

THEORETICAL ESSAY

THE NATURE OF COMPOUNDING

Pius ten HACKEN  

Institut für Translationswissenschaft – Leopold-Franzens-
Universität Innsbruck

ABSTRACT

This paper addresses the question of the definition of compounding from a terminological perspective. In terminology, concepts are defined by a selection of properties shared by prototypical cases. For scientific terminology, the selection is validated by the strength of the theories that can use the definition. It is shown that morphophonological criteria often adduced in the delimitation of compounding are not adequate in a universal definition. In order to come up with a better definition, a two-step procedure is proposed. In the first step, a universal definition is used to determine for constructions in a particular language whether they belong to compounding. In the second step, language-specific properties are used to identify instances of these constructions. A definition is proposed that takes a compound as a word with a binary, headed structure, a relation between the elements that is not determined by compounding and a non-head that is not introduced as an entity in the discourse. The use of this definition is illustrated with a number of constructions in different languages. It is shown that expressions commonly called *exocentric* and *copulative compounds* are generally not compounds in this definition, but that some expressions that have been labelled as such are in fact compounds. The two-step procedure demonstrated here for compounding can also be used for other linguistic terms.



OPEN ACCESS

EDITORS

– Miguel Oliveira, Jr. (UFAL)
– René Almeida (UFS)

REVIEWERS

– João Cyrino (UFBA)
– Emanuel Quadros (Yale)

DATES

– Received: 08/27/2020
– Accepted: 10/12/2020
– Published: 02/07/2021

HOW TO CITE

HACKEN, Pius ten (2021). The Nature of
Compounding. *Cadernos de
Linguística*, v. 2, n. 1, p. 01-21.

SAMENVATTING

In dit artikel wordt de vraag hoe *samenstelling* kan worden gedefinieerd vanuit een terminologisch perspectief behandeld. In de terminologie

worden concepten gedefinieerd op basis van een selectie van gemeenschappelijke eigenschappen van prototypische instanties. Voor wetenschappelijke terminologie wordt deze selectie gevalideerd door het succes van de theorieën waarin de definitie kan worden gebruikt. In het artikel wordt eerst aangetoond dat de morf fonologische criteria die vaak voor de afgrenzing van *samenstelling* worden gebruikt niet geschikt zijn voor een universeel toepasbare definitie. Om tot een betere definitie te komen, wordt een procedure met twee stadia voorgesteld. In het eerste stadium wordt de universeel geldige definitie gebruikt om te bepalen welke constructies in een taal samenstellingen vormen. In het tweede stadium worden taalspecifieke eigenschappen van deze constructies gebruikt om individuele uitdrukkingen te classificeren. Er wordt een definitie voorgesteld die een samenstelling als een binaire structuur met een hoofd karakteriseert, waarvan de semantische samenhang tussen de twee componenten niet door de constructie wordt bepaald en de component die niet het hoofd is niet referentieel is. De toepassing van deze definitie wordt geïllustreerd aan de hand van een aantal constructies in verschillende talen. Van uitdrukkingen die vaak *exocentrische* of *copulatieve samenstellingen* worden genoemd wordt aangetoond dat ze in het algemeen geen samenstellingen zijn, maar dat sommige van zulke uitdrukkingen in dit opzicht een uitzondering zijn. De procedure met twee stadia die hier gebruikt wordt leent zich ook goed voor diverse andere taalkundige termen.

KEYWORDS

Compounding; Terminology; Definition; Construction.

TREFWOORDEN

Samenstelling; Terminologie; Definitie; Constructie.

INTRODUCTION

In this paper, I will address the question of what is compounding. In pursuing this question, I will adopt a terminological approach. In terminology, it is common to start from a concept with rather indeterminate boundaries and work towards a more precisely determined concept that maintains the basic nature of the original idea. For compounding, I will start with some dictionary and textbook definitions and discuss some often-used formal criteria to make the boundary more specific. Then I will present my own approach, which takes as its point of departure that compounds are used first of all as names for concepts. I will show how this approach can work in practice and end with an evaluation of the definition and of the general approach it illustrates.

1. THE DEFINITION OF COMPOUNDING AS A PROBLEM

Compounding is the formation of compounds. For a general understanding of these concepts, we can consult dictionaries or look up the treatment in morphology textbooks. Some typical definitions are given in (1).

- (1) a. Compound: a word made up of two or more existing words
b. Compound: a word formed from two existing words or combining forms
c. Compounding is a word-formation process that combines two or more roots inside the same word.

The first two definitions are from general dictionaries of English. The word *compound* has several meanings and only the one for the relevant linguistic sense is quoted here. The definition in (1a) is from COED (2011), the one in (1b) from Collins (2014). Although the two definitions broadly agree, (1b) insists on binary branching and focuses on the formation process, whereas (1a) takes the resulting word as the starting point and takes two words only as a lower limit. The statement in (1c) is from Fábregas & Scalise (2012: 111). It is the first sentence of the chapter on compounding. Although it focuses on *compounding* rather than on *compound*, it is strikingly similar to the dictionary definitions, in particular (1a). The difference is the reference to *root* instead of *existing word*. A common property of such definitions is that they are quite vague. For the following discussion, I will take *symphony orchestra* as a typical example of a compound.

There have been various attempts in the linguistic literature to make the concept of compounding more precise. Compounds occur in many languages. It is natural to align

the concept across languages so that what is called *compound* in English corresponds to the concept in German, French, or Slovak. This does not mean that the translation of *symphony orchestra* into French is necessarily a French compound, but that when we discuss English compounding in an article written in French, the French translation of *compound*, *mot composé*, should identify the same set of items. This universal application is not entirely straightforward, because French discussions of *mots composés* take French examples as their starting point, whereas English discussions of compounding start from English examples. In evaluating proposed criteria, it is worth keeping in mind their application to different languages.

A frequently used criterion for compounding is stress. In (2), primary stress is indicated by an underline.

- (2) a. symphony orchestra
b. expensive orchestra

Whereas the compound in (2a) has a single primary stress on the first component, the A+N phrase in (2b) has two primary stresses, one for each word. Although there is more to say about the application of this criterion to English, its most immediate problem is that it cannot be generalized to languages such as Slovak, as illustrated in (3).

- (3) a. idem na obed ('I_go for lunch')
b. dovaril obed ('he_finished_preparing lunch')

Slovak has stress on the first syllable of a word. In (3b), each word has its own stress. In (3a), however, the preposition *na* takes the stress from the noun *obed* ('lunch'). Thus, the English-based stress criterion would make *na obed* ('for lunch') a compound, which is highly counterintuitive.

Whereas stress is a criterion that is universally applicable at least in principle, other criteria that have been proposed are not. In German, *Feld* ('field') is pronounced with a final /t/, because it is at the end of the word. This is called *Auslautverhärtung* ('final devoicing'). In the plural *Felder* ('fields'), the grapheme ⟨d⟩ is pronounced /d/, because it is followed by a vowel. In compounds, the first component is treated phonologically as a separate word. Therefore, *Feldarbeit* ('field work') is pronounced with a /t/ at the end of *Feld*, because the following vowel is separated by a word boundary. In English, this criterion cannot be used, because English does not have devoicing of word-final consonants. Similarly, in German *Arbeitsfeld* ('working field'), an -s- appears between the two components. This is not an inflectional ending of *Arbeit* ('work'), because, as a

feminine noun, *Arbeit* never has such an inflectional ending. It is called a *Fugenelement* ('linking element'). Although such linking elements are common (but not general) for compounds in German, they do not occur in English.

Another type of criterion concerns the morphological behaviour of the components. The effect of one such criterion can be seen in (4).

- (4) a. direttore musicale ('music director')
 b. direttori musicali ('music directors')

In (4), we have the singular and plural of an Italian combination of a noun and a relational adjective. As the English translations suggest, the meaning is conveyed by a clear compound in English. The plural marker is attached to the head in English, but in Italian both the noun and the adjective are marked as plural. For some linguists, e.g. Matthews (1974: 34-35), this is a reason to exclude (4) in Italian from compounding. As English does not have plural marking on adjectives, the criterion does not really apply to cases such as *musical instrument*.

The phonological and morphological criteria exemplified so far have the advantage that they generally give clear results as far as they apply, but they have a number of disadvantages. One disadvantage is that they are highly language-specific. The properties they address are often not relevant in all languages (e.g. final devoicing) or they have entirely different effects (e.g. stress). Moreover, even within a single language, they often apply only to some cases. German devoicing does not play a role in *Weltordnung* ('world order'), because *Welt* has a /t/ in all cases (cf. *Welte*, 'worlds'). Similarly, compounds such as *Feldarbeit* and *Weltordnung* do not have a linking element. From a terminological perspective, this makes these properties less useful as candidates for conditions, because they are only usable as sufficient, not as necessary conditions. Proper terminological definitions are composed of necessary conditions (cf. ten Hacken 2018).

In view of the problems with criteria of the type discussed so far, we can understand the assessments of the possibility of defining *compound* in (5).

- (5) a.[W]e cannot always make a clean distinction between compound words on the one hand and derived words or phrases on the other.
 b.[U]niversal definitions are not only theory-dependent [...] but also cross-linguistically never watertight—in many languages there are exceptions or fuzzy transitions to non-compounding[.]

In their introduction to the *Oxford Handbook of Compounding*, discussing the problem of defining *compound*, Lieber & Štekauer (2009: 4) make the remark in (5a). Given the context of a handbook, it seems an appropriate summary of a certain consensus in the field. In (5b), Dressler (2006: 24) goes further and implies some assumptions that are quite problematic from a terminological perspective. The start of (5b) suggests that theory-dependence should be evaluated negatively in terminological definitions. In fact, terms are necessarily theory-dependent and their definitions must be embedded in a theoretical framework. Therefore, the theory-dependence is inevitable and should not be subject to any positive or negative evaluation. In the second part, Dressler mentions the existence of exceptions, which implies a higher authority than the definition determining that an expression is actually a compound. However, in terminology, the boundaries of a concept are determined by the definition. Therefore, if we want to use *compound* as a theoretical concept and make claims about it, we should reject (5b) and interpret (5a) as an incitation to develop a proper definition.

2. A SEMANTIC PERSPECTIVE

As a starting point for a proper terminological definition, we can analyse a few typical examples of compounds and consider which properties of them we want to retain in a definition. A preliminary observation in this respect is that we should aim for a cross-linguistically valid definition. Without this aim, there are too many shared properties, many of which are due to accidental features of an individual language. On the basis of this decision, we can exclude criteria such as stress, which is subject to different language-specific constraints, for the primary identification of compounds. After all, compounding is not primarily a phonological phenomenon, but a type of word formation. Given that word formation is driven by the need to provide new names for concepts, a good starting point is the relationship between the form and the meaning.

As a first example, let us return to *symphony orchestra* and consider how the meaning of the word is related to its form. In accordance with the dictionary definitions in (1), *symphony orchestra* consists of two words. The contribution of the two components to the meaning of the compound is asymmetric, in the sense that a symphony orchestra is an orchestra. *Orchestra* is the head and *symphony* the non-head. The role of the non-head can be studied by contrasting it to paradigmatic alternatives. Table 1 gives the most frequent compounds with *orchestra* in COCA (2008-2020).

symphony orchestra	1560
chamber orchestra	262
youth orchestra	183
school orchestra	105
jazz orchestra	86
string orchestra	68
opera orchestra	51
baroque orchestra	39
festival orchestra	32

Table 1. Most frequent compounds with *orchestra* in COCA (2008-2020).

In *symphony orchestra*, the non-head indicates the type of music the orchestra plays. As such, it contrasts directly with *jazz orchestra* and *baroque orchestra*. In *youth orchestra*, it is a property of the players that is expressed by the non-head and in *festival orchestra*, the organization behind it. *School orchestra* is ambiguous between these two, as *school* can be interpreted as a criterion for membership or the organization behind it. In *opera orchestra*, we can see a parallel either to *symphony orchestra* or to *festival orchestra*, because such orchestras accompany operas and are attached to an opera house. The case of *chamber orchestra* is a bit more complex. OED (2000-2020 [2010]: *chamber* n.) defines it as “a small orchestra, typically comprising fewer than forty musicians”. We will come back to the relation with *chamber* below. At least in all other cases, we can say that an *X orchestra* is an orchestra that is characterized by *X*, where the nature of *X* determines how it can relate to *orchestra*.

As a next example, let us consider *piano trio*. This word has two senses, illustrated in (6) with quotations from OED (2000-2020 [2006]: *piano* n.²).

- (6) a. This evening at the Pavillion. Piano trios by Beethoven.
b. She was a talented pianist and played in a piano trio.

In (6a), *piano trio* refers to a piece of music, in (6b) to an ensemble. The relation between the two senses is a regular type of metonymy, which is found also for other kinds of chamber music and is in principle independent of compounding. It is obvious that *piano trio* refers to a kind of trio, i.e. three instruments. However, an ensemble of three pianos is not a piano trio. A piano trio consists of a piano, a violin and a cello. This is a meaning that is much more specific than can be derived from the components. Yet, the study by Smallman (1992) shows that it is an established genre within chamber music. It emerged in the 18th century and

composers writing well-known works for piano trio include Haydn, Beethoven, Brahms, Dvořák and Ravel.

Because of the large repertoire, there are specialized ensembles playing these works. A famous example is the Beaux Arts Trio, which gave its first performance in 1955 and its last in 2009. That it had an identity of its own is shown by the fact that individual players were replaced in the course of its history. As it started, it consisted of Menahem Pressler (b. 1923), piano, Daniel Guilet (1899-1990), violin, and Bernard Greenhouse (1916-2011), cello. When Daniel Guilet retired in 1969, he was succeeded by Isidore Cohen (1922-2005), but the name of the trio remained.

The fact that a piano trio consists of exactly the three instruments of piano, violin and cello cannot be explained when we start from the form *piano trio*. We must assume that the starting point was the concept. In the search for a name, the compound *piano trio* was selected. The meaning of *piano trio* is then the result of what ten Hacken (2019a: 64) calls *onomasiological coercion*. While applying to word formation in general, onomasiological coercion is particularly clear in compounding, because the relation between the two components needs to be determined. The effect is illustrated also when we compare *piano trio* with *clarinet trio* and *horn trio*. A clarinet trio consists of a piano, a clarinet and a cello and a horn trio of a piano, a violin and a horn. The composition of the trios is determined by what established itself as an ensemble in the history of musical tradition. The selection of the non-head can then be thought of as guided by the most salient or least expected instrument.

Although *piano trio* is an established name, we sometimes also find an alternative designation, leading to the synonyms in (7).

- (7) a. piano trio
b. trio for piano, violin and cello

Compared to the compound (7a), the expression in (7b) is much more explicit. First of all, it can only be used for a piece of music, not for an ensemble. Secondly, it specifies exactly which instruments constitute the trio. Whereas (7a) gives a name, (7b) is a description. (7b) is used mostly in the title of a piece of music. It is more formal than (7a) and considerably less frequent. BNC (2007) has 25 occurrences of (7a) and COCA (2008-2020) has 34. Neither has any occurrences of (7b).

From the hearer's perspective, (7b) is in a sense much more helpful than (7a). However, when a hearer who does not know the concept comes across (7a), there are three pieces of information that can be used to infer the meaning. First, the head indicates that we are talking about three instruments. Secondly, the non-head indicates that one of these instruments is a piano. Thirdly, the fact that (7a) is a compound used as a name suggests

that we are dealing with an established concept, i.e. a well-known genre. The degree of specification that can be achieved on this basis depends on the hearer's previous knowledge and experience.

At this point, it is worth returning to the meaning of *chamber orchestra*. When we try to understand the meaning on the basis of *chamber*, we run into problems. The original sense of *chamber* is a room in a house for private use. In the case of chamber music such as piano trios, a performance in such a room is quite normal. However, taking *chamber* in *chamber orchestra* as the place where the orchestra plays is less plausible when we consider that the OED definition refers to "fewer than forty musicians". A chamber orchestra typically performs in a concert hall. When we start from the naming perspective, we arrive at a better motivation. Because of *chamber music* as an established type of music, *chamber* is associated with a smaller group of musicians. In a period in which orchestras had expanded to eighty or more musicians and orchestras competed in size, *chamber orchestra* emerged as a good, i.e. understandable, name for an orchestra that consciously limited the number of musicians. Here onomasiological coercion has an even stronger influence than in *piano trio*.

On the basis of the typical examples of *symphony orchestra* and *piano trio*, we have seen that typical compounds consist of two components, one of which is the head. They are used for naming and their meaning is constrained by the meaning of their components, but determined by onomasiological coercion. These observations will then serve as the basis for a definition.

3. DEFINITION AS A TWO-STEP PROCEDURE

We have now seen two approaches to defining *compound*, each with their strengths and weaknesses. Criteria of the type discussed in section 1 are easy to apply, but cannot be generalized to all languages. Criteria of the type discussed in section 2 give a better perspective of cross-linguistic application, but they are quite abstract. In order to come up with a definition procedure that is at the same time valid for all languages and practical in its application to individual cases, I propose a two-step procedure. In a first step, the properties taken as criteria in the definition of *compound* are used to identify language-specific constructions that qualify as compounding. In a second step, properties of such a language-specific construction are used to determine which expressions are compounds.

The basic idea of such a two-step procedure for the definition of compounding can be traced back to ten Hacken (1994). The advantages of the criteria of the type discussed in section 1, which I will call *low-level criteria* here, are not lost, because they are used as tests for constructions. The calibration of their use and interpretation is motivated by invoking

the *high-level criteria* selected as part of the cross-linguistically valid definition. What exactly counts as a construction is a question that can be approached pragmatically. It is determined by how the high-level criteria can be applied and the low-level criteria selected. The notion of *construction* used in ten Hacken (1994) is not determined by Booij's (2010) homonymous concept, although the two are probably largely compatible.

In a terminological approach to defining, the starting point is a number of typical examples. The next step is the selection of shared properties of these examples that are deemed crucial. On the basis of the discussion in sections 1 and 2, we can identify two sources of properties that can be used as criteria. On one hand, a compound behaves as a word with two components. On the other, compounds have a characteristic way of being assigned a meaning. The definition I propose is (8).

(8) Definition of compound

A compound is a word with a structure $[X Y]_Z$ or $[Y X]_Z$, such that:

- The denotation of Z is a subset of the denotation of Y;
- The range of possible relations between X and Y depends on the semantics of X and Y;
- X does not have independent access to the discourse.

In line with standard terminological practice, e.g. Wüster (1979), (8) specifies a hyperonym and special conditions that distinguish the concept from its co-hyponyms. The hyperonym determines a binary structure. For *symphony orchestra* and *piano trio*, this condition is obviously fulfilled. If we find three basic elements, as in *radio symphony orchestra*, we can only fit it into the definition in (8) if we assume that it has one of the structures in (9).

- (9)** a. [radio [symphony orchestra]]
 b. [[radio symphony] orchestra]

Semantically, (9a) is more plausible. In this case, Y is instantiated as *symphony orchestra*, whose internal structure does not play a direct role in the analysis of (9a). The reason for giving both orders of X and Y in (8) is that the semantic and syntactic status of the two components is not equal, but the role is not universally correlated with their order. In discussing compounds in French, Arnaud (2016) gives the example of *opéra rock* ('rock opera'), which illustrates that in French the order is reversed compared to English.

The first bullet point in (8) highlights the role of Y on the basis of its semantic contribution. Although we may use the term *head* for Y, we should keep in mind that it is not the same notion as Williams's (1981) *head*. Williams defines *head* on the basis of morphosyntactic properties and assigns this status also to affixes. The reference to *denotation* in (8) implies that Y is a content word with a lexical rather than purely grammatical meaning. In English, Y is the right-hand component, i.e. *orchestra* in *symphony orchestra* and *trio* in *piano trio*.

The second bullet point in (8) addresses the semantic relation between the two components. The list of different compounds with *orchestra* as the second component in Table 1 illustrates how the meaning of the first component influences the range of possible interpretations. However, in view of cases such as *piano trio*, it cannot be stated that the relation is determined by the two components. As we saw in section 2, onomasiological coercion determines the meaning of the compound. The reference to a range of possible relations again distinguishes compounds from derivations. *Viewer* can refer to a person or an instrument, and the choice between them is determined by onomasiological coercion, but this is not a choice between different relations of the components. If there is a relation involved in *viewer*, it is one between *view* and *viewer*, not between *view* and *-er*.

The third bullet point in (8) is about the syntactic position of the non-head. Its impact can be seen in (10).

- (10) a. I went to hear a symphony_i orchestra. It_i was very loud.
 b. I went to hear a symphony orchestra_i. It_i was very loud.
 c. I went to hear a piano_i trio. *It_i was out of tune.

By introducing *symphony orchestra*, one does not introduce *symphony* into the discourse. Therefore, *it* in (10a-b) cannot refer to *symphony*, but only to *orchestra*. Similarly, in (10c) *it* cannot refer to *piano*. As illustrated in (11), some proper names do not have this restriction, because they inherently refer to the object they name.

- (11) Brahms_i symphonies are generally considered too difficult for a youth orchestra. He_i requires too much technical skill and musical maturity.

Because *Brahms* in (11) refers to an individual, independently of its context of use, it can serve as the antecedent of *he*, even if embedded in a compound. Therefore, (11) cannot be used to exclude *Brahms symphony* as a compound.

In (8), we have a definition of *compound* that fulfills basic terminological criteria for a definition. The properties it builds on are widely accepted as properties of compounds. Thus,

Allen (1978) argues for the IS A and Variable R conditions, which correspond to the first two bullet points and Lieber (2004: 46) lists properties corresponding to the first and third bullet points. The application of this definition was illustrated for some prototypical English N+N compounds in order to explain the individual criteria. At the same time, the less superficial nature of these criteria makes their cross-linguistic applicability plausible.

4. SOME APPLICATIONS

The N+N construction illustrated by *symphony orchestra* and *piano trio* can be taken as prototypical for compounding. Another construction that is often considered compounding is the Saxon genitive. Marchand (1969: 65-69) discusses cases such as *driver's seat* as compounds. As it is generally accepted that spelling is a derivative property, the question may in principle be asked whether *driver's seat* should not be analysed as a variant spelling of *drivers seat*, with a plural noun as non-head. Often, German linking elements are homonymous to plural endings. Thus in *Blumenladen* ('flower shop'), the first noun is *Blume*. It has a plural form *Blumen*, but *-en* is also a common linking element. In the two-step procedure, we can choose the examples of a construction to be used for determining whether or not it is a compounding construction. Therefore, I will take *women's* as the genitive noun, for which the morphophonological ambiguity of the 's does not arise. In (12), we have an example of *women's magazine*, taken from BNC (2007).

(12) One national women's magazine recently carried an advertorial with no fewer than ten participating, but of course non-competing, manufacturers.

OED (2000-2020 [2011]: *woman*) describes *women's magazine* as "a magazine designed primarily for women", which indicates that it has a head *magazine* and the relation with the non-head was selected in a way so as to name a particular concept. The final clause of (8) can be applied by continuing with a pronoun after (12), e.g. *They were enthusiastic about it*. In this case, *they* can only refer to *manufacturers*, not to *women*. This can be explained because, as the non-head of a compound, *women* in (12) does not actually introduce any women in the discourse.

The conclusion that *women's magazine* is a compound does not imply that all N 's+N are compounds. Another corpus example, taken from NOW (2016-2020) is (13).

(13) It's 2015 and yet we still can't separate women's accomplishments from their looks.

In (13), *women's accomplishments* cannot be a compound, because *their* refers to *women*. An easy way to separate the cases such as (12) and (13) is the position of adjectives. In (12), *national* comes in front of *women*. The alternative, **one women's national magazine* is ungrammatical. In (13), if we insert *academic*, it must come between the two nouns, *women's academic accomplishments*. Therefore, the test of adjective placement can be used to distinguish compounds such as *women's magazine* in (12) from syntactic constructs such as *women's accomplishments* in (13).

This brief discussion of the Saxon genitive constitutes a good example of the two-step procedure for recognizing compounds. By means of the definition in (8), we can determine that there are N's+N compounds in English. The compounding construction has to be distinguished from a similar syntactic construction, which can be achieved by a language-specific test based on the position of an adjective.

The affinity between genitive constructions and compounding is not only found in English. Many German linking elements can be analysed as (historical) genitive endings. Thus, in *Handelshaus* ('trade house'), the linking element *-s-* is homophonous to the genitive ending of *Handel*. In Romance languages, the genitive is expressed by a periphrastic construction with a preposition, illustrated for French in (14).

- (14) a. usine d'automobiles ('factory of cars', i.e. car factory)
 b. goutte d'eau ('drop of water', i.e. water drop)
 c. bouteille de vin ('bottle of wine' or wine bottle)

The French preposition *de*, translated as 'of' in (14), has a meaning and function that corresponds to the morphological genitive in other languages. Ten Hacken (2013) compares English, with its Saxon genitive, Polish, with a morphological genitive as in *fabryka samochodów* ('factory cars_{gen}'), and French. In all cases, we get an ambiguity of the type we saw in the contrast between (12) and (13). In (14c), this ambiguity is realized in a single expression as illustrated in (15).

- (15) a. une bouteille de vin ouverte ('a_f bottle_f of wine_m open_f)
 b. une bouteille de vin blanc ('a_f bottle_f of wine_m white_m)

In (15), the glosses indicate gender of all words that are marked for it. In (15a), the adjective has feminine gender and modifies *bouteille*. This corresponds to an interpretation as 'wine bottle', which aligns it with compounds. In (15b), the adjective has masculine gender and modifies *vin*. This corresponds rather to the interpretation 'bottle of wine', i.e. as a syntactic construction in which *bouteille* indicates a quantity. What we can observe here is,

then, a correlation between syntactic and semantic properties. How this correlation can be used to come up with a set of precise criteria for delimiting the compounding construction in French remains to be determined.

Another type of construction with a recognized affinity to compounding is the combination of relational adjectives with nouns (RA+N). As we saw in the discussion of (4) in section 1, Matthews (1974) excludes them from compounding because the adjective agrees with the noun. This criterion does not apply in English, because adjectives do not show agreement. Levi (1978) includes RA+N into her account, but she avoids the term *compounding*, using *complex nominal* instead, and only discusses English data. As argued by ten Hacken (2019b), there are good reasons for considering RA+N a compounding construction, because the way they receive their meaning is similar to that of prototypical N+N compounds.

The central properties that mark a relational adjective are that it corresponds to a noun and that its meaning is only determined by the relation to the designation of the noun. This is illustrated by pairs such as (16) and (17).

- (16) a. urban police
b. city police

- (17) a. urban development
b. city development

In (16a) and (17a), we have RA+N expressions. The noun corresponding to *urban* is *city*. Although it is possible that the two expressions in (16) and (17) are assigned different meanings, any such difference is due to onomasiological coercion, not to differences in the expression. A more likely scenario is that the pairs in (16) and (17) are competing expressions. The frequency in BNC (2007) suggests that, at least in British English, (16b) and (17a) have won this competition, although the other expressions are also attested.

A question that has to be addressed when RA+N expressions are treated as compounds is how to distinguish relational adjectives from other adjectives. The Italian contrast in (18) illustrates this issue.

- (18) a. banca popolare (lit. ‘bank popular’)
b. calciatore popolare (lit. ‘football_player popular’)

The Italian adjective *popolare* can be a relational adjective corresponding to *popolo* (‘people’), as in (18a), or a qualitative adjective expressing popularity, as in (18b). For (18a)

there is no lexicalized corresponding expression in English, but in German there is the corresponding compound *Volksbank*, which has *Volk* ('people') as its non-head. There are various well-known tests that exploit the difference between these two uses. Two of them are illustrated in (19).

- (19) a. #banca molto popolare
 b. calciatore molto popolare
 c. #la banca è popolare
 d. il calciatore è popolare

In (19a-b), the adjective *popolare* is modified by *molto* ('very'). In the case of (19b), the modifier amplifies the property it conveys in (18b), but in (19a), it requires a reinterpretation of the adjective compared to the normal reading of (18a). The same contrast occurs in (19c-d), where the adjective is used predicatively with the copula *è* ('is'). The # in (19) indicates the change in interpretation of the adjective with respect to (18). Clearly, the precise set of tests will have to be calibrated for each language, but the general idea behind it is that a relational adjective expresses a relation to an underlying noun, which, as *urban* in (16) and (17) illustrates, need not be morphologically related to the adjective. The meaning of an RA+N expression is determined in much the same way as for (other) compounds.

The dictionary definitions in (1) apply not only to typical compounds such as *symphony orchestra* and *piano trio*, but also to a range of other expressions. Bloomfield (1933) gives a typology of compounds that is based on earlier work, in particular in Sanskrit grammar. This typology has had a significant influence on subsequent accounts, mostly in a somewhat simplified format with three types, exocentric, determinative and copulative. Determinative compounds have a head and a non-head, so it is no surprise that all compounds discussed so far are of this type. Indeed, the definition in (8) takes the asymmetry between the head and the non-head as a defining criterion. It is therefore a legitimate question what happens with the expressions that have been labelled as exocentric or copulative compounds in the literature.

In (20), some examples of expressions that have been labelled as exocentric compounds are given. They represent different types of expression that have been assigned to this category.

- (20)a. blackcap
 b. breakwater
 c. spaceship

In (20a), we have an A+N noun. As such it contrasts with *blackbird*. Whereas *blackbird* refers to a kind of bird, *blackcap* does not refer to a kind of cap, but to someone or something marked by a black cap. The type in (20a) is often called *bahuvrīhi*. In both *blackbird* and *blackcap*, onomasiological coercion determines the actual meaning, but the meaning assignment is not parallel. In the case of *blackbird*, the meaning is a species of birds, *Turdus merula*, whose male specimens are marked by a black colour. In the case of *blackcap*, the meaning is another species of birds, *Sylvia atricapilla*, whose male specimens are marked by a black area on the top of their head. The two species are illustrated in Fig. 1 and Fig. 2.



Figure 1: Blackbird, © Sannse, Wikipedia



Figure 2: Blackcap, © Ron Knight, Wikipedia

For *blackbird*, the meaning arises in a way that is compatible with (8). A hearer who does not know the word immediately gets the information that it refers to a bird. In *blackcap*, a different rule is at work. A hearer who does not know this already does not get the information that it refers to a bird. Therefore *blackbird* is a compound, but *blackcap* is not. Ten Hacken (2010) proposes an alternative analysis for *bahuvrīhi* expressions such as *blackcap*.

The example in (20b) is a V+N noun. This type is not very common in current English and Bauer (2008: 46) calls it unproductive. However, in Romance languages it is a highly productive type of word formation. Some Portuguese examples are given in (21).

- (21) a. guarda-chuva (lit. ‘guard_v-rain’, i.e. umbrella)
 b. porta-lápis (lit. ‘carry_v-pencil’, i.e. pencil holder)

This type clearly violates the headedness constraint, because neither of the two components indicates the class to which the referent of the entire word belongs. The English translation in (21b) indicates that there is a parallel with synthetic or verbal compounds of the type N+[V+affix]_N. However, this parallel has its limitations, as shown in (22).

- (22)a. toca-discos (lit. 'play_v-records', i.e. record player)
 b. *joga-estrela (lit. 'play_v-star', intended as 'star player')
 c. jogador estrela (lit. 'player star', i.e. star player)

In English, synthetic compounds with the head *player* can have a non-head corresponding to the object of *play*, e.g. *record player*, or not, e.g. *star player*. In Portuguese, there are two verbs corresponding to *play*, *tocar* for music and *jogar* for sports. As shown by (22a), the V+N construction can be used if the N is the object of the V. However, in other cases, as in (22b), this is not possible. Instead, the left-headed N+N compound in (22c) can be used, although *craque* is the more common translation. What we find, then, is that the English type N+[V+affix]_N is a compounding construction according to (8), but the Portuguese V+N construction imposes a much more restricted relation between the two components.

In (20c), we have an N+N noun of a type that has sometimes also been labelled as *exocentric*. It is not a *bahuvrīhi*, but its meaning can hardly be described as a kind of ship. If (20a) is an example of metonymy, (20c) can be described as using metaphor. This difference is significant, because neither component of *blackcap* expresses what it is. What *blackcap* expresses is only what it has. In *spaceship*, however, the referent is identified as a kind of ship, even if *space* compels one to take some liberty in interpreting *ship*. Therefore, the metaphorical interpretation of *ship* is in line with the use of onomasiological coercion in the application of (8) and (20c) should be categorized as a compound.

The third type of compound in the classification based on Bloomfield (1933) is copulative compounds. Some examples of expressions that have been called copulative compounds are given in (23).

- (23)a. singer-songwriter
 b. Czechoslovakia
 c. girlfriend

All three examples in (23) consist of two components, but the relation between the meaning of the whole expression and its components is different. The opposition between (23a) and (23b) can be expressed conveniently on the basis of the structure [X Y]_Z. In (23a), Z designates a person who is at the same time X and Y, whereas in (23b), Z designates a country that consists of the combination of X and Y. In English, the type in (23b) is mostly restricted to proper names. In Modern Greek it is more common. Some examples are given in (24).

- (24)α. σαββατοκύριακο (lit. ‘Saturday-Sunday’, i.e. weekend)
 b. αντρόγυνο (lit. ‘man-woman’, i.e. [married] couple)

The type illustrated in (24) clearly violates the headedness condition in (8), because a weekend is neither a kind of Saturday nor a kind of Sunday, but only the combination of the two days. This also applies to (23b).

Along the lines of (23a), we can also have longer lists of attributes. Thus, *orator-statesman-lawyer-philosopher* has been used as to qualify Cicero. In this case, the interpretation of the four nouns as combined in a list violates the binary structure imposed by (8). It would be arbitrary to assign it one of the five possible binary structures, because no adjacent pair is more strongly related to each other than to the other components. In fact, the order may be determined by prosodic considerations or some kind of logical sequence, but it is not fixed on the basis of headedness. Therefore, the type illustrated in (23a) is not a compound according to (8).

For (23c), it has sometimes been claimed that it is of the same type as (23a). Thus, Marchand (1969: 41) states that they are “not grammatically different”. However, whereas (23a) is symmetric, (23c) is not. It is true that a girlfriend is at the same time a girl and a friend, but whereas a girlfriend is a kind of friend, we cannot say that a girlfriend is a kind of girl. This shows that (23c) is not symmetric, but headed. The specialized meaning of ‘female half of an unmarried couple’ is the result of onomasiological coercion operating on a meaning in accordance with (8). Therefore it is a compound.

In this section, I have illustrated some consequences of the definition in (8). First, the discussion of genitive constructions and relational adjectives illustrates how the two-step procedure of a cross-linguistically valid definition and language-specific tests for constructions can work in practice. Then, the discussion of various types of expressions that have been labelled *exocentric* and *copulative compounds* illustrates how the definition can be invoked to draw a borderline between compounds and non-compounds in the domain of such cases.

5. CONCLUSION

In this paper, the question of the delimitation of compounding was approached from a terminological perspective. In the relevant sense, *compound* is a term of linguistics. The fact that it is used in other fields, e.g. chemistry, is irrelevant for the linguistic term. At the same time, *compound*, taken in the linguistic sense, is a word in the mental lexicon of linguists and other speakers who know it. In the mental lexicon, the meaning of words is prototype-based

and speaker-specific. In terminology, the purpose is to delimit concepts more precisely and in such a way that different speakers arrive at the same conceptual boundaries. Hence the starting point is a concept with vagueness originating from two sources. On the one hand, for individual speakers (linguists), there are prototypical instances and a gradual transition to non-instances without a clear, natural borderline. On the other hand, different linguists each have their own prototypical concepts in their mental lexicon. The purpose of a terminological definition is to introduce a standard with precise boundaries. Such a standard is necessary for testing theoretical claims. The definition changes the nature of the concept. Instead of a structured concept with more or less typical instances, the terminological concept is determined exclusively by its boundary.

The result of the discussion in this paper can be viewed in two different ways. First, I proposed a particular definition of *compound*, formulated in (8), and explored its effects on a number of controversial and less controversial instances. At a more basic level, however, I demonstrated how a two-step procedure can be used for the definition of concepts such as *compound*. This two-step procedure takes into account that interesting theories in linguistics generally make cross-linguistic claims, but data are necessarily language-specific. Definition then involves a universal and a language-specific step, mediated by the level of what I call *construction*. The basic idea of this two-step procedure can also be applied to linguistic concepts such as *adjective*, *subject*, and *diminutive*.

In formulating the definition in (8), I made the assumption that what determines whether an expression is a compound or not is the way its meaning is assigned. Therefore, phonological and inflection-based properties were not retained in the cross-linguistic definition used to identify language-specific constructions, although they can play an important role in the second, language-specific step, mediating between constructions and individual expressions. The central intuition used for the language-independent definition is that compounds are words with a binary asymmetric structure that combine the two components without imposing a specific relation between them. As with all naming units, the meaning of a particular compound is ultimately determined by onomasiological coercion.

Compared to the definitions in (1), (8) is much more specific. Whereas the definitions in (1) might make *without* a compound, (8) clearly excludes this. In cases such as RA+N constructions, whether the definitions in (1) include them depends on one's interpretation of *word*. In (8) they are included. The main difference is not that one definition uses other terms and the other one does not. The main difference is that (8) gives a number of specific properties that identify constructions cross-linguistically that share the way the expressions receive their meaning. Definitions such as (1) can be seen as deficient in the sense that they do not determine a boundary for the concept. They only give a property that compounds share, but do not indicate why this should be. The notion of *word* used in (8) is

specified by the further conditions it imposes. As a result, (8) diverges in some cases from traditional classifications. However, by taking the way meaning is assigned as a starting point, the constructions identified by (8) share important properties. As such, (8) is a better definition from a terminological perspective.

REFERENCES

- Allen, Margaret Reece (1978), *Morphological Investigations*, PhD Dissertation, University of Connecticut.
- Arnaud, Pierre J.L. (2016), 'Categorizing the modification relations in French relational subordinative [NN]_N compounds', in ten Hacken, Pius (ed.), *The Semantics of Compounding*, Cambridge: Cambridge University Press, pp. 71-93.
- Bauer, Laurie (2008), 'Les composés exocentriques de l'anglais', in Amiot, Dany (ed.), *La Composition dans une perspective typologique*, Arras: Artois Presses Université, pp. 35-47.
- Bloomfield, Leonard (1933), *Language*, London: Allen & Unwin.
- BNC (2007), *British National Corpus*, XML edition, University of Oxford, <http://www.natcorp.ox.ac.uk/>
- Booij, Geert (2010), *Construction Morphology*, Oxford: Oxford University Press.
- COCA (2008-2017), *The Corpus of Contemporary American English*, Mark Davies (ed.), <http://corpus.byu.edu/coca/>
- COED (2011), *Concise Oxford English Dictionary*, 12th edition, Stevenson, Angus & Waite, Maurice (eds.), Oxford: Oxford University Press.
- Collins (2014), *Collins English Dictionary*, 12th edition, Glasgow: HarperCollins.
- Dressler, Wolfgang U. (2006), 'Compound Types', in Libben, Gary & Jarema, Gonia (eds.), *The Representation and Processing of Compound Words*, Oxford: Oxford University Press, pp. 23-44.
- ten Hacken, Pius (1994), *Defining Morphology: A Principled Approach to Determining the Boundaries of Compounding, Derivation, and Inflection*, Hildesheim: Olms.
- ten Hacken, Pius (2010), 'Synthetic and Exocentric Compounds in a Parallel Architecture', *Linguistische Berichte Sonderheft* 17:233-251.
- ten Hacken, Pius (2013), 'Compounds in English, in French, in Polish, and in General', *SKASE Journal of Theoretical Linguistics* 10: 97-113.
- ten Hacken, Pius (2018), 'Terms between Standardization and the Mental Lexicon', *Roczniki Humanistyczne* 66 (1): 59-77.
- ten Hacken, Pius (2019a), *Word Formation in Parallel Architecture*, Berlin: Springer.
- ten Hacken, Pius (2019b), 'Relational Adjectives between Syntax and Morphology', *SKASE Journal of Theoretical Linguistics* 19 (1): 77-92.
- Levi, Judith N. (1978), *The syntax and semantics of complex nominals*, New York: Academic Press.
- Lieber, Rochelle (2004), *Morphology and Lexical Semantics*, Cambridge: Cambridge University Press.

Lieber, Rochelle & Štekauer, Pavol (2009), 'Introduction: Status and Definition of Compounding', in Lieber, Rochelle & Štekauer, Pavol (eds.), *The Oxford Handbook of Compounding*, Oxford: Oxford University Press, pp. 3-18.

Marchand, Hans (1969), *The Categories and Types of Present-Day English Word-formation: A Synchronic-Diachronic Approach*, 2nd edition, München: Beck.

Matthews, Peter H. (1974), *Morphology: An Introduction to the Theory of Word Structure*, Cambridge: Cambridge University Press.

NOW (2016-2020), *Corpus of News on the Web (NOW): 10 billion words from 20 countries, updated every day*, ed. Davies, Mark, <https://www.english-corpora.org/now/>

OED (2000-2020), *Oxford English Dictionary*, Third edition, edited by John Simpson and Michael Proffitt, Oxford: Oxford University Press, www.oed.com

Smallman, Basil (1992), *The Piano Trio: Its History, Technique, and Repertoire*, Oxford: Oxford University Press.

Williams, Edwin (1981), 'On the notions *Lexically related* and *Head of a Word*', *Linguistic Inquiry* 12:245-274.

Wüster, Eugen (1979), *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*, L. Bauer (ed.), Wien: Springer.