG')). The absent word-form is *meridhes*, having the values feminine and plural.

⁵⁹ This is a stereotypical answer which was spontaneously given by natives of different languages (German, English, Hungarian, Italian). When judging constructions without an explicit 'UNIT phrase', they all agreed that these would be odd in everyday language but characteristic of waiters' speech.

- (38) a. Zwei [two] **Bier** [SG] // **Zwei Kaffee** [SG] // ?**Zwei Milch** [SG] // ?**zwei Suppen** [soup-PL] *bitte!*
 'Two beers // two cups of coffee // two milks // two soups / two bowls of soup, please!'
- b. Zwei Glas [two glass-SG] **Bier** // Zwei Tassen [two cup-PL] **Kaffee** // Zwei Glas [two glass-SG] **Milch** // Zwei Tassen [two bowls-PL] *Suppe, bitte.*
 'Two glasses of beer // two cups of coffee // two glasses of milk // two bowls of soup, please!'

The use of an indefinite article for indicating SORTS is generally acceptable for all lexical units in the domain of DRINKS and also in the domains of other LIQUIDS and amorphous SUBSTANCES, especially when the SORT interpretation is made explicit by a modifying adjective as shown in (39). German shows a split between SORT uses in singular and plural. An explicit mention of a 'SORT noun' heading the 'DRINK noun' in a compound (cf. (39) d.) or constituting a 'SORT phrase', like in English, is clearly favored with some lexical units (*Kaffee, Milch*) as against a construction without an explicit 'SORT noun' (cf. (39) b. and c.).⁶⁰ Here, the morphological capacity of constructing a plural form plays an important role, in addition to conventionalized pragmatic conditions.⁶¹ We receive results analogous to (39) b., c. and d. in combination with COUNT quantifiers like *viel* (inflected) ('many') (cf. (39) d.).

- (39) a. *Das ist* [that is] **ein ausgezeichnetes** [excellent] **Bier** // **ein ausgezeichnetes Kaffee** // **eine ausgezeichnete Milch** // **ein ausgezeichnetes Öl**.
 'That is a great beer // an excellent coffee // a great milk // an excellent oil.'
- b. *Sie verkaufen* [they sell] **ausgezeichnete Biere** [beer-PL] // ??**ausgezeichnete Kaffees** [coffee-PL] // ***ausgezeichnete Milche / Milchen** [milk-PL] // **ausgezeichnete Öle** [oil-PL].
 'They sell great kinds of beer / excellent sorts of beer // excellent coffees // great kinds of milk / excellent sorts of milk // excellent oils.'

⁶⁰ Note that adjectives biasing a SORT interpretation, as found in (39) c., do not alter the acceptability judgments.

⁶¹ Both are, of course, historically interdependent. Pragmatic conditions may result in morphological gaps. Pragmatic conditions, however, change more rapidly than morphological restrictions. From a synchronic point of view, morphological restrictions may clearly have a retarding effect on the generalization of a construction for a specific use.

- c. *Sie verkaufen **sehr gute belgische** [very good Belgian] **Biere** // ??**sehr gute südamerikanische** [very good South American] **Kaffees** // ***sehr gute holländische** [very good Dutch] **Milche / Milchen** // **sehr gute griechische** [very good Greek] **Öle**.
'They sell very good (sorts of) Belgian beer // very good (sorts of) South American coffee // very good (sorts of) Dutch milk // very good (sorts of) Greek oil.'*
- d. *Sie verkaufen **sehr gute belgische Biersorten** [beer-sort-PL] // ... **Kaffeessorten** [coffee-sort-PL] // ... **Milchsorten** [milk-sort-PL] // ... [oil-sort-PL].
'They sell very good sorts of Belgian beer // ... sorts of ... coffee // ... sorts of ... milk // ... sorts of ... oil.'*
- e. *Im Kaufhaus gibt es jetzt **?viele südamerikanische Kaffees** // in-the department store there is / are now many **viele südamerikanische Kaffeessorten**.
'You can get a lot of different South American coffees / a lot of different sorts of South American coffee at the department store now.'*

We can observe the fact that the plural forms of *Bier* and *Öl* are equally good alternatives for the explicit *Biersorten* and *Ölsorten* (in opposition to the plural forms of *Kaffee* and *Milch*) which typically occur in text types representing a taxonomic hierarchy, for instance, a list of beverages and the 'yellow pages'. In such hierarchies, it is the plural form that generally occurs instead of the singular form on the higher levels. (40) is an extract from a list of beverages in a bar in Munich and (41) from the 'yellow pages' for Cologne. Actually, we find *Biere* ('beers') as header for 6 sorts of beer on the list, in contrast, for instance, to *Sekt* ('sparkling wine', 'champagne'), including subordinate SORTS as well. In the phone book, *Öle* ('oils') occurs in a plural form like *Öfen* ('ovens'), rather than being combined with a superordinate dummy element (*Ware* ('ware'), *Stoff* ('material')) as its head, allowing morphological plural formation and/or SORT interpretation.

(40) List of Beverages:

- a. First Level: PL: *Heiße Getränke* ('Hot Drinks'), *kalte Getränke* ('Cold Drinks'), ***Biere*** ('Beers'), *Aperitifs* ('Aperitifs'), *Drinks* ('Alcoholic Drinks'), *Cocktails* ('Cocktails'); SG: *Sekt* ('Sparkling Wine', 'Champagne'), *Champaner* ('Champagne')
- b. Second Level: PL: *Liköre* ('Liqueurs'), *Digestifs* ('Digestifs'); SG: *Gin* ('Gin'), *Cognac* ('Cognac')

(41) Yellow Pages

PL: *Öfen und Herde* ('Ovens and Stoves'), *Papier- und Schreibwaren* ('Stationary (Products)'), *Alt- und Abfallstoffe* ('Waste and Scrap Materials'), ***Öle und Fette*** ('Oils and Fats'); SG: *Obst und Gemüse* ('Fruit and Vegetables'),

To summarize, at least four lexical classes which behave differently have been demonstrated above ((a) *Bier*, *Öl*, (b), *Kaffee*, (c) *Milch*, (d) *Suppe*). It could be maintained that these differences are not a matter of linguistic knowledge. This would be dangerous, however, since exactly such differences reflect the native (and not predictable) way of constructing, interpreting and judging utterances. Strictly speaking, such an attitude insisting on a distinction between linguistic and extralinguistic knowledge should have the consequence of abstaining from native judgments at all.

3.2.2 MOTIVATION, CRIME, and SENSATION

I will introduce the lexical units to be investigated in the three subfields of 'abstract nouns' through the explicit specification of a shared context called 'frame'. 'Frame' represents a lexically and grammatically specified context with a slot (indicated by 'X') in which a phrase headed by an appropriate word-form of the listed lexical units can be substituted. The purpose of this technique is to guarantee that the selected senses of the lexical forms in question (i.e., the selected lexical units) share their behavior in at least one context, which allows a comparison in further contexts.

3.2.2.1 MOTIVATION

(42) Frame: *Sie zeigte X*. ('She showed /demonstrated X')

Lexical units: *Einsatz* ('effort', 'devotion'), *Interesse* ('interest'), *Engagement* ('involvement', 'commitment')

All of the three lexical units (*Einsatz*, *Interesse*, *Engagement*) can be used with the uninflected (MASS) version of the quantifier *viel* in a singular form, but not with the inflected (COUNT) version of *viel* in a plural form (cf. (43)). Similar to English, the phrase with a MASS quantifier alternates with a phrase containing an indefinite article and an intensifying adjective (*viel Interesse* ('much Interest') / *ein starkes Interesse* ('a strong interest')) (cf. (19), (43)). It is probably correct to say that the use of a configuration of an indefinite article and a (lexically appropriate) adjective is generally allowed. The lexical forms in question all permit the construction of morphologically plural word-forms. There is, however, a considerable difference between *Einsatz* and *Engagement* on the one hand, and *Interesse* on the other hand.

In the first case, the plural word-form can only be associated with lexical senses other than the MOTIVATION sense (*Einsatz* ('mission', 'insert', in addition to 'effort'), *Engagement* ('engagement in a theater, etc.', in addition to 'involvement'), as corresponding ambiguity tests

unequivocally show.⁶² These ambiguity-proved distinct lexical units have a full NUMBER paradigm. In the second case, the plural form of *Interesse* (see also English *many interests*) can only be associated with a sense ('interest', 'preference'), closely related to the MOTIVATION sense. If these senses are seen to constitute distinct lexical units (as usually found in dictionaries; see, for instance, DUDEN), then they show complementary defective NUMBER paradigms (i.e., *singulare tantum* and *plurale tantum*) and, consequently, cannot be tested for ambiguity. The fact that these (hypothetically distinct) lexical units are closely related becomes apparent through their similar collocation behavior. The lexical context in (43) excludes a meaningful interpretation of *Einsatz* and *Engagement*, used in plural, in contrast to *Interesse*. In other lexical contexts such as found in (44), such interpretations are possible.

(43) *viel* (uninflected), *viel* (inflected), *ein*

a. *Sie zeigte viel Einsatz* [SG] // *einen beispiellosen* [incomparable, extraordinary, etc.] *Einsatz* [SG] // **viele Einsätze* [PL].

'She showed a great deal of effort.' // 'She demonstrated extraordinary effort.' //

'She showed a great deal of effort.'

b. *Sie zeigte viel Interesse* [SG] // *ein starkes* [strong] *Interesse* [SG] // *viele Interessen* [PL].

'She showed a great deal of interest // strong interest // interest in many things.'

c. *Sie zeigte viel Engagement* [SG] // *ein außergewöhnliches* [extraordinary, extreme] *Engagement* [SG] // **viele Engagements* [PL].

'She demonstrated a great deal of involvement // extreme involvement // a great deal of involvement'

(44) *Sie zeichnete sich durch* [distinguished herself through, made a name for herself through] *viele Einsätze* // *viele Interessen* // *viele Engagements aus* [PRE].

'She made a name for herself through many operations // in her many interests // in her many (permanent) engagements (with various theaters).'

Neither of the three lexical units is allowed to be used in phrases following an indefinite article, unless modified by an adjective or a relative clause (cf. (46)) in our initially defined frame. Bare phrases are, in contrast, allowed here, however, especially in the case of *Einsatz*,

⁶² Applying the identity test to the following sentence shows that no cross-interpretation is possible:

Peters Engagement *hat uns vielmehr überrascht als Pauls.*

P.'s involvement / engagement (in theater) AUX us much more surprised than P.'s

That is, we cannot interpret this sentence as 'Peter's involvement has surprised us much more than Paul's being engaged by a theater' and vice versa.

they are judged as worse than uses with an indefinite article and a modifying element (cf. (45)).

(45) bare phrase

*Sie zeigte **Einsatz** // **Interesse** // **Engagement**.*

'She showed effort // interest // involvement.'

(46) *ein* with and without modifiers

a. *Sie zeigte ***einen Einsatz** // ***ein Interesse** // ***ein Engagement**.*

'She showed effort // interest // involvement.'

b. *Sie zeigte **einen Einsatz, der in der Geschichte des Vereins beispiellos war**.*

'She showed an effort which was incomparable in the history of the club.'

Bare phrases are allowed to be used in subject position and in predicate position as well as having, in both cases, a 'metalinguistic' flavor: 'X means...' or 'That is what I call X.' (cf. (47), (48)). Thus, bare phrases combined with the copula imply the assertion of identity rather than the ascriptive predication of QUALITY, with the consequence that they are excluded from the PAST TENSE (see, however, (49)).

(47) ***Einsatz** ist [is]...*

'Effort means...'

(48) *Das ist [that is] **Einsatz!***

'That's real effort!'

(49) *Das war [that was] **ein beispielloser Einsatz**.*

'That was an exemplary effort.' (cf. (43))

3.2.2.2 CRIME

(50) Frame: *Er begeht // erleidet X.* ('He commits // suffers X')

Lexical units: *Mord* ('murder'), *Diebstahl* ('theft', 'larceny'), *Verbrechen* ('crime'),

Raubüberfall ('robbery'), *Raub* ('robbery'), *Unfall* ('accident')

We are dealing with two frames here. The first is designed for lexical units characterized by the feature 'intentionality' and the second for lexical units with the opposite value of this feature. Only the lexical units satisfying the first frame fall under the label CRIME, but, for

sake of contrast, the lexical unit *Unfall*, which is said to behave as a 'typical' COUNT noun, was also taken into account. Both frames define so-called 'lexical phrases'.

The use of an indefinite article is generalized in both frames (cf.(51) a. and b.), despite the fact that a specific interpretation is not necessarily implied here as in (51) c. (see the translation variations in English). In contrast to this, in prepositional phrases introduced by the preposition *wegen*, the lack of an indefinite article is usualized under non-specific interpretation for 'CRIME nouns' (cf. (52) a.). The only exception is the hyperonym *Verbrechen*, which requires an indefinite article in this position as well, just as does *Unfall* (cf. (52) b.)

(51) *ein*

a. *Er begeht **einen Mord** // **einen Diebstahl** // **ein Verbrechen** // **einen Raubüberfall** // **einen Raub.***

'He commits (a) murder // a theft / larceny // a crime // (a) robbery // (a) robbery.'

b. *Er erlitt **einen Unfall***

'He had an accident.'

c. *Wir [we] müssen [must] **einen Mord** // **einen Diebstahl** // **ein Verbrechen** // **einen Raubüberfall** // **einen Raub** // **einen Unfall** aufklären [clear up].*

'We must clear up a murder // a theft // a crime // a robbery // a robbery // an accident.'

(52) bare phrase

a. *Er wurde [AUX] wegen [for, due] **Mordes** // **wegen Diebstahls** //*

****wegen Verbrechens** // **wegen Raubüberfalls** // **wegen Raubes** verurteilt [accused].*

'He was accused of murder // of (a) theft // of a crime // of (a) robbery // of (a) robbery.'

b. *Er wurde wegen eines Unfalls freigestellt [let go].*

'He was let go due to an accident.'

In the predicate position, we encounter an alternation between phrases with and without an indefinite article both in the PRESENT and PAST TENSE (cf. (53)). This is not an entirely systematic syntacto-semantic alternation. There is a strong tendency for using the bare phrase for predicating a contrastive QUALITY (see (54) a. which is a stereotype formula in German crime films). The semantic effect of the indefinite article, as shown in the first sentence of (54) b., is, however, less clear. Its use with an adjective may be interpreted as an instance of a general pattern of German for overruling finer lexical differences (see also (57) c.).

(53) *Das ist // war (**ein**) **Mord**.*

that is // was (a)

'This / It is murder.' / 'That was (a) murder.'

(54) a. *Es* [it] *war Mord*. *Nein, Chef* [no, chief], *es war Selbstmord* [suicide] // *ein Unfall*.

'It was murder.' No, Chief Inspector, it was suicide // an accident.'

b. *Es war ein Mord*. // *Es war ein kaltblütiger Mord*.

'It was (a) murder. // It was cold-blooded murder.'

The behavior of the lexical units in question in the predicate position is not fully analogous to that found in the lexical phrase defining the initial frame (cf. (50), (51) a. and b.) and in prepositional phrases (cf. (52)). *Raub* and *Raubüberfall* are the critical elements here. The lexical unit, which is usually employed for spatio-temporally located INSTANCES, is the morphologically complex *Raubüberfall* (*Raub* + *Überfall* ('raid')), which differs from *Raub* by its ability to have plural forms.⁶³ This, however, does not imply that *Raub* is a ('good') MASS noun in opposition to *Raubüberfall* or, that the two have complementary distribution with respect to the standard distinguishing criteria for MASS and COUNT nouns. As we have seen in (51) and (52) (see also (55) c.), no difference between the two can be found. A difference can only be observed in the predicate position, where *Raub* with and *Raubüberfall* without an indefinite article are judged as 'odd'. With the exception of *Raub*, plural uses are allowed for all lexical units in question (cf. (56)). Furthermore, all give preference to the (inflected) COUNT version of *viel* over the (uninflected) MASS version (cf. (57)), whereas *viel Raub* and *viele Raube* are equally good or bad.

(55) bare phrase, *ein*, *ein* with modifier

a. *Es war Diebstahl* // **Verbrechen* // ??*Raubüberfall* // *Raub* // **Unfall*.

'It was (a) theft // a crime // (a) robbery // (a) robbery // an accident.'

b. *Es war ein Diebstahl* // *ein Verbrechen* // *ein Raubüberfall* // ?*ein Raub* // *ein Unfall*.

'It was (a) theft // a crime // (a) robbery // (a) robbery // an accident.'

c. *Es war ein schlimmer Diebstahl* // *ein schlimmes Verbrechen* // *ein schlimmer Raub* // *ein schlimmer Raubüberfall* // *ein schlimmer Unfall*.

'It was a serious theft // a serious crime // a serious robbery // a serious robbery // a serious accident.'

⁶³ DUDEN provides a plural form (*Raube*) and marks it as 'rarely'. Out of 28 persons I tested in a series of acceptability tests for nominalizations, 9 persons radically rejected this plural form, 17 accepted it and 2 were indecisive. The search in a corpus of 3 months' worth of newspaper texts has brought about the following result: *Raub* occurs only in singular and in all plural contexts, *Raubüberfall* is systematically used.

- (56) *Das waren schlimme Morde* [PL] // *schlimme Diebstähle* [PL] // *schlimme Verbrechen* [PL] // *schlimme Raubüberfälle* [PL] // ??*schlimme Raube* [PL] // *schlimme Unfälle* [PL].

'These were serious murders // serious thefts // serious crimes // serious robberies // serious robberies // serious accidents.'

- (57) *viel* (uninflected), *viel* (inflected)

- a. *Die Polizei* [the police] *hat* [AUX] *viele Morde* [PL] // *viele Diebstähle* [PL] // *viele Verbrechen* // *viele Raubüberfälle* [PL] // ??*viele Raube* [PL] // *viele Unfälle* [PL] *registriert* ['recorded'].

'The police recorded many murders / a great deal of murders // many thefts / a great deal of thefts // many crimes / a great deal of crime // many robberies / a number of robberies // many robberies / a number of robberies // many accidents / a great deal

of

accidents.'

- b. *Die Polizei hat *viel Mord* [SG] // **viel Diebstahl* [SG] // **viel Verbrechen* [SG] // **viel Raubüberfall* [SG] // ??**viel Raub* [SG] // **viel Unfall* [SG] registriert.

See translation of (57) a.

In 3.2.2.2, we have found at least three lexical groups which behave differently: (a) *Mord*, *Diebstahl*, and probably *Raubüberfall*, (b) *Verbrechen*, *Unfall*, and (c) *Raub*. The difference between group (a) and (b) is also reflected in the subject position: only group (a) displays an alternation between bare phrases and phrases with an indefinite article in the subject position (compare this with the results in 3.2.2.1).

- (58) a. *Mord* // *Ein Mord ist ein tragisches Ereignis* [tragic event].

'Murder / A murder is a tragic occurrence.'

- b. **Unfall* // *Ein Unfall ist ein tragisches Ereignis*.

'An accident is a tragic occurrence.'

3.2.2.3 SENSATION

- (59) Frame: *Ich empfinde X*. ('I feel X')

Lexical units: *Freude* ('joy', 'pleasure'), *Angst* ('fear'), *Schmerz* ('pain'), *Haß* ('hatred'), *Liebe* ('love'), *Glück* ('happiness'), *Kummer* ('sorrow'), *Leid* ('suffering', 'sorrow', 'pain'), *Ärger* ('annoyance', 'anger')

'Abstract nouns' in the domain of (mental) SENSATION may display, in German, a lexical ambiguity between a sense SENSATION and a sense EVENT CAUSING CORRESPONDING SENSATION. Although this is a systematic pattern, its application is relatively restricted. In our sample, it applies only to the lexical units *Freude* and *Ärger*. Dictionaries show considerable divergences with respect to whether or not they explicitly list both senses as distinct 'word senses'. Though the second sense can be considered as an INSTANCE sense, it is not simply, in Jespersen's words (see above) "a single instance of quality", but rather, an INSTANCE of an event. In the following examples, this ambiguity is taken into account. This will also be the case for both interpretations of *Schmerz* (i.e., mental and physical sensation). In contrast to this, the homonymous sense of *Glück* ('luck') will be excluded from the tests and the lexicalized metonymical senses of *Liebe* ('beloved person', 'love affair'), too.

With the exception of *Schmerz*, all lexical units in the sample fit as singular bare phrases in the frame under (59) (cf. (60)). However, *Schmerz* can be used, like the other lexical units, in a singular form quantified by the uninflected MASS version of *viel* (cf. (61)).

(60) bare phrase

Ich empfinde Freude // Angst // Schmerzen [PL] // *Haß // Liebe // Glück // Kummer // Leid // Ärger.*

'I feel joy / pleasure // fear.' // 'I feel pain / I have a sensation of pain (in my right foot).'

//

'I feel hatred // love // happiness // sorrow.' // 'I suffer.' // 'I feel annoyed.'

(61) *viel* (uninflected)

Ich habe [AUX] *sehr* [very] *viel Freude // sehr viel Angst // sehr viel Schmerz // sehr viel Haß // sehr viel Liebe // sehr viel Glück // sehr viel Kummer // sehr viel Leid // sehr viel Ärger in meinem Leben* [in my life] *erfahren* ['experienced'].

'I have experienced a great deal of joy // a great deal of fear // a great deal of pain //

a great deal of hatred // a great deal of love // a great deal of happiness //

a great deal of sorrow // a great deal of anger and annoyance in my life.'

This semantic field confirms the general preference in German for using indefinite articles in the presence of modifying adjectives as already noted for other subfields of 'abstract nouns' above. There is no lexical unit which would prohibit indefinite articles in this context (cf. (63)). We find finer lexical differences with respect to the optional use of an indefinite article in so-called 'thetic utterances'⁶⁴ (cf. Sasse 1995) as exemplified in (62). In (62) a., we observe

⁶⁴ In this type of utterance, the subject is said to fall within the scope of assertion rather than constituting an autonomous constituent (like generic subjects, for instance) about which something is asserted.

a decreasing ability for free alternation between uses with and without an indefinite article. *Kummer* and in particular *Haß* in (62) b. show a conventionalized preponderance to be used without indefinite article inthetic utterances, even if they are modified by an adjective. In (62) b., as in (60), *Schmerz* deviates from the other lexical units in showing a 'clear' COUNT behavior, i.e., it is restricted to only being used with an indefinite article.

(62) bare phrase, *ein*

a. (**Eine**) *tiefe* ['deep'] *Freude* // (**Ein**) *unermeßliches* ['immense'] *Leid* //

(**Ein**) *starker* ['strong'] *Kummer erfüllte* ['filled with'] *ihn* [him].

'He was filled with (a) deep joy // with (an) immense sorrow //

with (an) immense sorrow.'

b. (**Eine**) *schreckliche* [terrible] *Angst* // (**Ein**) *stechender* [piercing] *Schmerz* //

Blinder [blind] *Haß* ergriff [overcame] *ihn*.

'A terrible fear / (A) great fear // (A) penetrating pain // Blind hatred overcame him.'

(63) a. **Eine leidenschaftliche Liebe** beherrschte ihn in jenen Jahren.

passionate / ardent commanded him in those years

'He was commanded by an ardent love in those years.'

b. **Ein stilles Glück** erlebte er in jenen Jahren.

quiet experienced he

'He experienced a quiet happiness in those years.'

Conventionalized prepositional phrases and 'multi-words' occur systematically without an article, as seen in (64).

(64) a. *mit* // *aus Freude*, *mit* // *aus Angst*, *mit* // *aus Liebe*, etc.

'with pleasure' // 'for fun', 'with fear' // 'out of fear', 'with love' // 'out of love'

b. *etw. macht jmdm. Freude* // *Angst* // *Kummer* // *Ärger*

's.th. is a lot of fun for s.o. // s.th. scares s.o. // s.th. gives s.o. a lot of worry //

s.th. annoys s.o.'

c. *jd. schürt Haß*

's.o. stirs up hatred'

A well-known difference between German and English consists in the fact that in German, generic subjects or objects in certain subdomains of 'abstract nouns' are not confined to bare phrases, but can, alternatively, be used with a definite article. From this follows that corresponding phrases with a definite article are ambiguous between a generic (TYPE) and a specific (TOKEN) reading. The semantic field of mental SENSATION exhibits such a determiner variation, whereas the lexical units which were discussed above under the label MOTIVATION

behave like their English counterparts. In the field of mental SENSATION, we encounter the same variation between uses with and without a definite article in the predicate position as in the subject and object position (cf. (65) a. and b.). *Schmerz* constitutes an exception here too, permitting only phrases with a definite article, and one may be inclined to ascribe this fact to the 'COUNT status' of *Schmerz*.

- (65) a. *Das ist ?Freude // Angst // *Schmerz // Haß // Liebe // Glück // Kummer // Leid // ??Ärger.*
 'That can be called 'joy'.' // 'That is fear // pain // hatred // love // happiness // sorrow // suffering.' // 'That can be called 'annoyance'.'
- b. *Das ist die Freude // die Angst // der Schmerz // der Haß // die Liebe // das Glück // der Kummer // das Leid // der Ärger.*
 i 'That can be called 'joy'.' // 'That's fear // pain // hatred // love // happiness // sorrow // suffering.' // 'That can be called 'annoyance'.'
 ii 'That is the specific feeling you have told me about...'
- c. *Das ist schmerzlich [ADJ] // ärgerlich [ADJ].*
 'That's (really) painful // annoying.'
- d. *Das ist eine Freude / eine freudige [ADJ] Sache ['thing', 'event'] // ein Ärgernis / eine ärgerliche [ADJ] Sache.*
 'That's a real joy.' // 'That's an annoying situation.'
- e. *Das ist *ein Schmerz / ein schmerzliches [ADJ] Erlebnis [experience].*
 'That's a painful experience.'
- f. *Dás ist ein Schmerz! // Dás sind Schmerzen!*
 'It's / That's so painful.'

The situation is, however, more complicated. (65) a. and b. are two alternative ways of expressing the fact that a list of properties (mentioned verbally or experienced in a situation) is recognized as satisfying the conditions to be named in a certain way. While *Schmerz* is the clearest outsider, the acceptance of these two alternative constructions also varies among the remaining lexical units in our sample. The construction with a bare phrase is judged as questionable, especially in those cases where the lexical unit in question shows an ambiguity between the senses SENSATION VS. EVENT CAUSING SENSATION, i.e., in the case of *Ärger* and, to a lesser extent, in the case of *Freude*. In those cases in which both constructions are judged as equally good (*Liebe*, for instance), the following question arises: Does the determiner variation systematically correlate with a semantic difference or not? One possible solution could be to assume a semantic difference and associate it with the difference between identity-stating sentences (use with a definite article) and categorizing ('ascriptive') sentences (use without a definite article), i.e., with the difference between referential and predicative uses in the dimension of referentiality. But in this case, we must allow that lexical units

restricted either to bare phrases (cf. *Einsatz* in 3.2.2.1) or to phrases with definite articles are ambiguous about identity stating and categorizing when used in predicate position.

It is worth noting that *Schmerz* and *Ärger* have adjectival forms (*schmerzlich*, *ärgerlich*) (cf. (65) c.) as quasi-alternative expressions for (65) a. These sentences are, however, not characterized by a 'metalinguistic' flavor. The use of the indefinite article for signaling an EVENT CAUSING SENSATION is only possible for *Freude* (cf. (65) d.). Instead of *Ärger* (though it has a corresponding lexical sense) a related lexical form *Ärgernis* is used and *Schmerz* does not allow a metonymical interpretation as painful event at all (cf. (65) e.). And *Schmerz*, even in the more concrete sense of a physical SENSATION, occurs only under restricted discourse conditions with an indefinite article, unless modified by an adjective.

(66) illustrates the morphological ability to construct plural forms. This cross-cuts the lexically established ambiguity structure and the determiner behavior. We find two types of morphological alternatives for expressing multiplicity of INSTANCES (cf. (67)) when no morphological plural is possible: compounding with a head like *Gefühl* ('feeling', 'sense') (cf. (66) b.) or a related lexical form exhibiting morphological plural (cf. (66) c.).

(66) a. +PL: *Freude, Angst, Schmerz*

b. -PL: *Haß, Liebe, Glück* > *Haßgefühle* ('hatred-sense-PL'), *Liebesgefühle* ('love-sense-PL'), *Glücksgefühle* ('happiness-sense-PL')

c. -PL: *Ärger, Kummer, Leid* > *Ärgernisse, Kummernisse, Leiden*

(67) *Wenn* [when] ***die Freuden // die Ängste // die Schmerzen // die Glücksgefühle // die Haßgefühle // die Liebesgefühle // die Kummernisse // die Ärgernisse // die Leiden*** *vergessen sind* [are forgotten] ...

'When the joys // the fears // the painful experiences // the feelings of happiness // the feelings of hatred // the feelings of love // the sorrows // one's sufferings // the annoyances are forgotten...'

We can summarize the results of the investigation in 3.2.1 and 3.2.2 as follows. We have found general lexical patterns with respect to determination and quantification which characterize different semantic fields. Furthermore, within each semantic field, more than two lexical classes with different behavior have to be distinguished, which cannot be grasped by a simple use of criteria, such as compatibility with an indefinite article or ability to build plural forms. We have observed four different factors which are constitutive for lexical classes and triggering the syntactic behavior and semantic interpretation of lexical units: (a) purely morphological factors, (b) lexically established ambiguity structures, (c) conventionalizing of pragmatic factors, and (d) idiosyncratic behavior which is especially unstable.

3.2.3 Morphologically Complex Unitary Concepts

We are faced with three different, but partly convergent, research activities which center on the phenomenon called 'morphologically complex unitary concepts' here. First, there is a long typological tradition of investigating incorporation and compounding cross-linguistically, with respect to semantic and pragmatic properties of incorporated elements and to the relation of the verbal and nominal lexical units involved. It has been observed that there is a cross-linguistic tendency towards a habitual or progressive reading of incorporated constructions and towards a non-specific, non-individual or 'mass-like' interpretation of the incorporated element (cf. Mithun 1984: 856). Moreover, the lexical relation of the verbal and nominal lexical units determines the likelihood of incorporation in that frequently combined elements will be interpreted as unitary concepts and more easily incorporated and conventionalized. The second line of research concerns the study of 'idioms' in a broader sense of the term, which concentrates on language-specific properties and classification of morphologically complex, but semantically simple expressions. For our purposes, what is of importance is that type of idiom research which also takes paradigmatic aspects into account by exploring lexicalized patterns whose instances are not necessarily 'idiomatic' with respect to the opacity of the constituents, but idiomatic with respect to paradigmatic alternatives. The third research activity is located in computer lexicography. In this context, a good deal of experience has been made in extracting 'collocations' from machine-readable dictionaries and text corpora (cf. Smajda/McKeown 1990; Zernik (ed.) 1991; Boguraev/Pustejovsky 1990). Here, it was possible to explore all types of frequently occurring combinations of lexical elements, not only traditional collocations. Especially interesting are investigations made from the perspective of the noun, for instance, "What does one normally do with books?" (cf. Boguraev 1991: 249).⁶⁵ Such investigations close the gap between typological studies on incorporation and studies on idioms and lead to a more general notion of 'morphologically complex unitary concepts.'

In this potential domain of unitary concepts, the use of singular bare phrases in the object position is not very common. It is strictly lexicalized in (68) b. (in contrast to English, where a definite article is also allowed), and in some single (non-recurrent) phrases, such as (68) a.

- (68) a. *Ich lese Zeitung.*
 I read newspaper

⁶⁵ Pustejovsky (1991) uses this type of information, represented as the "Qualia structure" of lexical units, for metonymy resolution.

'I am reading the newspaper.'

b. *Ich spiele Klavier // Schach // Tennis...*

I play piano // chess // tennis...

'I play piano // chess // tennis...'

All candidates for a unitary concept under (69) are used either with a bare plural phrase or with a singular phrase containing an indefinite article. All these phrases have the property of OBJECTNESS in the corresponding dimension in that morphological NUMBER unequivocally forces a particular interpretation. That is, using bare plural phrases in a non-specific way necessarily implies the multiplicity of the entities in question. And the other way around, if an utterance like (69) b. is made in the sense of 'I am a car-owner', rather than 'I have a specific car', a subsequent correction such as *Genauer gesagt, drei.* ('To be exact, three.') is not appropriate.⁶⁶

(69) a. *Ich schreibe einen Brief / Briefe.*

I write a letter-SG/ letter-PL

'I am writing a letter // letters.'

b. *Ich habe ein Auto.*

I have a car

'I have a car.'

c. *Er gießt Blumen //schält Kartoffeln //putzt Fenster //wäscht Teller...*

He waters plants //peels potatoes //cleans windows //washes plates...

'He waters plants / is watering (the) plants // peels potatoes / is peeling

(the) potatoes // cleans windows / cleaning (the) windows // washes dishes /

is washing (the) dishes...'

In the examples under (70), German opts for the use of a definite article (cf. English examples under (27)). The respective phrases are generally ambiguous about non-specific and specific, definite interpretation since German lacks verbal marking of aspectual properties (i.e., marking the difference between habitual and actual situation). Sometimes, however, disambiguation is achieved by prepositions. The aspectual disambiguation and the correct selection of the metonymically related senses in (70) b. vs. (70) c., for instance, are triggered by the difference between *auf* and *in*. Contracted and non-contracted forms of prepositions and determiners (*im* vs. *in dem* ('in the')) show an overlapping distribution. Whereas the former is ambiguous, the latter is restricted to a specific, definite interpretation. GENDER is less strong

⁶⁶ The same holds true for English. Nevertheless, such utterances are sometimes attested in English. Du Bois (1980: 211), for instance, cites the following utterance made by a musician:

"I finally found out what the best is. I have a Mercedes - three of them in fact." (Downbeat 4/6/78).

than NUMBER. (70) e. may be followed by an utterance in which a specific female doctor is explicitly mentioned.

- (70) a. *Ich fahre mit der Straßenbahn // mit dem Zug // mit dem Taxi...*
 I go with the tram // train // taxi...
 'I go by tram // by train // by taxi...'
- b. *Er ist im Gefängnis // auf der Universität...*
 He is in-the jail // on / at the university...
 'He is in jail // at university / at college'⁶⁷
- c. *Er ist im Gefängnis // in [in] der Universität...*
 'He is in the jail // at the university / at the college right now.'
- d. *Ich gehe zur Schule.*
 I go to-the school
 'I go to school.'
- e. *Ich gehe zum [to-the] Arzt [doctor].*
 'I am going to the doctor.'

Whereas locative phrases allow no variation between singular phrases with or without article, this is exactly what can be observed in instrumental phrases typical for attributive uses in the dimension of referentiality. (71) b. is the first clear example in this section where QUALITY (i.e., the method of fishing) is foregrounded instead of OBJECTNESS (cf. the English example (27) i.). There is no implication about the number of worms or fishnets used. Definite articles would result in a definite specific interpretation, in contrast to (71) a., where the default interpretation is that only one typewriter, computer, etc., is employed. (72) illustrates further lexical differences found with potential unitary concepts expressed by a verbal head and a prepositional phrase.

- (71) a. *Ich arbeite mit (der) Schreibmaschine // mit (dem) Computer...*
 I work with (the) typewriter // computer...
 'I work with typewriters // with computers...'
- b. *Er angelt mit Wurm // mit Netz...*
 He fishes with worm // fishnet...
 'He fishes with worms // with a fishnet / with fishnets...'

- (72) *Es riecht hier nach Katze / Katzen // *Bohne / Bohnen.*
 It smells here after cat-SG / cat-PL // bean-SG / bean-PL
 'Here, it smells of cat // beans.'

In German, too, generically-used singular bare phrases can be found with 'prototypical COUNT nouns' in the domain of 'concrete nouns', such as *Schreibmaschine* or *Katze* (cf. *Car is the*

⁶⁷ Cf. differences between British and American English.

best mode of transport., example (18) b.), though bare plural phrases, singular phrases with a definite article or with an indefinite article (for indefinite generics) are, as in English, more common. (73) c. is the second example in this section which points to the foregrounding of QUALITY without NUMBER sensitivity in a non-subject/non-object position.

- (73) a. *Schreibmaschine* ist besser als *Computer*.
 typewriter is better than computer
 'Typewriters are better than computers.'
- b. *Katze* [cat] ist besser als *Hund* [dog].
 'Cats are better than dogs.'
- c. *Ein Leben ohne Computer// ohne Frau...*
 a life without computer // woman...
 'Life without a computer /computers // a woman / women...'

3.2.4 Grinder Effects

Lexical units in the 'FOOD field' are clearly sensitive for 'grinder effects' triggered by the contrast between grammatical MASS and COUNT contexts. The COUNT quantifier in (74) a. forces an interpretation according to which reference is made to intact apples; (74) b. can be used for referring to a certain amount of apple puree or to a certain number of apple slices.

- (74) a. *Gibst du mir ein paar Äpfel?*
 give you me some / a few apple-PL
 'Will you give me some apples?'
- b. *Gibst du mir ein bißchen [some / a little bit of] Apfel [SG]?*
 'Will you give me some apple / some pieces of apple?'

In (75), three related test-frames are introduced for proving the possibility of foregrounding QUALITY vs. OBJECTNESS. It is predicted that only those lexical units which allow foregrounding of QUALITY will be appropriate in singular bare phrases in this context. This is actually the case for *Wasser*, as can be seen in (76) b. and d. *Ei*, *Kartoffel*, and *Tomate*, in contrast, can take only bare plurals (cf. (76) a. vs. (76) c.); that is, this specific context is biased against interpretation as 'grinded' SUBSTANCE, and simultaneously retaining characteristic SHAPE and focusing on QUALITY rather than on OBJECTNESS is not possible. However, in contexts in which interpretation as amorphous SUBSTANCE is appropriate, bare singular phrases are allowed and inversely, these can only be interpreted as amorphous SUBSTANCE (cf. (77)).

- (75) *Frames: Der Korb ist voll von / mit X.* ('The basket is full of X')
X ist / sind im Korb. / Im Korb ist / sind X. ('X is / are in the basket' /
 'There is / are X in the basket')
 Lexical units: *Ei* ('egg'), *Kartoffel* ('potato'), *Tomate* ('tomato')
- (76) a. *Der Korb ist voll mit Eiern* [PL] // *mit Kartoffeln* [PL] // *mit Tomaten* [PL].
 'The basket is full of eggs // potatoes // tomatoes.'
 b. *Der Eimer* [bucket] *ist voll mit Wasser* [water].
 'The bucket is full of water.'
 c. **Ei* [SG] // **Kartoffel* [SG] // **Tomate* [SG] *ist im Korb.*
 'Eggs // potatoes // tomatoes are in the basket.'
 d. *Wasser* [SG] *ist im Eimer.*
 'Water is in the bucket.'
- (77) a. *Hast du schon Ei in deinem Teig?*
 have you already egg in your dough
 'Do you already have eggs in your dough?'
 b. *Leberkäse mit Ei*
 'Meat Loaf with Egg (Southern German food)'

The German lexical form *Glas* displays the same ambiguity between MATERIAL + ARTIFACTS MADE OF MATERIAL as its English translation *glass*. In German, too, grammatical MASS/COUNT contexts can be exploited for a disambiguation (cf. (78) a., b. and c.), but note that the indefinite article can represent a neutral context (cf. (78) d. vs. e.) due to the general pattern of German using the indefinite article in the presence of modifying adjectives and/or for SORT interpretations.

- (78) a. *Das ist Glas.*
 that is glass
 'That is glass.'
 b. *Das ist ein Glas.*
 'That is a glass.'
 c. *Das ist ein Stück Glas.*
 'That is a piece of glass.'
 d. *Das ist ein sehr gutes Glas.*
 'That is very good glass, i.e., a very good material.'
 e. *Das ist ein schönes Glas.*
 'That is a beautiful glass, i.e., a beautiful vessel.'

According to DUDEN, the German *Kreide* has two senses (MATERIAL + PIECE OF MATERIAL) which correspond to the distinct lexical forms *chalk* and *crayon* in English. We should expect that the use of an uninflected (MASS) quantifier vs. an inflected (COUNT) quantifier will correlate with these two senses; this, however, is not the case (cf. (79) a. and b.). The MASS quantifier turns out to be a neutral context with respect to the interpretation of MATERIAL VS. PIECE OF MATERIAL, whereas the COUNT quantifier is not suitable for signaling the distributive multiplicity of crayons. For this purpose, a compound is used containing *Kreide* as modifying element: *Kreidestift* (*Kreide* + *Stift* ('stick')). In principle, the definite article constitutes a neutral environment (cf. (79) c.) and phrases with an indefinite article, as found in (79) e., are not applicable for referring to single pieces. Nevertheless, the verbal predicate can bias the singulative interpretation of a definite phrase (cf. (79) d.). This fact points to a lexical preponderance of the MATERIAL sense.

- (79) a. *Wir haben sehr viel Kreide gekriegt.*
 we AUX very much chalk / crayon received
 'We have received a lot of chalk / a lot of crayons.'
- b. ??*Wir haben sehr viele Kreiden [PL] gekriegt.*
 'We have received a lot of crayons.'
- c. *Die [the] Kreide ist [is] im anderen Zimmer [in the other room].*
 'The chalk / The crayon is in the other room.'
- d. *Die Kreide ist [AUX] zerbrochen [broken into pieces].*
 'The crayon has been broken into pieces.'
- e. *Gib [give] mir *eine [a] Kreide / ein Stück [a piece] Kreide.*
 'Give me a crayon / a piece of chalk.'

Kreide can, of course, be used in the same phrasal context as *Lehm*, *Tinte*, and *Soft*, indicating a MATERIAL sense (cf. (80) a.). However, lexical units primarily denoting discrete objects such as *Bleistift* or *Kugelschreiber* in a singular form and with an optional indefinite article can also appear in this context (cf. (80) b.). One could conclude that *Bleistift* and *Kugelschreiber* are subject to a productive metonymy resulting in the senses 'pencil lead' and 'ink'. But such a process would not have a restrictive effect on the use of the indefinite article as in the case of *Lehm*, etc. (**mit einem Lehm*). Another solution would be to assume ambiguity of *Bleistift* in *mit einem Bleistift* and *mit Bleistift* in (80) b. and to interpret the first as instrument producing MATERIAL, the second as (metonymically derived) MATERIAL. German natives, however, have no intuition whether or not the use of the indefinite article correlates with such a semantic difference and thus are not able to confirm the second hypothesis.⁶⁸

⁶⁸ A detailed corpus study of finer contextual differences is needed to clarify this issue.

- (80) a. *Du hast dich mit Kreide //mit Lehm//mit Tinte //mit Saft beschmiert.*
 you AUX yourself with chalk/crayon// mud // ink // juice soiled
 'You've got yourself dirty with chalk // mud // ink // juice.'
- b. *Du hast dich mit (einem) [a] Bleistift [pencil] // mit (einem) Kugelschreiber [pen] beschmiert.*
 'You've got yourself dirty with (a) pencil / with pencil lead // with (a) pen / with ink.'

The *Kreide*-case is a good illustration of a genuine problem we are commonly faced with: We must simultaneously deal with (a) lexically established preferences between closely related senses, (b) entirely productive processes, and (c) the fact that the formal alternation of grammatical markers may have semantic correlates in some contexts and may be a case of free alternation in others.

3.3 Hungarian

In section 2.3.1, it has been shown that Hungarian is similar to English with respect to the adjective-noun ambiguity in specific semantic fields but behaves differently in other respects (determiner behavior in the predicate position). The following sections, in which the Hungarian data are presented, are almost analogous to the presentation of the German data but appear in a different order, which follows from the discussion of specific properties of Hungarian. The domain of 'abstract nouns' will not be discussed in such a detailed fashion as with the German. Rather, I will try to elaborate some significant contrastive differences between Hungarian on the one hand, and German and English on the other. After some introductory remarks about the system of quantifiers and determiners, 'morphological complex unitary concepts' will be discussed. This will be followed by a section on the domain of 'abstract nouns' and a section on the domain of DRINKS. Finally, it will be investigated as to whether or not Hungarian shows 'grinder effects'.

The most important linguistic facts concerning the MASS/COUNT distinction in Hungarian can be summarized as follows:

- All quantifiers and numerals can be used and are generally used with word-forms in the singular (cf. (82)). The most important quantifiers are listed under (81). Of these, only those listed under (83) allow, in principle, a free alternation between singular and plural forms. All other quantifiers are strictly confined to singular phrases. Moreover, plural phrases in combination with quantifiers displaying NUMBER ambivalence primarily occur in certain conventionalized phrases such as in (83), and, these are often translations from

German or other languages (for instance, *minden körülmények között* from German *unter allen Bedingungen*).⁶⁹

- (81) *sok* ('much', 'many', 'a lot of', 'a great deal of', etc.) *kevés* ('little', 'few', 'a few'), *minden* ('all', 'each', 'every'), *összes* ('all', 'total', 'collective', 'every'), *valamennyi* ('all (of them)', 'every (one of them)', 'a little', 'some (unstressed)'), *számos* ('numerous', 'a number of'), *számtalan* ('innumerable'), *néhány* ('a few', 'several'), *egypár* ('a few', 'a couple of'), *annyi* ('so much', 'so many'), *ugyanannyi* ('just as much', 'just as many'), *a többi* ('the remaining', 'the other'), numerals (*kettő* ('two'), *ezer* ('thousand'), etc.)
- (82) *sok ház* ('many houses'), *sok ember* ('many people'), *sok pénz* ('much money'), *sok idő* ('a lot of time'), *sok boldogság* ('much happiness')
- (83) *sok, minden, összes, a többi*
Petőfi összes verse [SG] / *versei* [PL] ('collected poems of Petőfi'), *minden körülmények* [PL] *között* ('by all means', 'under any circumstances')

- Most quantifiers do not subdivide the nominal lexicon into formal classes.⁷⁰ *Sok*, for instance, quantifies both *ház* and *boldogság* in a singular phrase (cf. (82)), whereas in corresponding English and German phrases, different lexical forms (*many* vs. *much*) or different word-forms (*viel* (inflected) vs. *viel* (uninflected)), correlating with different NUMBER values in both cases, are applied. The Hungarian system of quantification also contains a pair of quantifiers, i.e., two distinct lexical forms which imply different semantic interpretations of phrases in a complementary way: The 'question words' *mennyi?* and *hány?*, as demonstrated in (84), systematically correlate with a collective and a distributive interpretation, respectively. Basically, all lexical units denoting solid objects can be combined with either of them, resulting in distinct interpretations in the sentence. Nevertheless, *mennyi?* and *hány?* are characterized by a lexical restriction of the elements they can modify. This is found in the difference between measure or unit nouns on the one hand, and all other nouns on the other hand. Measure or unit nouns require the use of the

⁶⁹ There is an hypothesis which states that the plural use was indeed brought into usage in the last century. At that time, a lot of western expressions were adopted in a great campaign for extending the Hungarian vocabulary. This hypothesis is probably not correct. Nevertheless, we can observe today an increasing tendency for singular use, also in such phrases whose plural use is partly conventionalized as in *Petőfi összes verse* (SG) in (83).

⁷⁰ Here, I do not consider the fact that amount phrases such as 'X liter of...' certainly show co-occurrence restrictions in Hungarian as well, since such restrictions are not relevant for the rise of paradigmatic structure in the domain of quantification.

distributive version (*hány?*). Thus, we find two alternative expressions for asking for an amount of apple ((84)) or for time ((85)). Interestingly enough, it was exactly this difference which was identified by Cao Xuân Hao as the only relevant formal distinction in Vietnamese, measure/unit nouns being lexically specified with respect to FORM/SHAPE, the others being unspecified. The situation in Hungarian, however, is somewhat different from that in Vietnamese since Hungarian lexical units are not neutral with respect to conceptualization with or without characteristic FORM, as we shall see later (cf. 3.3.4). Those lexical units which are lexically marked as having a characteristic FORM are neutral with respect to the grammatical distinction between distributivity and collectivity imposed by *mennyi?* and *hány?*.

(84) *mennyi?* ('how much?'), *hány?* ('how many?')

a. *Mennyi alma termett az idén?*

apple was produced this year

'How many apples were produced this year?'

b. *Hány mázsa [centner] alma termett az idén?*

'How many centners of apples were produced this year?'

(85) *Hány óra* [hour, o'clock] *van* [is]? / *Mennyi idő* [time] *van*?

'What's the time?'

- The morphological formation of the plural is similarly easy like in English. We find some lexically conditioned allomorphic variation which is diachronic in origin, but otherwise, plural formation is entirely predictable in present-day Hungarian. Consequently, there are no morphological gaps as in German. Example (86) shows the correct plural of lexical forms which, in one of their senses, denote MATERIALS ((86) a.), abstract concepts in the semantic fields of CRIME and SENSATION ((86) b.) (compare the German data in 3.2.2 with this), and cultural uniques such as *humanizmus* ((86) a.). All these plural forms are actually used when pragmatically appropriate SORT or INSTANCE interpretations are possible. That is, there are no morphologically or syntactically motivated and lexically established NUMBER restrictions which would force the use of alternative constructions (i.e., explicit 'SORT phrases', compounds headed by a 'SORT noun', or the use of other lexical forms which allow plural) in such contexts. In other words, Hungarian has no real *singulare tantum* and, typically, no real *plurale tantum*. Only for geographical names are a few *pluralia tantum* attested, which are translation copies of the source structure (*az Egyesült Államok* ('the United States'), *Alpok* ('Alps')).

(86) a. *ezüstök* ('silver-PL'), *selymek* ('silk-PL'), *vizek* ('water-PL'), *tejek* ('milk-PL')

b. *rablások* ('robbery-PL'), *boldogságok* ('happiness-PL'), *szenvedések* ('suffering-PL')

c. *nyelvészetek* ('linguistics-PL'), *bürokráciák* ('functionarism-PL', 'red-tape(dom)-PL'), *piacgazdaságok* ('market economy-PL'), *humanizmusok* ('humanism-PL')

- As already demonstrated in 2.3.1, all Hungarian nouns, not only nouns in specific semantic fields such as in German, can appear without an indefinite article in the predicative position (cf. (87)). The use of a bare singular phrase is not even restricted to the predicate position. Bare singular phrases systematically occur with all nouns in other syntactic positions as well, for instance, in the position of direct object (cf. (88)) or in that of subject in possessive or existential sentences (cf. (90)). They contain no NUMBER implications at all. In sentences (88) and (89), it remains linguistically open as to how many entities (i.e., how many apples, newspapers, stamps, letters) are involved in the situation expressed by the object argument and the finite verb. Subsequent sentences, which make the number of entities involved explicit, are basically possible. We have seen that in German, a discourse sequence in which first the ownership of a car and then the exact number of cars owned is asserted is not appropriate, if (the German lexical unit for) car occurs in the first sentence in a singular phrase and the exact number of cars is higher than one. This is exactly what constitutes a usual discourse sequence in Hungarian, as illustrated in (90). It must be emphasized that only singular bare phrases exhibit this 'transnumeral' property. Phrases determined by a definite article or an indefinite article, but also bare plural phrases, are clearly NUMBER-sensitive.⁷¹

(87) *Ő férfi // tanár.*

He man // teacher

'He is a man // a teacher'

(88) *Kati almát eszik // újságot olvas // bélyeget gyűjt.*

K. apple-SG-ACC eats // newspaper-SG-ACC reads // stamp-SG-ACC collects

'Kati is eating an apple / apples // is reading the newspaper // collects stamps.'

(89) *Egész délután levelet írtam.*

all afternoon letter-SG-ACC wrote-I

'I was writing a letter / letters all afternoon.'

⁷¹ Link (1991: 418) had assumed that Hungarian bare plural phrases are 'transnumeral', the individuating role of singular being neutralized. In my opinion, the opposite is the case. The singular alone has no individuating function in Hungarian, rather, the determiners can evoke individuation. In contrast, the plural implies distributivity even without determiners, either under the condition of spatio-temporally concreteness or spatio-temporally abstractness, rather than "unstructured multiplicity" ("eine eher strukturlose Vielzahl") (cf. *ibid.*).

- (90) *Nekem is van gyerekem. Hány? Három nagy lányom van.*
 Me as well is child-my-SG How many? three grown daughter-my-SG is
 'I have children as well. How many? I have three grown daughters.'

We can summarize the most important linguistic facts in Hungarian as follows. Seen from the perspective of the lexicon, the standard criteria for distinguishing MASS and COUNT nouns in Western languages cannot be applied. There are no mutual restrictions between morphological NUMBER marking and quantifiers which could be organized into a macroparadigm. There is one paradigmatic cell which could be interpreted in terms of a MASS/COUNT distinction, namely, the opposition between 'question words'. This, however, is not paired with lexical sensitivity among lexical units other than measure or unit nouns.

3.3.1 Morphologically Complex Unitary Concepts

Bare singular phrases do not occur everywhere in the sentence. They generally occur directly before the finite verb, unless they constitute contrastive generics or the sentence contains a focus on a different constituent. Focused constituents also occupy the preverbal position and lead to the inversion of bare singular phrases. One basic feature of Hungarian syntax is the following systematic alternation: A certain lexical unit is either used in an indeterminate singular phrase occurring in preverbal position, or in a definite (singular or plural) phrase occurring in postverbal position. Phrases with an indefinite article or with a quantifier cross-cut this alternation and may occur both preverbally and postverbally. The semantic interpretation of the bare singular phrase, which I have called 'verb modifier' in Behrens (1982), is puzzling in the same way as that of incorporated elements in other languages. MacWhinney (1989: 232), for instance, writes:

"In Hungarian, such nouns tend to take on a meaning of mass or generic quality. They are clearly not definite."

To be exact, it is the QUALITY as opposed to OBJECTNESS that is foregrounded in the meaning of the Hungarian verb modifiers. And this foregrounding of QUALITY is also found in English phrases based either on MASS nouns or generically-used COUNT nouns. It comes as no surprise that all 'morphologically complex unitary concepts' are constructed in Hungarian, with a singular bare phrase in the preverbal position (cf. (91) a., (92) a., and (93) a.). Recall that we have found the following constructions in the corresponding English and German sentences: (a) phrases with an indefinite article or bare plural phrases, (b) singular definite phrases alternating with singular bare phrases, (c) plural bare phrases, and, only exceptionally, (d) singular bare phrases alone. In Hungarian, in turn, we rarely encounter non-specifically used

bare plurals. The postverbal direct-object phrases, locative phrases and instrumental phrases in the b.-sentences signal definite and specific interpretation. From this, it follows that the alternation in question has the side effect of selecting the correct sense in the case of systematic metonymical ambiguities such as shown in (92).

- (91) a. *Piri virágot* *öntöz* // *virágot tesz a vázába* //
 P. plant / flower-SG-ACC waters // puts the vase-into //
krumplit *hámoz* // *ablakot* *pucol.*
 potato-SG-ACC peels // window-SG-ACC cleans
 'Piri waters plants / is watering (the) plants // is putting flowers in the vase //
 peels potatoes / is peeling (the) potatoes // cleans windows / cleaning (the) windows.'
 b. *Piri megöntözte* [PRE-watered-OBJ] *a* [the] *virágot* [SG-ACC] // *a virágokat* [PL-ACC].
 'Piri has watered the plant // the plants'.

- (92) a. *Jani iskolába jár* // *börtönben* *van.*
 J. school-to goes-HABIT // jail-in is
 'Jani goes to school // is in jail.'
 b. *Villamossal megyek az iskolába.*
 tram-with go-I the school-to
 'I'll go to the school by tram.'

- (93) a. *Computerrel dolgozik.*
 computer-with works
 'She / He works with computers.'
 b. *Sok baja van a computerrel.*
 much / many trouble-his-SG is the
 'She / He has a lot of trouble with the computer.'

We may conclude that attributive use and foregrounding of QUALITY are strongly associated in the case of the preverbally occurring singular bare phrases. This association is characteristic of English and German compounds, while attributively used autonomous phrases are mainly associated with OBJECTNESS. It is worth noting that in Hungarian, in the subsequent discourse, anaphoric reference can be made to such QUALITY-foregrounded phrases, as illustrated in (94) a. (cf. 3.2.3). In other words, QUALITY-foregrounding is combined with spatio-temporal concreteness in such cases. In contrast to (94) a., the phrase *egy regényt* in (94) b. would imply a specific interpretation and the foregrounding of OBJECTNESS. In non-standard Hungarian, the use of an indefinite article can also be observed in the predicative position (cf. (95)). But whereas there is a clear semantic difference between

the presence and absence of the indefinite article in non-predicate positions, in predicate positions, we find only stylistic and social correlates.

- (94) a. *Regényt* vagy *kisregényt* írok. Egy epizódja Füreden
 novel-SG-ACC or short novels-SG-ACC write-I one episode-his F.-in
*történik.*⁷²
 happens
 'I am writing a novel or rather, a short novel. One of the episodes is set in Füred.'
- b. *Írok egy [a] regényt.*
 'I am writing a novel.'

- (95) *Ez egy szép ház // egy jó jel.*
 this a nice house // a good sign
 'This is a nice house // a good sign.'

3.3.2 SENSATION

As mentioned above, Hungarian does not have morphological restrictions for plural but only pragmatically-based ones. I am not familiar, for instance, with *gyűlöletek* ('hatred-PL'), and *szeretetek* ('love-PL') (cf. (96)). I would not, however, be surprised at seeing them in an appropriate context. Generally, plural forms of 'abstract nouns' are more frequently used in Hungarian than in English or German (cf. (96)).

- (96) a. *örömök* ('joy-PL'), *félelmek* ('fear-PL'), *fájdalmak* ('pain-PL'), *?gyűlöletek* ('hatred-PL'),
?szeretetek ('love-PL'), *bánatok* ('sorrow-PL'), *szenvetések* ('suffering-PL', 'sorrow-PL', 'pain-PL'), *bosszúságok* ('annoyance-PL, anger-PL')
- b. *A könyvekben emberekről volt szó... és örömekről és szenvedésekről.*
 the books-in human-PL-about was about... and joy-PL-about and suffering-PL-about
 'The books were about people... and about joy and suffering.'

Lexical units in the field of mental SENSATION participate in the same alternation pattern between preverbal phrase without article and postverbal phrase with a definite article, as described above (cf. (97)). Here, it makes no difference whether or not the phrase in question

⁷² Taken from Örkény (1992: 118).

is a subject (cf. (97) a., b., c. and d.) or an adverbial phrase ('source' or 'goal') (cf. (97) e. and f.). 'Classical generics', i.e., generics which are constructed as subjects of generic propositions, rather than as subjects ofthetic propositions as in (97) a. and b., require a definite article in the domain of 'abstract nouns'. There is some evidence that the postverbally-used subject in athetic proposition is identical with the topical subject in a generic proposition: In both cases, reference is made to the name of an established abstract concept. Insofar as the phrase in question contains a modifying adjective, as in (97) a., phrases without an article are preferred in the preverbal position.

- (97) a. *Szörnyű izgalom ragadta el.*
 terrible excitement overcame-OBJ PRE
 'A great excitement overcame her / him.'
- b. *Elragadta az [the] izgalom.*
 'Excitement overcame her / him.'
- c. *Arcán öröm ragyogott.*
 face-her/his-on joy glown
 'Her / His face was aglow with joy.'
- d. *Ragyogott az [the] arcán az öröm.*
- e. *Mari / Mari arca [face-her] ragyogott az örömtől [joy-from].*
 'Mari / Mari's face was aglow with joy.'
- f. *Arca örömrre [joy-into] derült [lit (up)]*
 'Her / His face lit with joy.'

Indefinite articles⁷³ are not used in this domain at all; we find, however, numerals as shown in (98).

- (98) *Egy [one] bánatom [sorrow-my] // örömöm [joy-my] ... van [is].*
 'I have only one sorrow // only one joy ...'

3.3.3 DRINKS and Co.

In Hungarian, lexical forms associated with a sense denoting DRINKS can be used (a) in bare phrases (cf. (99) e), (b) in phrases modified by a quantifier like *egy kis* ('a little') (cf. (99) f.), (c) in phrases directly modified by a numeral (cf. (99) a. and b.) and (d) in phrases modified by a numeral and a 'CONTAINER noun' (cf. (99) c. and d.). Note that 'CONTAINER nouns' syntactically behave as modifying elements in Hungarian rather than as heads: It is the 'DRINK

⁷³ As in German, the indefinite article (*egy*) is only prosodically distinguished from the numeral *egy* ('one').

noun' which is inflected for CASE. The acceptability of the use of numerals without an explicit expression for CONTAINER depends on the existence of a default CONTAINER or default unit of measure and the situational context (restaurant, at home, etc.) as in German.

Some of the lexical units are strongly conventionalized for use without a 'CONTAINER noun' in the restaurant context (see *két kávé* (99) a.). With these, a default CONTAINER is explicitly mentioned, chiefly, as long as it is contrastively focused as in (99) c. Other lexical units lacking a default CONTAINER, like *bor* ('wine'), regularly occur with an explicit expression for a CONTAINER, as shown in (99) d. Since the default CONTAINER varies depending on the situational context, the acceptability of the same construction varies, too (see, for instance, (99) b. which is clearly better in a store context rather than in a restaurant context). In contrast to German, all lexical units can be and are frequently used in phrases directly modified by a numeral for SORT interpretations (compare *bor* again in (100) a. and b.; note that *két bor(t)* is, in both sentences, ambiguous about PORTION and SORT reading). One of the reasons for this is very simple: Hungarian lacks NUMBER agreement and morphological plural restrictions. The other reason is that Hungarian lexical units in the nominal domain are basically characterized by the ability to be used in phrases which refer to SORTS, just as German nouns are generally characterized for generic use.

We have seen that in German, three alternative constructions may indicate SORT interpretations: (a) an explicit SORT expression constitutes, as a 'free form', the head of the phrasal construction (appositive, genitive or prepositional construction as in *zwei* ('two') *Sorten* ('sorts') *Wein* ('wine')), (b) an explicit SORT expression constitutes the 'lexical head' of a compound (as in *zwei Weinsorten*), and, in the case of selected lexical units, (c) numerals and quantifiers are directly combined with the appropriate word-form of the lexical unit in question. In Hungarian, the (c)-construction alternates with a construction in which an explicit SORT expression appears as an affix appended to the numerals and quantifiers rather than to the 'content words'. Thus, the Hungarian phrases *két* ('two') *bor* (cf. (100) a. and b.) and *sok* ('many') *bor* (cf. (100) c.) are, in one reading (the SORT reading), synonymous to *kétféle* ('two-SORT-ADJ') *bor* and *sokféle* ('many-SORT-ADJ') *bor* (cf. (100) d.). The ability of all numerals and a great number of quantifiers to take a SORT affix establishes a paradigmatic row in the grammatical apparatus of quantification (*ezer* ('thousand') > *ezerféle*, *minden* ('all') > *mindenféle*, etc.) which is located in the dimension of taxonomy. This is the second instance of paradigmatic organization in the domain of quantification; the first was, as we remember, the paradigmatic organization of 'question words' according to the grammatical distinction between distributivity and collectivity. As was the case with the latter, the recurrent

grammatical contrast achieved by the affixation of *-féle* is likewise not paired with lexical sensitivity.⁷⁴

- (99) a. *Két sör* // *Két kávé* // *?Két bort* *kérünk.*
 Two beer-SG-ACC // coffee-SG-ACC // wine-SG-ACC ask-we
 'Two beers // Two cups of coffee // Two glasses of wine, please.'
- b. *Két tejet* [milk-SG-ACC] *kérünk.*
 'Two milks, please.'
- c. *Két pohár* [glass] *sört* *kérünk.* *Nem* [not] *két korsót* [mug-ACC].
 'Two glasses of beer, please. Not two mugs.'
- d. *Két üveg* [bottle] *bort* // *két pohár* [glass] *bort* *kérünk.*
 'Two bottles of wine // two glasses of wine, please.'
- e. *Sört* // *kávé* // *bort* // *tejet* *kérünk.*
 'Some beer // some coffee // some wine // some milk, please.'
- f. *Kérek* [ask-I] *egy kis* [a little] *sört* // *egy kis kávé* // *egy kis bort* // *egy kis tejet.*
 'I would like (to have) a little beer // a little coffee // a little wine // a little milk.'
- (100) a. *Hoztam két bort.*
 brought-I two wine-ACC
 'I have brought two sorts of wine / two bottles of wine.'
- b. *Minda két bor nagyon finom.*
 all the two wine very delicious.
 'Both wines are delicious (i.e., two sorts of wine, or two portions of wine).'
- c. *Finom badacsonyi* [B.-ADJ] *borokat* [PL-ACC] // *Sok* [many] *finom badacsonyi*
bort [SG-ACC] *lehet ott kapni* [one can buy there].
 'One can buy excellent wines from Badacsony there.' // 'One can buy many
 excellent wines from Badacsony there.'
- d. *Kétféle* [two-sort-ADJ] // *Sokféle* [many-sort-ADJ] *badacsonyi bort* *lehet ott kapni.*
 'One can buy two sorts of wine from Badacsony // many sorts of wine from
 Badacsony there.'

When numerals and quantifiers are not explicitly mentioned, the use of plural forms for signaling the multiplicity of SORTS (cf. (100) c.) and PORTIONS is very common. I have cursorily tested the occurrences of PORTION reference in phrases with and without an explicit

⁷⁴ There are probably frequency differences or finer contextual differences concerning the synonymy relation of constructions containing *sok* vs. *sokféle* not accessible to native intuition. This would be a good task for further corpus research.

expression for CONTAINER in an Hungarian novel.⁷⁵ The results confirm the assumption made above about the importance of default CONTAINERS (compare (99) a. with (101) c., an example found in the novel). In the case of lexical units denoting DRINKS and connected with a default CONTAINER, a number of plural phrases lacking a CONTAINER expression are also attested (cf. (101) a. and b.). In all examples found in the novel, such plural phrases focus on the distributive multiplicity of PORTIONS.

- (101) a. *Ő a napközben lenyelt konyakoktól és részben a vodkától, amit én vettem elő, már igencsak fáradt volt.*
 He the in the course of the day gulped down cognac-PL-from and partly the vodka-SG-from which I got out PRE already very tired was
 'He was already very tired from the numerous glasses of cognac he had gulped down, and partly, from the vodka which I got out.'
- b. *Végigitta a szokásos vasárnapi fröccsöket.*
 drank through the usual Sunday-ADJ
 drink consisting of wine and soda-PL-ACC
 'As usual on Sundays, he drank his way through his normal number of glasses of 'fröccs'.'
- c. *Kért egy kávét (a hatodikát), de nem itta meg.*
 asked a coffee-ACC (the sixth-ACC) but not drank-OBJ PRE
 'He ordered a coffee (his sixth one) but did not drink it.'

We can summarize by saying that morphological plural, as marker of distributive multiplicity, is not biased toward spatio-temporally concrete reference or to reference to UNITS instead of SORTS; consequently, pluralization is independent from the lexical specification in the dimension of form, i.e., from the question as to whether or not the SHAPE of the entities denoted is lexicalized as a dispositional property.

3.3.4 Grinder effects

One could hypothesize that Hungarian lexical units in the nominal domain are generally unspecified with respect to FORM/SHAPE as dispositional property. However, as we will see in this section, this hypothesis turns out as wrong. The best way to show this is to investigate

⁷⁵ The novel is *Zakariás*, by Ákos Kertész (1990). It is the life history of an alcoholic actor and, consequently, contains a great number of relevant phrases.

contexts in which we can observe 'grinder effects' in MASS environments in English (cf. (25)) and German (cf. 3.2.4).

In Hungarian, the sense alternation between ANIMAL + ANIMAL USED AS FOOD is much more restricted than in English. The lexical form which is regularly used for ANIMALS can be used for the corresponding FOOD only in explicit 'FOOD names', for instance on menu cards. In other contexts, a compound occurs which is headed by the lexical form *hús* ('meat') (cf. *bárány* and *bárányhús* in (102) c.). *Hal* ('fish'), lacking a corresponding compound for FOODS, is an exception. Hungarian has a complex quantifier *egy kis* which has the same compositional and ambiguity structure as the English *a little* (i.e., *some* vs. *a small*). Applied to *hal*, *egy kis* evokes the FOOD sense (cf. (102) a.); the phrasal interpretation 'a small fish' is blocked in the sentence structure of (102) a., if uttered with a neutral intonation. That is, the interpretation 'indefinite article + adjective' is possible for this type of ambiguous lexical forms only if the phrase in question occupies the focus position, as in (102) b., and the adjective *kis* bears a contrastive stress. Otherwise, a different lexical form (*kicsi* ('small')) has to be used. In other words, the use of *egy kis* does not lead to sentence ambiguity. In everyday speech, the quantifier *egy kis* never modifies the other lexical forms which primarily denote FOODS (*bárány*, etc., cf. (102) c.); here, the compound form has to be used. Combined with *alma* ('apple'), *egy kis* can be understood as a quantifier, but the phrasal interpretation is per default 'a small number of intact apples' rather than 'a small number of apple slices, etc.' (cf. (102) d. vs. e.). The reason for this is the following: *Alma*, like *bárány*, has a lexicalized sense which contains information about the FORM/SHAPE of the discrete objects to which corresponding phrases refer. The quantifier *egy kis* has the semantic role of collective quantification without evoking a shift to a 'grinded' SUBSTANCE. Only if such a sense is already lexically established, as in the case of *hal*, does the construction *egy kis X* make a reference to a formless SUBSTANCE. Otherwise, a partitive construction is preferred for achieving this semantic effect (cf. (102) d.).

- (102) a. *Adjál egy kis halat.*
 give a little fish-ACC
 'Give me a little fish.'
- b. *Egy kis / kicsi halat adjál.*
 'Give me a small fish.'
- c. *?Adjál egy kis bárányt [lamb-ACC]. > Adjál egy kis bárányhúst [lamb-meat-ACC].*
 'Give me a little lamb-meat.'
- d. *?Adjál egy kis almát [apple-ACC]. > Adjál egy kicsit [little-ACC]
az almából [apple-from].
 'Give me some pieces of apple.'*

e. *Adjál egy kis almát.*

'Give me some apples.'

The Hungarian lexical form *üveg* displays an ambiguity between MATERIAL + ARTIFACTS MADE OF MATERIAL similar to the English *glass* and the German *Glas*: The respective senses are, however, 'glass' and 'bottle'. As expected, most grammatical contexts do not disambiguate, i.e., quantifiers like *sok* ('much', 'many'), numerals, morphological plurals and the absence of a determiner/quantifier permit both interpretations (cf. the predicatively used bare phrase in (103) a.). A determiner or quantifier is obligatory only before measure or unit nouns used as free word-forms, as before the unit noun *darab* in (103) b.⁷⁶ Here, we encounter the same type of lexical sensitivity for a grammatical distinction (i.e., a lexical distinction between measure and unit nouns on the one hand, and all other lexical units on the other hand) as in the case of 'question words'. This is actually the only case in which 'countability', in the strict sense, plays a role in the Hungarian lexicon.

(103) a. *Ez üveg.*

this glass / bottle

'This is glass / a bottle.'

b. *Ez egy [a] darab [piece] üveg / (egy) üvegdarab.*

'This is a piece of glass.'

It is well-known that languages show significant cross-linguistic differences with respect to lexical patterns related to the difference between the homogeneous collections of small and similar objects and single elements of such collections. Four patterns are usually found:

- (a) Simple lexical forms denote collections while compounding or other sorts of morphological derivation are applied for the generation of complex lexical forms with a singulative sense.
- (b) There are no lexical forms with an exclusively singulative interpretation, but periphrastic constructions consisting of a classifying unit noun and a collective noun are productively generated for singulative reference.

⁷⁶ Recall that it is not the measure/unit noun which is inflected for case in measure phrases but the lexical unit denoting SUBSTANCES/MATERIALS. We have considered this as the main argument for the head status of the latter, since all other common criteria for headedness (gender or number agreement, etc.) fail in Hungarian. However, we could probably consider the fact that absence of a determiner/quantifier is not allowed in just such a context as an evidence for the head status of the measure/unit nouns, with the consequence that nominals like *darab üveg* ('piece glass') or *üveg bor* ('bottle wine') have two different heads.

- (c) Simple lexical forms usually denote single objects; these lexical forms either constitute the base of derivational processes resulting in complex lexical forms with a collective sense or it is the plural word-form of these lexical forms which indicates collective interpretation on the phrasal level.
- (d) Simple lexical forms show ambiguity between the collective and singulative interpretation.

Hungarian chiefly opted for the first and fourth, English for the second and third pattern. The examples under (104) show transparent but conventionalized lexical units with the compound structure headed by the unit nouns *szem* ('grain', 'piece'), *szál* ('longish, thin object'), *csepp* ('drop'), etc.⁷⁷ I consider this fact, i.e. the recurrence of this pattern on the one hand, and the conventionalization of the complex forms on the other hand, as an evidence for the (positive or negative) lexical specification of the SHAPE of the entities denoted.

- (104) a. *szem*: *babszem* ('a bean'), *lencseszem* ('a lentil'), *borsószem* ('a pea'), *búzaszem* ('grain of wheat'), *kukoricaszem* ('grain of maize'), *szőlőszem* ('a grape'),
homokszem ('grain of sand'), *porszem* ('grain of dust')
- b. *szál*: *hajszál* ('a single hair'), *fűszál* ('blade of grass')
- c. *csepp*: *vízcsepp* ('water drop'), *esőcsepp* ('rain drop')

Some of the simple lexical forms allow an alternation between the collective and the singulative sense (cf. (105) a.), which is clearly a lexically marked property. If the complex form, which is positively specified in the dimension of form appears in a singular bare phrase, the lexically established conceptualization does not change (cf. (105) b.). This is because the use of a singular bare phrase is sensitive to the distinction between QUALITY and OBJECTNESS, but not to the distinction between FORM and SUBSTANCE, that is, only to one of the two components associated with the MASS/COUNT distinction in English. In Hungarian, QUALITY and FORM are indeed quite compatible.

- (105) a. *Ne vedd a fűvet / a fűszálat a szádba!*
 Don't take the grass-ACC/ blade of grass-ACC the mouth-your-into
 'Don't put the blade of grass into your mouth.'

⁷⁷ One may argue that there is no difference between the productive use of a 'qualifying' unit noun used as a free form and the productive use of corresponding affixes or compound heads. This is probably correct; in Hungarian, however, the latter construction, as seen in (104), is not entirely productive. Furthermore, the same unit nouns, when used as prenominal 'free words', can act as negative polarity items overruling 'natural lexical affinities' (cf. *egy* ('one') *szem férfi* ('man'), *egy csepp idő* ('time')).

- b. *Fűszállal csiklandozta az orromat.*
 blade of grass-with tickled the nose-my-ACC
 'He / She tickled my nose with a blade of grass.'

We receive the same results if we apply the test-frames for proving the possibility of foregrounding QUALITY vs. OBJECTNESS under (106). In discussing the German data, we have made the prediction that only those lexical units which allow foregrounding of QUALITY will be appropriate in this type of context. This is exactly the case in Hungarian: *Tojás*, *krumpli*, and *paradicsom*, despite their all having the lexical specification of FORM property, commonly occur in these test-frames in singular form (cf. (107) a.), independent of word-order variation, as seen in (107) b. and c. Neither of the sentences under (107) evoke the conceptualization of a 'mass of egg/potato/tomato'. The implication is always that the basket contains intact eggs, potatoes and tomatoes; that is, the singular form is neutral with respect to the number of entities involved and the lexical context biases, especially in (107) a., the interpretation of more than one discrete object. In (107) d. - in contrast to (107) c. - QUALITY is not foregrounded and the distributive multiplicity is made explicit.

- (106) Frames: *A kosár tele van X-szel.* ('The basket is full of X')
X van a kosárban. / A kosárban X van. ('X is in the basket' /
 'There is X in the basket')
 Lexical units: *tojás* ('egg'), *krumpli* ('potato'), *paradicsom* ('tomato')

- (107) a. *A kosár tele van tojással* [SG-with] // *paradicsommal* [SG-with] // *krumplival* [SG-with].
 'The basket is full of eggs // potatoes // tomatoes.'
 b. *Tojás* [SG] // *Krumpli* [SG] // *Paradicsom* [SG] *van a kosárban.*
 'Eggs // potatoes // tomatoes are in the basket.' / 'There are ...'
 c. *A kosárban tojás* [SG] // *krumpli* [SG] // *paradicsom* [SG] *van.*
 'There are eggs // potatoes // tomatoes in the basket.' / 'Eggs ...'
 d. *A kosárban tojások* [PL] // *krumplik* [PL] // *paradicsomok* [PL] *vannak* [are].
 'There are eggs // potatoes // tomatoes in the basket.'

Since the grammatical environment of singular bare phrases is not concerned with the lexical distinction established in the dimension of form, lexically sensitive systematic and/or conventionalized ambiguity between MATERIAL + NATURAL PIECE OF MATERIAL or MATERIAL + ARTIFACTS MADE OF MATERIAL will survive as sentence ambiguity unless the lexical context or knowledge of pragmatic likelihood selects one particular sense. The Hungarian lexical form *kréta* displays the often observed systematic ambiguity corresponding to the senses of *chalk* and *crayon*. The phrase *egy krétát* in (108) a. receives the interpretation as ARTIFACT,

due to the fact that the SORT interpretation of the SUBSTANCE is ruled out by the lexical context. The unit noun *darab* has, as mentioned, an affinity to the SUBSTANCE sense. In (108) b., however, which contains a singular phrase in the preverbal position, both interpretations are possible as shown by the English translation. Similarly, in (109) a. the means of payment can be (a) 'silver as homogeneous substance', (b) 'silver-ware' or (c) 'silver-coins', according to the ambiguity structure of *ezüst*. The (b) interpretation follows the pattern as illustrated in (109) b., where discrete lambs, apples and cars have to be conceived of as means of payment. Note that Hungarian has no way of expressing the semantic differences evoked by determiners, prepositions and NUMBER marking in the English translation of (109) b.

- (108) a. *Adj egy krétát // egy darab krétát!*
 give a chalk / crayon-ACC // piece
 'Give me a crayon // a piece of chalk.'
- b. *Az egész falat krétával mázolta be.*
 the entire wall chalk / crayon-with cover with paint PRE
 'She / He covered the entire wall with chalk.' / 'She / He scribbled over the entire wall with a crayon.'
- (109) a. *Ezüsttel fizetett.*
 silver-with paid
 'She / He paid in silver / in silver-ware / in silver coins...'
- b. *Báránnyal // almával // autóval fizetett.*
 lamb-with // apple-with // car-with
 'She / He paid with a lamb / in lamb // with an apple / in apples // with a car / in cars.'

The last examples underline the important difference between strong grammatical implication, i.e., implication triggered by grammatical contexts, and weak lexical or pragmatic implications. Since the grammatical context in (109) a. makes no implication and the lexical context (i.e., the predicate) does not show strong preferences in favor of a certain interpretation combined with *ezüst*, at least three sentence interpretations are possible. However, this fact presupposes a certain amount of lexical ambiguity.

4 Summary and Conclusion

In this study, I have investigated a special type of categorization problem which arises when categories cannot be exclusively assigned to either the lexicon or to the grammar but result

from the language-specific interactions of these two domains. I have tried to show that well-known categorization problems in the domain of the MASS/COUNT distinction, but also in the domains of ASPECT/AKTIONSART and of parts-of-speech, cannot be generally attributed to 'prototypicality of linguistic categories'. They rest, to a great extent, in the (conceptual) incompatibility of some basic concepts concerning the lexicon and its interaction with grammar, which I have traced back to Bloomfield. One of the odd ideas which developed in the long history of linguistic research was that one can abstract from complicated lexical structures, such as different types of ambiguities and productive lexical patterns, and classify lexical forms according to one particular salient distinction in a grammatical domain, with the expectation that the assigned classes would predict the behavior of all lexical forms in question in all contexts, i.e., that they would allow 'lexical insertion' on a high level of abstraction. According to another common linguistic concept, rooted in the Post-Bloomfieldian tradition, it was thought that formal variation could be studied independently from systematic semantic variation and vice versa, and likewise, that a lexical or grammatical category must be either a purely 'formal' category or a purely 'semantic' category.

The approach presented here is founded on the following basic ideas: Ambiguity and synonymy, as universal principles of natural language, characterize elements both of the lexicon and of the grammar (i.e., syntactic positions, 'function words', affixes, etc.). Instead of a rigid concept of a 'one-form-one-meaning principle' concerning static and abstract linguistic entities, a dynamic concept of form-meaning correspondence is adopted, which guides actual uses and native judgments. Consequently, form-meaning correspondences (rather than formal means or semantic primitives alone) must be the basic units of analysis in that all possible semantic interpretations in a certain grammatical context, and all possible grammatical contexts under a certain semantic interpretation have to be mutually examined. This approach is based on insights similar to those underlying the approach of syntacto-semantic verbal alternations developed in recent years. In this paper, it has been shown that, in the nominal domain, alternations mediating between lexicon and grammar are of similar importance. It has turned out to be extremely important and fruitful to connect the study of systematic ambiguities with that of the MASS/COUNT distinction. For modeling cross-linguistic aspects of the lexicon-grammar interaction, two logically independent parameters have been introduced: 'lexical sensitivity', responsible for lexical subregularities and lexical built-in preferences for grammatical distinction, and 'paradigmatic uniformity', responsible for paradigmatic patterning of grammatical meaning.

Three languages, English, German, and Hungarian, have been compared with respect to their lexicon-grammar interaction in the domain of quantification and determination. First, the influence of English on our understanding of the MASS/COUNT distinction has been studied: Alternative approaches have been evaluated according to different criteria and recurrent

problems, for instance those with so-called 'dual nouns', have been systematized. Here, special attention was paid to arguments for 'meaning shifts' and the 'basic' vs. 'secondary' status of specific senses and dimensions. As a necessary prerequisite to the examination of the German and Hungarian data, the different dimensions (the dimensions of referentiality, objectness, spatio-temporally concreteness, taxonomy, and form) in which the MASS/COUNT distinction can be dealt with have been disentangled. In a fine-grained analysis, different syntactic positions and lexical and situational contexts on the one hand, and different semantic/pragmatic implications on the other hand, have been taken into account.

The analysis in selected semantic fields in German has shown that each field can be characterized by its own specific lexical pattern with respect to determination and quantification. Nevertheless, within each semantic field, more than two lexical classes have to be distinguished. We have observed four different factors which cross-classify the lexicon and cannot be captured by the simple use of standard distinguishing criteria: (a) purely morphological factors, (b) lexically established ambiguity structures, (c) conventionalizing of pragmatic factors, and (d) idiosyncratic behavior, probably reflecting language change. As regards English and Hungarian, factors (b), (c), and (d) also have a cross-classificatory effect on the lexicon, but purely morphological factors are less important for the formation of lexical classes in English and are not at all important in Hungarian.

We have observed considerable differences between English and German on the one hand, and Hungarian on the other hand, concerning (a) the relation of lexical sensitivity and paradigmatic uniformity and (b) the affinity between values on the five dimensions mentioned. The grammatical organization of determiners/quantifiers and morphological NUMBER specification points, in English to a greater, in German to a lesser, extent, to a macroparadigm which can be interpreted as a binary distinction between MASS and COUNT values of a 'MASS/COUNT category'. In both languages, however, the respective macroparadigms do not satisfy the criterion of uniformity and are paired with lexical sensitivity. Consequently, the conclusion, drawn from the existence of 'dual nouns', namely, that the MASS/COUNT distinction has nothing to do with the lexicon, turns out to be wrong for these two languages. That is, a lexicon claimed to represent native-like knowledge and to predict grammatical behavior cannot be characterized by a binary division between MASS and COUNT classes but only by a series of classes, as mentioned above. In contrast, Hungarian has neither a unified paradigm in the MASS/COUNT domain nor lexical sensitivity among 'content words' for the recurrent grammatical distinctions found in independent subparadigms. This corresponds to the fact that the standard criteria for distinguishing MASS and COUNT nouns in Western languages cannot be applied here. The most significant difference between languages like English and German on the one hand and Hungarian on the other is the fact that in the latter, the values (a) FORM vs. SUBSTANCE and (b) OBJECTNESS vs. QUALITY are located in two

independent dimensions and are not implicationally connected. The lack of 'grinder effects' and the lack of a non-specifically used bare plural exemplify exactly this independence. Furthermore, Hungarian has a strong formal and semantic correlation between attributive use and foregrounding of QUALITY, which can only be observed in a similarly straightforward way with compounds in English and German. In this sense, Hungarian has no MASS/COUNT category at any level of linguistic description and the complex notion of 'individuation' used for the characterization of the English MASS/COUNT distinction has no unitary grammatical correlate. Nevertheless, a comparison on multidimensional levels is possible and this allows interesting typological generalizations, for instance, in the area of morphologically complex unitary concepts.

My conclusion is that a MASS/COUNT category, such as that found in English, cannot be considered as a universal category at all, either as a grammatical one or as a semantic category. Rather, the semantic ingredients of this language-specific category can be expected to be conflated and integrated in different ways in the lexico-grammar of other languages.

In general, it has become clear that none of the presented strategies for mapping between lexicon and grammar (multiple categories, mixed categories, dynamic category change, ellipsis, change of grammar) can be seen as superior and as providing a general solution for problems in mapping in a certain domain for all languages. Their adequacy depends heavily on the language-specific optimization based on syntagmatic and paradigmatic regularities, which ultimately determines the distribution of ambiguity/indeterminacy between lexicon and grammar.

There are two tasks which I consider important for future research: Corpus research provides new methods for determining the primary vs. secondary status of lexical senses and dimensions in addition to old ones based on linguistic tests. Exploiting these new methodological resources would be an important step in avoiding philosophical and not necessarily correct foundations of 'meaning shift'. My second suggestion concerns the interaction of different alternations. Whereas recursivity is well studied in the domain of morphological processes, we do not know very much about the recursivity of productive semantic alternations which are not morphologically marked. This study has unraveled some clear indications for the existence of such recursivity.

References:

- Akiyama, M. Michael / Wilcox, Sharon A. (1993), Naming as a Function of Linguistic Form-Class and Object Categories. In: *Journal of Child Language* 20(2), 419-35.
- Allan, Keith (1980), Nouns and countability. In: *Language* 56(3), 541-567.
- Basili, Roberto / Pazienza, Maria T. / Velardi, Paola (1992), Computational Lexicons: the Neat Examples and the Odd Exemplars. In: *Proceedings of the Third Conference on Applied Natural Language Processing*. Trento 1992, 96-102.
- Bäuerle, Rainer (1994), Zustand - Prozeß - Ereignis. Zur Kategorisierung von Verb(al)phrasen. In: *Wuppertaler Arbeitspapiere zur Sprachwissenschaft 10*. Wuppertal: Bergische Universität-Gesamthochschule Wuppertal.
- Behrens, Leila (1982), Zur funktionalen Motivation der Wortstellung: Untersuchungen anhand des Ungarischen. München: Veröffentlichungen des Finnisch-Ugrischen Seminars an der Universität München.
- Behrens, Leila (1990), Bemerkungen zur Aspektklassifizierung. In: *CIS-Bericht-90-20*. München.
- Behrens, Leila (1995), Alternationen - ein Schlüssel für die Universalienforschung. In: *Zeitschrift für Sprachwissenschaft* 13(2), 149-200.
- Bergenholtz, Henning (1983), Grammatik im Wörterbuch: Wortarten. In: *Studien zur neuhochdeutschen Lexikographie IV*.
- Bierwisch, Manfred (1983), Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In: R. Mutsch, W. (eds.). *Untersuchungen zur Semantik*. Berlin: Akademie-Verlag, 61-69.
- Birkenmaier, Willy (1980), Die Kategorie 'zählbar - nichtzählbar' beim russischen Substantiv. In: *International Review of Applied Linguistics in Language Teaching* 18(1), 31-40.
- Bloomfield, Leonard (1914), *An Introduction to the Study of Language*. London: Bell.
- Bloomfield, Leonard (1933/1973), *Language*. London: George Allen & Unwin LTD.
- Boguraev, Branimir / Briscoe, Ted (1989), *Computational lexicography for natural language processing*. London / New York: Longman.
- Boguraev, Branimir (1991), The Contribution of Computers. In: Boguraev, Branimir (ed.), *Building a Lexicon*. *International Journal of Lexicography* 4(3) (Special Issue), 227-260.
- Brown, E.K. / Miller, J.E. (1980), *Syntax: A Linguistic Introduction to Sentence Structure*. London et al.: Hutchinson.
- Bunt, Harry C. (1985), *Mass terms and model-theoretic semantics*. Cambridge et al.: Cambridge University Press.
- Bußmann, Hadumod (1990), *Lexikon der Sprachwissenschaft*. 2. völlig neu bearb. Aufl. Stuttgart: Kröner.
- Bybee, Joan L. (1985), *Morphology*. *Typological Studies in Language*, Vol. 9. Amsterdam / Philadelphia: John Benjamins B.V.
- Bybee, Joan L. / Pagliuca, William (1985), Cross-linguistic comparison and the development of grammatical meaning. In: Fisiak, Jacek (ed.), *Historical Semantics; Historical Word-Formation*. Berlin: Mouton, 59-83.

- Cao, Xuân Hao (1988), The Count/Mass Distinction in Vietnamese and the Concept of 'Classifier'. In: *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 41(1), 38-47.
- Carlson, Lauri (1981), Aspect and Quantification. In: Tedeschi, Philip J. / Zaenen Annie (eds.), *Tense and Aspect. Syntax and Semantics* 14. New York et al.: Academic Press.
- Carroll, John B. (1978), Continuous/Discrete: A Reinterpretation of the Mass/Count Feature of English Common Nouns. In: Jazayery, Mohammad Ali / Polomé, Edgar C. / Winter, Werner (eds.), *Linguistic and Literary Studies in Honor of Archibald A. Hill, vol. II: Descriptive Linguistics*. The Hague: Mouton, 19-30.
- Comrie, Bernard (1976), *Aspect: an introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.
- Corrigan, Roberta (1989), Introduction: Linguistic and non-linguistic categorization: Structure and process. In: Corrigan, Roberta / Eckman, Fred / Noonan Michael (eds.), *Linguistic categorization*. Amsterdam / Philadelphia: John Benjamins, 1-28.
- Cruse, D. Alan (1986), *Lexical Semantics*. Cambridge: Cambridge University Press.
- Dowty, David R. (1979) *Word Meaning and Montague Grammar*. Dordrecht: D. Reidel.
- Dowty, David R. (1991), Thematic Proto-Roles and Argument Selection. In: *Language* 67(3), 547-619.
- Du Bois, John W. (1980), Beyond Definiteness: The Trace of Identity in Discourse. In: Chafe, Wallace L. (ed.), *The Pear Stories. Cultural, and Linguistic Aspects of Narrative Production*. Norwood, New Jersey: Ablex.
- Franks, Bradley (1991), Sense Generation and Concept Combination. In: *DYANA R2.4*, 73-138.
- Franks, Bradley / Braisby, Nick (1991), Sense Generation. In: *DYANA R2.4*, 56-71.
- Fries, C.C. (1954), Meaning and linguistic analysis. In: *Language* 30, 57-68.
- Gathercole, Virginia C. (1985), 'He has too much hard questions': the acquisition of the linguistic mass-count distinction in *much* and *many*. In: *Journal of Child Language* 12(2), 395-415.
- Geeraerts, Dirk (1989), Introduction: Prospects and Problems of Prototype Theory In: *Linguistics* 27, 587-612.
- Gillon, Brendan S. (1992), Towards a Common Semantics for English Count and Mass Nouns. In: *Linguistics and Philosophy* 15(6), 597-639.
- Gordon, Peter (1985), Evaluating the semantic categories hypothesis: The case of the count/mass distinction. In: *Cognition* 20, 209-242.
- Gordon, Peter (1988), Count/Mass Category Acquisition: Distributional Distinctions in Children's Speech. In: *Journal of Child Language* 15(1), 109-128.
- Harris, Zellig S. (1951/1960), *Structural Linguistics (Formerly Entitled: Methods in Structural Linguistics (1951))*. Chicago / London: The University of Chicago Press.
- Harweg, Roland (1987), Stoffnamen und Gattungsnamen. In: *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 40(6), 792-804.
- Hendrikse, Rusandré (1993), Review: Frans Plank (ed.): *Paradigms: The Economy of Inflection*. Berlin: Mouton de Gruyter, 1991, x, 317 pp. In: *Cognitive Linguistics* 4-2, 183-190.
- Hopper, Paul / Thompson, Sandra (1984), The Discourse Basis for Lexical Categories in Universal Grammar. In: *Language* 60, 703-752.

- Householder, F. W. (1973), On arguments from asterisks. In: *Foundations of Language* 10, 365-376.
- Jackendoff, Ray S. (1983/1988), *Semantics and Cognition*. Cambridge, Ma / London: The MIT Press.
- Jespersen, Otto (1925), *The Philosophy of Grammar*. London: Allen & Unwin.
- Kilgarriff, Adam (1993), Inheriting Verb Alternations. In: *Proceedings of the Sixth Conference of the European Chapter of the Association for Computational Linguistics*. Utrecht 1993, 213-221.
- Krifka, Manfred (1991), Massennomina. In: von Stechow, Arnim / Wunderlich, Dieter (eds.), *Semantik. Ein internationales Handbuch der zeitgenössischen Forschung. / Semantics. An International Handbook of Contemporary Research*. Berlin / New York: de Gruyter, 399-417.
- Langacker, Ronald W. (1987a), Nouns and Verbs. In: *Language* 63(1), 53-94.
- Langacker, Ronald W. (1987b), *Foundations of cognitive grammar I*. Stanford, CA: Stanford University Press.
- Langacker, Ronald W. (1991), *Foundations of cognitive grammar II*. Stanford, CA: Stanford University Press.
- Lehrer, Adrienne (1987), English quantifiers from noun sources. In: Steele, Ross / Threadgold, Terry (eds.), *Language Topics. Essays in honour of Michael Halliday*. Amsterdam / Philadelphia: John Benjamins, 94-105.
- Levin, Beth (1993), *English Verb Classes and Alternations. A Preliminary Investigation*. Chicago / London: The University of Chicago Press.
- Levy, Yonata (1988), On the Early Learning of Formal Grammatical Systems: Evidence from Studies of the Acquisition of Gender and Countability. In: *Journal of Child Language* 15(1), 179-187.
- Link, Godehard (1991), Plural. In: von Stechow, Arnim / Wunderlich, Dieter (eds.), *Semantik. Ein internationales Handbuch der zeitgenössischen Forschung. / Semantics. An International Handbook of Contemporary Research*. Berlin / New York: de Gruyter, 418-440.
- Lønning, Jan Tore (1987), Mass Terms and Quantification. In: *Linguistics and Philosophy* 10(1), 1-52.
- Lyons, John (1977), *Semantics*. Cambridge et al.: Cambridge University Press.
- Lyons, John (1991), Bedeutungstheorien. In: von Stechow, Arnim / Wunderlich, Dieter (eds.), *Semantik. Ein internationales Handbuch der zeitgenössischen Forschung. / Semantics. An International Handbook of Contemporary Research*. Berlin / New York: de Gruyter, 1-24.
- MacWhinney, Brian (1989), Competition and lexical categorization. In: Corrigan, Roberta / Eckman, Fred / Noonan Michael (eds.), *Linguistic categorization*. Amsterdam / Philadelphia: John Benjamins, 195-241.
- Matthews, Peter H. (1974), *Morphology. An Introduction to the Theory of Word-Structure*. Cambridge et al.: Cambridge University Press.
- Matthews, Peter H. (1993), *Grammatical Theory in the United States from Bloomfield to Chomsky*. Cambridge et al.: Cambridge University Press.
- McPherson, Leslie Maggie Perrin (1991), A Little Goes a Long Way: Evidence for a Perceptual Basis of Learning for the Noun Categories COUNT and MASS. In: *Journal of Child Language* 18(2), 315-38.
- Meulen, Alice G. B. ter (1984), Events, Quantities and Individuals. In: Landman, Fred / Veltman, Frank (eds.), *Varieties of Formal Semantics*. Dordrecht: Foris, 259-280.
- Miller, James (1985), *Semantics and Syntax. Parallels and Connections*. Cambridge et al.: Cambridge University Press.

- Mithun, Marianne (1984), The evolution of noun incorporation. In: *Language* 60(4), 847-894.
- Morrow, Daniel G. (1986), Grammatical Morphemes and Conceptual Structure in Discourse Processing. In: *Cognitive Science* 10, 423-455.
- Mourelatos, Alexander P.D. (1981), Events, Processes, and States. In: Tedeschi, Philip J. / Zaenen, Annie (eds.), *Tense and Aspect (Syntax and Semantics 14)*. New York et al.: Academic Press.
- Mufwene, Salikoko S. (1984), The Count/Mass Distinction and the English Lexicon. In: Testen, David / Mishra, Veena / Drogo, Joseph (eds.), *Papers from the Parasession on Lexical Semantics*. Chicago: Chicago Linguistic Society, 200-221.
- Newmeyer, Frederick J. (1994), Competing motivations and synchronic analysis. In: *STUF* 4, 67-77.
- Pawley, Andrew (1985), Lexicalization. In: *Languages and Linguistics: The Interdependence of Theory, Data, and Application (GURT 85)*. Georgetown University Press.
- Pelletier, Francis J. / Schubert, Lenhart K. (1989), Mass expression. In: Gabbay, D. / Guentner, F. (eds.), *Handbook of Philosophical Logic, Vol. IV*. Dordrecht et al.: D. Reidel, 327-407.
- Pelletier, Francis J. (1975), Non-Singular Reference. Some Preliminaries. In: *Philosophia* 5.
- Pinkal, Manfred (1985), *Logik und Lexikon. Die Semantik des Unbestimmten*. Berlin / New York: de Gruyter.
- Pinker, Steven (1989), *Learnability and Cognition*. Cambridge, Ma / London: The MIT Press.
- Pustejovsky, James (1991), The Generative Lexicon. In: *Computational Linguistics* 17(3).
- Quine, W.V.O. (1960), *Word and Objects*. Cambridge, MA: MIT Press.
- Quirk, Randolph / Greenbaum, Sidney / Leech, Geoffrey / Svartvik, Jan (1972), *A Grammar of Contemporary English*. London: Longman.
- Sampson, Geoffrey (1992), Probabilistic parsing. In: Svartvik, J. (ed.), *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlin / New York: Mouton / de Gruyter, 425-447.
- Sasse, Hans-Jürgen (1991), Aspect and Aktionsart. In: Vetters, Carl / Vandeweghe, Willy (eds.) *Perspectives on Aspect and Aktionsart*. (Belgian Journal of Linguistics 6). Bruxelles: Editions de l'Université de Bruxelles, 31-45.
- Sasse, Hans-Jürgen (1993), Das Nomen - eine universelle Kategorie? In: *STUF* 3, 187-221.
- Sasse, Hans-Jürgen (1995), "Theticity" and VS Order: A Case Study. In: Sasse, Hans-Jürgen / Matras, Yaron (eds.), *Verb-Subject Order and Theticity in European Languages*. Berlin: Akademie-Verlag.
- Searle, John R. (1969), *Speech Acts: an essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Shapiro, Lewis P. / Zurif, Edgar / Carey, Susan / Grossman, Murray (1989), Comprehension of Lexical Subcategory Distinctions by Aphasic Patients: Proper/Common and Mass/Count Nouns. In: *Journal of Speech and Hearing Research* 32(3), 481-488.
- Sharvy, Richard (1978), Maybe English has no count nouns: Notes on Chinese semantics. In: *Studies in Language* 2, 345-365.

- Smajda, A. Frank / MecKeown Kathleen R. (1990), Automatically Extracting and Representing Collocations for Language Generation. In: Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics, 6-9 June 1990. Pittsburgh: ACL.
- Soja, Nancy N. / Carey, Susan / Spelke, Elizabeth S. (1991), Ontological Categories Guide Young Children's Inductions of Word Meaning: Object Terms and Substance Terms. In: *Cognition* 38(2), 179-211.
- Talmy, Leonard (1985), Lexicalization Patterns: Semantic Structure in Lexical Form. In: Shopen, T. (ed.): *Language Typology and Syntactic Description 3*. Cambridge: Cambridge University Press, 57-149.
- Thompson, Sandra (1989), A discourse approach to the cross-linguistic category 'Adjective'. In: Corrigan, Roberta / Eckman, Fred / Noonan Michael (eds.), *Linguistic categorization*. Amsterdam / Philadelphia: John Benjamins, 245-265.
- Tsohatzidis, S.L. (1990), Introduction. In: Tsohatzidis, S.L. (eds.), *Meanings and Prototypes*. Studies in linguistic categorization. London/New York: Routledge, 1-13.
- Wierzbicka, Anna (1991), Semantic rules know no exceptions. In: *Studies in Language* 15-2, 371-398.
- Wilensky, Robert (1990), Extending the Lexicon by Exploiting Subregularities. In: Proceedings of COLING-90, Vol. 2. Helsinki 1990, 407-412.
- Wilks, Yorick (1992), Form and content in semantics. In: Rosner, Michael / Johnson, Roderick (eds.), *Computational linguistics and formal semantics*. Cambridge: Cambridge University Press, 257-282.
- Zelinsky-Wibbelt, Cornelia (1989), Machine Translation Based on Cognitive Linguistics: What Lexical Semantics Contributes to the Semantic Unity of a Sentence. In: Eurotra-D Working Papers Nr. 16.
- Zelinsky-Wibbelt, Cornelia (1991), Token Reference vs. Type Reference: Implications for Machine Translation. In: *Machine Translation* 6(3), 183-92.
- Zernik, Uri (ed.), *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*. Hillsdale, New Jersey: Lawrence Erlbaum.

Dictionaries / Grammars:

- COBUILD GRAMMAR, Sinclair John (ed.), (1992), Collins Cobuild. English Usage. London: Harper Collins Publishers.
- COLLINS, Hanks, Patrick (ed.) (1979), *Collins Dictionary of the English Language*. London / Glasgow: Collins.
- DUDEN, Drosdowski, G. / Wiss. Rat u.d. Mitarbeiter der Dudenredaktion (eds.) (1983), DUDEN. Deutsches Universalwörterbuch. Mannheim et al.: Dudenverlag.
- HALÁSZ, Halász, Előd (1992) (10., változatlan kiadás), *Magyar-Német Szótár*. Vol. 1./2. Budapest: Akadémiai Kiadó.
- LANGENSCHIEDT, Messinger, Heinz / Langenscheidt-Redaktion (eds.) (1986), *Langenscheidts Großwörterbuch der englischen und deutschen Sprache*. Berlin et al.: Langenscheidt.
- MNYÉSZ, A Magyar Tudományos Akadémia Nyelvtudományi Intézete (1979-1984), *A Magyar Nyelv Értelmező Szótára*. Vol. 1.-7. Budapest: Akadémiai Kiadó.

ORSZÁGH, Ország, László (1974), Magyar-Angol Szótár. Vol. 1./2. Budapest: Akadémiai Kiadó.

Text sources:

Kertész, Ákos (1990), *Zakariás*. Budapest: Magvető Kiadó.

Örkény, István (1992), *Levelek egypercben*. Budapest: Szépirodalmi Könyvkiadó.

Abbreviations in interlinear translation:

ACC accusative

OBJ 'objective conjugation' marking definite objects

PL plural

PRE verbal prefix

SG singular

1 INTRODUCTION.....	1
2 CATEGORIZATION BETWEEN LEXICON AND GRAMMAR.....	6
2.1 COMPLEX CATEGORIES AND THE PROTOTYPE APPROACH	6
2.2 PARADIGMS AND LEXICAL SENSITIVITY	8
2.2.1 <i>Paradigmatic Uniformity</i>	8
2.2.2 <i>Lexical Sensitivity</i>	11
2.2.2.1 Form-Classes, Grammatical Categories and Ambiguity Phenomena	12
2.2.2.1.1 Bloomfield.....	12
2.2.2.1.2 Post-Bloomfieldians	13
2.2.2.1.3 Contemporary Approaches	15
2.2.2.2 Some Heuristic Proposals	16
2.2.3 <i>Interrelations Between Paradigmatic Uniformity and Lexical Sensitivity</i>	19
2.3 LINGUISTIC STRATEGIES FOR MAPPING BETWEEN LEXICON AND GRAMMAR	20
2.3.1 <i>Adjectives and Nouns: A Brief Case Study</i>	20
2.3.2 <i>Lack of Necessary Criteria for Lexical Categories</i>	25
2.3.3 <i>Lack of Sufficient Criteria for Lexical Categories</i>	26
2.3.3.1 Strategy of Multiple Categories	26
2.3.3.2 Strategy of Mixed Categories	28
2.3.3.3 Strategy of Dynamic Category Change	29
2.3.3.4 Strategy of Ellipsis.....	30
2.3.3.5 Strategy of Grammar Change	31
3 MASS AND COUNT	31
3.1 ENGLISH AND ITS INFLUENCE ON THE MASS/COUNT THEORY	31
3.1.1 <i>Alternative Approaches</i>	31
3.1.2 <i>Evaluation of Linguistic Argumentation</i>	35
3.1.2.1 Presentation of the First Examples	36
3.1.2.2 Presentation of Criteria for Distinguishing Between MASS and COUNT Nouns.....	37
3.1.2.3 Nouns as Lexical Units vs. Nouns as Grammatical Forms	39
3.1.2.4 'Form-Classes/Lexical Classes' vs. 'Syntactic/Grammatical Categories', Denote/Denotation vs. Refer/Reference.....	40
3.1.2.5 Arguments for the 'Ungrammaticality' of Certain 'Forms'	42
3.1.2.6 Arguments for the 'Basic' vs. 'Secondary'/'Extended' Status of a Specific Sense or Dimension	45
3.1.3 <i>Five Dimensions in the MASS/COUNT Domain</i>	52
3.2 GERMAN	55
3.2.1 <i>DRINKS and Co.</i>	58

3.2.2 <i>MOTIVATION, CRIME, and SENSATION</i>	62
3.2.2.1 <i>MOTIVATION</i>	62
3.2.2.2 <i>CRIME</i>	64
3.2.2.3 <i>SENSATION</i>	67
3.2.3 <i>Morphologically Complex Unitary Concepts</i>	72
3.2.4 <i>Grinder Effects</i>	76
3.3 <i>HUNGARIAN</i>	80
3.3.1 <i>Morphologically Complex Unitary Concepts</i>	84
3.3.2 <i>SENSATION</i>	86
3.3.3 <i>DRINKS and Co.</i>	87
3.3.4 <i>Grinder effects</i>	90
4 SUMMARY AND CONCLUSION	95