

Compounds in English, in French, in Polish, and in General

Pius ten Hacken

The notion of compound can be taken as a theoretical concept only if it has a precise definition. In many current discussions it is assumed that such a definition is not available or not possible. Here, I will show how translation can be used as a heuristic to determine a concept of compound that is semantically coherent. This concept includes genitive constructions and constructions of a relational adjective with a noun, but not prepositional constructions in which the preposition expresses part of the meaning. An essential component of the use of compounds for naming is shown to be onomasiological coercion.

Keywords: *compound, definition, genitive, relational adjective, preposition, onomasiological coercion*

1. Introduction

It is often claimed that the boundaries of concepts such as *compound* should be seen in terms of a continuum between clear compounds and clear members of another class, in this case derivation or phrase. Here I will argue that this approach is neither necessary nor desirable. I will start by discussing the nature of theoretical concepts, taking *compound* as an example (section 2). Then, I will present some preliminary considerations that have to be taken into account when discussing compounding as a word formation device (section 3). Section 4 sets out a methodology by which we can arrive at a set of defining properties for compounding. Sections 5, 6 and 7 present some results that can be achieved in using this methodology. Finally, section 8 summarizes the conclusions we can draw from the reasoning in earlier sections.¹

2. Compound as a theoretical concept

The dictionary definition in (1) provides a convenient starting point for the discussion of the definition of *compound*.

- (1) [Compound:] a word made up of two or more words

This particular definition is from Chambers (1998), but similar definitions can be found in other dictionaries. The difficulty of applying (1) to a particular expression to determine whether it is a compound arises from the two occurrences of ‘word’ in (1). Both of them serve to characterize *compound* as ‘marked’, less common. The first occurrence refers to the result. It contrasts with *phrase*, which is a less marked outcome of the combination of two or more words. The second occurrence refers to the components. It contrasts with *affix* in the sense that complex words are more typically put together from a word and an affix. It is these two boundaries that Lieber & Štekauer (2009: 4) identify as problematic in (2):

- (2) [W]e cannot always make a clean distinction between compound words on the one hand and derived words or phrases on the other.

The fact that (2) is found in the introduction to a *Handbook of Compounding* shows that uncertainty about the exact boundaries of the concept is part of many people's perception of the current state of research. The formulation in (2) leaves open whether this lack of a 'clean distinction' is due to disagreement between researchers or accepted fuzziness of the concept.

An example of the former situation is the term *government*. In the framework of Chomsky's (1981) Government and Binding framework, a number of different definitions of *government* have been proposed. Linguists working with different definitions did not agree on whether in a particular structure a node α would govern a node β . However, no linguist would claim that it is not clear whether α governs β . They would argue for or against a particular definition, but each definition would decide quite clearly whether α governs β or not. Newmeyer (1986: 204-205) identifies the question of how government is best defined as one of the most important theoretical issues in GB-theory at the time.

An indication that in the case of *compound* the issue of definition is different from the case of *government* is Dressler's (2006: 24) statement in (3):

- (3) [U]niversal definitions [of compounding] are not only theory-dependent [...] but also cross-linguistically never watertight—in many languages there are exceptions or fuzzy transitions to non-compounding[.]

What (3) suggests is that Dressler sees the entire idea of a universal definition of compounding as unrealistic. His objections are that the definitions that have been proposed are first that they are theory-dependent and second that they have counterexamples ('exceptions'). He seems to suggest that linguists should accept that the transition between compounding and non-compounding in many (or at least some) languages is fuzzy. This is reminiscent of the position Bloomfield (1933: 223-224) takes with respect to the boundary of inflection and derivation when he states that "[t]his distinction cannot always be carried out."

In psycholinguistic research, it has long been accepted that general-language concepts such as *cup* have fuzzy boundaries. In a famous experiment, Labov (1973) asked informants to classify objects as cups, bowls, and vases. He discovered that the boundaries between these concepts was fuzzy in the sense that there was no specific point at which all informants placed the boundary between *cup* and *bowl*. Instead, the judgements would vary between informants and, for the same informant, between instances in different contexts. On the basis of this type of observations, Rosch (1978) proposed her prototype theory. For concepts like *cup* and *bowl*, each speaker has a prototype, which is encoded as information about the concept in the mental lexicon. Judgements on categorization are based on the perceived distance to prototypes. The perception of the distance can be influenced by various factors, such as which similar items the speaker has recently been asked to categorize.

My claim here is that whereas natural concepts are prototype-based, theoretical concepts are not. If we accept (3) and do not see (2) as a challenge to change the situation, we abandon *compound* as a theoretical concept. This means that we cannot make any testable claims on compounds, but only use the term in a pre-theoretical sense. Only on the basis of a proper definition of *compound* is it possible to evaluate, for instance, whether a particular theory of compounds can account for all instantiations. It should be noted that if we choose to abandon *compound* as a theoretical concept, this does not solve the more general definition problem. Without *compound*, we will need other theoretical concepts in order to make testable claims about them. These concepts also require proper definitions.

In the last paragraph, I changed the focus from *definition* to *proper definition*. The reason for this specification is that some definitions are not suitable for use as a device to determine the boundaries of a concept. A definition such as (1), while perfectly adequate to evoke a prototype, is too vague to determine the boundaries of *compounding*. In Dressler's words, it is not 'watertight'. A proper definition is one that has no exceptions and takes a decision in borderline cases. Whether a definition has exceptions is in part a matter of how we use it. If the definition of a theoretical concept clashes with intuitions we have about that concept, we must accept the categorization that the definition gives. Here a parallel with zoological taxonomy is enlightening. In many people's intuition, *cat*, *dog*, *bird*, and *fish* are approximately equally specific in designating animal species. Zoologically, however, *cat* and *dog* are quite closely related species (both are carnivores), whereas *bird* is a class (i.e. at the same taxonomic level as *mammal*) and *fish* constitutes a paraphyletic group, i.e. a group that does not correspond to any node in the taxonomy. We should not see the intuitions as counterexamples to zoological taxonomy, but rather accept that theoretical concepts of zoology are not the same as general-language concepts with the same names.

The fact that a theoretical definition cannot have counterexamples does not mean that we cannot evaluate it. In fact, my use of *proper definition* already implies an evaluation. A definition is not proper if it does not impose a decision in borderline cases. We can also discuss how good a definition is. However, this type of evaluation is a matter of judgement rather than one of testing. A good definition of a theoretical concept identifies a useful concept. In this sense, competing definitions of *compounding* are no different from competing definitions of *government* in Chomskyan syntax or of *species* and other taxonomic levels in zoology.

Dressler's claim in (3) raises two specific issues in relation to definitions of compounding. One is that definitions are theory-dependent. Although (3) suggests that this is somehow problematic, it is in fact a general property of definitions, because a theory-free definition would only rely on unaided primary intuitions. For a theoretical concept it is not problematic and indeed inevitable that the definition refers to other theoretical concepts. As mentioned in ten Hacken (2008), in a field such as mathematical linguistics the terms constitute a network held together by the definitions. This network encodes essential parts of the knowledge of the field.

The other point Dressler raises in (3) is the cross-linguistic validity of definitions. It may seem attractive to define compounding on a language-specific basis, because compounds in the same language have more properties in common. They constitute a more homogeneous class, so that it is easier to find criteria that distinguish them. However, with any set of logically independent criteria, there may be borderline cases where some of them are met but others not. The choice of the most important criteria remains arbitrary when considered in the space of one language, but taking into account other languages may reduce the arbitrariness of the selection of criteria. Using cross-linguistic data comparison to decide on language-specific issues is a technique that is well-known from Chomskyan linguistics, e.g. Chomsky (1986: 37-38).

Therefore, it is worth looking for a proper definition of compounding. Such a definition takes priority over any individual judgements or opinions, determines the boundaries of the concept, uses other theoretical concepts, and is cross-linguistically valid.

3. Compounding as a word formation device

In searching to define compounding, it is essential to keep in mind that it is a word formation process. This statement may seem trivial, but it is sometimes not sufficiently taken into account when a set of rules for compounding is proposed. Superficially, we could take a compound such as *doghouse* to be generated by a rule such as (4):

(4) $N \rightarrow N N$

Selkirk (1982: 16) proposes an account of compounding that takes a set of rules such as (4) as its basis. She then notes that compounds in English are right-headed (1982: 20) and that the non-head specifies the head in some way (1982: 22). However, this gives a very limited sense of the meaning of a compound such as *doghouse*. A prototypical doghouse may be an object such as in Fig. 1.



Figure 1 *Doghouse (or kennel in the UK)*

The object in Fig. 1 is not a typical house. It is a small structure for a dog. A rule such as (4) and any supplementary indications on the possible meanings of compounds will not be sufficient to account for the fact that *doghouse* refers to an object such as Fig. 1. The point is even stronger for compounds such as *spaceship*. A spaceship is not at all a ship.

The explanation why a doghouse is an object such as in Fig. 1 depends on the fact that compounding is a word formation device. As such it is used to name new concepts. The reason why *doghouse* refers to objects such as Fig. 1 and not to ‘any house that is in some way related to dogs’ is that speakers of English perceived the need to name the concept of Fig. 1. This concept then contributed essential attributes to the meaning of the compound. We can call this phenomenon *onomasiological coercion*.

Lexicalization has often been treated as a circumstantial property of compounds. Thus, Levi (1978: 10-12) excludes lexicalized compounds from the scope of her theory. However, it is actually a central property of compounding that it is used to produce lexicalized expressions. In this sense, the compounding rule (4) is to be distinguished from the syntactic rule which looks the same. In a language like Dutch, we find the contrast in (5):

- (5)
- | | | |
|----|----------------------|---|
| a. | <i>een wijnfles</i> | ‘a wine bottle’ |
| b. | <i>flessenwijn</i> | ‘bottles wine’, i.e. wine sold in bottles (as opposed to casks) |
| c. | <i>een fles wijn</i> | ‘a bottle wine’, i.e. a bottle of wine |

All three expressions in (5) can be generated by means of (4), but only (5a) and (5b) are compounds. (5c) is an example of a syntactic construction indicating quantity. The compounds refer to specific concepts. (5a) is a bottle of a certain shape. (5b) is wine of a certain (minimal) quality. These properties are added by onomasiological coercion.²

4. Cross-linguistic evidence for a definition

In order to determine which criteria can be used to define *compounding*, we have to consider data from different languages. On one hand, it seems straightforward to use translation as a component of the methodology. On the other hand, we have to be extremely careful in interpreting the data collected in this way. It is well known that a certain type of equivalence plays a central role in translation, cf. Munday (2008: 36-54), but this is not necessarily equivalence at the lexical level. Baker (2011) organizes her entire textbook on the basis of different levels of equivalence, devoting most of her attention to equivalence at sentential and textual levels. Functionalist approaches, as described by Nord (1997), play down the role of equivalence. Thus, in his *skopos* theory, Vermeer (2000) rates the instructions to the translator and the coherence of the target text resulting from translation higher than the correspondence of the target text to the source text.

An example of the problems that can arise in using translation is the triple of English, French, and Polish in (6):

- (6) a. *doghouse*
b. *niche*
c. *buda dla psa* ‘shed for [a/the] dog’

In all trilingual examples, the (a) example is English and the (b) and (c) examples are the French and Polish translations, respectively. Whereas (6a) is a prototypical example of a compound, (6b) in French is a simplex word which can also mean ‘recess, alcove’. Polish (6c) involves the preposition *dla* (‘for’). In interpreting the Polish translation, one should keep in mind that Polish does not have articles, so that (6c) can be translated as ‘shed for a dog’ or ‘shed for the dog’. Obviously, (6b) is not a compound, because it is not possible to identify two components. For (6c), it is at least not obvious that it is a compound, because there seem to be three components. We will come back to such prepositional constructions in section 6.

As will be shown in subsequent sections, however, provided that it is applied with the appropriate care, translation can be a very useful technique in the identification of properties of compounds in different languages. The starting point for the data-oriented component of this research was the list of compounds Levi (1978: 280-284) gives as an appendix to her monograph on what she calls ‘complex nominals’. In Levi’s theory, complex nominals include both noun-noun compounds and combinations of relational adjective and noun. In the first instance, only the 257 noun-noun compounds were considered. They were translated into French and Polish by native speakers of these languages with a background in translation.

The collection of compounds by Levi (1978) has the advantage of giving a systematic classification of meanings. However, some of them were quite specific to the geographic, temporal and intellectual environment Levi was working in and were not understood by the translators. They were discarded from the analysis. The distribution of the translations among different constructions is represented in Fig. 2.

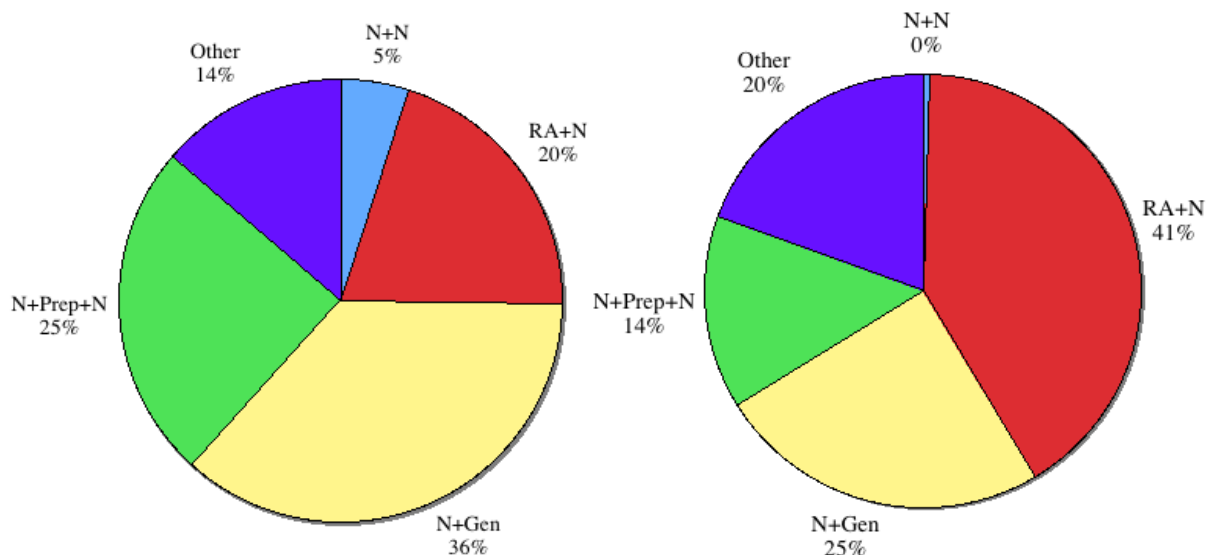


Figure 2 Distribution of translations in French (left) and Polish (right)

The most striking feature of the pie charts in Fig. 2 is the low proportion of noun-noun compounds. In French, such compounds are not readily accepted, although some have become quite common, e.g. *centre-ville* ‘centre-town’, i.e. town centre. An example from our data set is *prix plafond* ‘price ceiling’, i.e. ceiling price. In Polish, there was only a single case, *pletwonurek* ‘flipper-diver’, i.e. frog man, in our data set. As Szymanek (2010: 219) states, such Polish compounds have a linking element, typically *-o-*, and are generally quite rare. The class marked ‘other’ in Fig. 2 includes translations such as *niche* in (6b), where no two words can be distinguished, or paraphrases, as in (7) in Polish:

- (7) *dziura w ubraniu wygryziona przez mole*
 ‘hole in [a] piece_of_clothing chewed by moths’, i.e. moth hole

There are three large classes of translations for English noun-noun compounds in French and Polish. One is a combination of a noun with a relational adjective, a second combines a noun with a genitive, and a third connects two nouns by means of a preposition. Together, they make up around 80% of the translations in French and Polish. They will be considered in more detail in the following sections.

5. Genitive constructions

An example of genitive constructions in French and Polish corresponding to an English noun-noun compound is given in (8):

- (8) a. *car factory*
 b. *usine d’automobiles* ‘factory of cars’
 c. *fabryka samochodów* ‘factory cars_{GEN}’

In all three expressions in (8), we can observe a distribution of the meaning among two components. The genitive markers in French and Polish can be seen as indicators of the

construction. Polish has morphological case marking on nouns, so that (8c) has the genitive plural of *samochód* ‘car’. In French, morphological case was lost around the 13th century and the genitive is expressed by means of the preposition *de* ‘of’. English also has a genitive marker and it is used in certain types of compounds, as in (9):

- (9) a. *children’s bible*
 b. *kinderbijbel* ‘child bible’

As the Dutch translation (9b) shows, other languages use noun-noun compounds as equivalents. The form *kinder* in (9b) is not the same as the singular *kind* (‘child’) or the plural *kinderen* (‘children’), but has an extension *-er*. In ten Hacken (1994: 258-262) I propose to analyse *kinder* as a stem variant of *kind*, along the lines of Aronoff’s (1994) analysis of the Latin verbal paradigm. Booij (2002: 22-23) adopts a similar analysis. In a sense, the ‘s in English and *de* in French can be seen as markers in the same way as Dutch *-er* in (9b).

Whereas (9a) can be a compound, there are also contexts in which it is clearly not one. The ambiguity is illustrated in (10):

- (10) a. *The children’s bible is full of pictures.*
 b. *This book is a children’s bible. It is full of pictures.*
 c. *This bible belongs to the children. It is full of pictures.*

Both (10b) and (10c) can be paraphrases of (10a), but they are not synonyms. In (10b), *children’s bible* is a single concept, a kind of bible. In (10c), *bible* and *children* refer to separate concepts. The intuition that there is a single concept in one but not in the other is supported by the possibility of pronominal reference, as illustrated in (11):

- (11) a. *Anna bought them_i a children’s bible, but they_i did not like the pictures.*
 b. **Anna bought a children_i’s bible, but they_i did not like the pictures.*

The pronoun *they* in (11) cannot refer back to *children* as part of *children’s bible*. In (11), no interpretation of *children’s bible* along the lines of (10c) is possible, because the article *a* does not agree with *children*. The different interpretations in (10) also correspond to different structures assigned to the subject NP in (10a), as illustrated in (12):

- (12) a. [the [children’s bible]]
 b. [[the children’s] bible]

The structure in (12a) corresponds to the paraphrase in (10b) and the one in (12b) to (10c). This explains that ‘a children’s bible’ in (11) can only have a structure on the pattern of (12a), because *a* agrees with *bible*, but not with *children*. The structure is also visible in the placement of adjectives modifying *bible*, which precede *children’s* in (12a), but follow it in (12b).

In French, it is crucial to make a distinction between *de* with and without the definite article. An example of this contrast is (13):

- (13) a. *donneur de sang* ‘donor of blood’
 b. *donneur du sang* ‘donor of _the blood’

Both expressions in (13) can be translated as *blood donor*, but (13b) assumes that *sang* ('blood') has been introduced as an entity in the context in which it is used, whereas (13a) only introduces a single concept. This contrast is of the same type as illustrated for (10) in English. The syntactic contrast in (13) suggests that we can use definite articles as a diagnostic. However, in some contexts the definite article can also be used to indicate genericity, as in *pays des ours* 'land of the bears', i.e. bear country. For Polish, Szymanek (2010: 218) considers genitive constructions as syntactic phrases. The absence of articles in Polish makes it more difficult to distinguish the two readings. Therefore, the interaction of genitive constructions with articles is more complex and cannot be used in all cases as a criterion to distinguish compounds from phrasal expressions.

In view of these data, we can take two positions. On one hand, we can consider genitive constructions as ambiguous between a phrasal and a compounding interpretation. This requires a more precise characterization of the boundary between these two interpretations, but, as illustrated in (8) and (9), it increases the cross-linguistic correspondence between compounds as translational equivalents. Compounds in this interpretation are primarily naming units and the categorization takes semantic criteria into account. The semantics of the criteria is above all the reference to a single as opposed to two distinct entities by means of the genitive construction. On the other hand, we can consider genitive constructions as phrases. This is the position adopted by Szymanek (2010). In this perception, *compound* is a purely formal category. The consequences include an increased number of phrasal naming units and cross-linguistic mismatches in the sense that compounds in one languages correspond to phrases in another.

6. Relational adjectives

Relational adjectives (RAs) are adjectives that do not express a property but rather a relation to a corresponding noun. As shown in Fig. 2, combinations of a noun with a relational adjective are often used in the translation of English compounds, in particular in Polish. An example of an English compound with N+RA equivalents in French and Polish is given in (14):

- (14) a. *cell division*
 b. *division cellulaire* 'division cell_{ADJ}'
 c. *podział komórkowy* 'division cell_{ADJ}'

The suffix *-owy*, as illustrated in (14c), is the most frequent source of relational adjectives in Polish. In this case, *komórkowy* is derived from *komórka* 'cell'. It is hardly possible to assign a specific meaning to such adjectives. Dictionaries often use formulas such as 'or or relating to [N]' (with N the base noun) to characterize their meaning. Kallas (1999: 485-494) distinguishes 18 different categories of meaning for denominal adjectives and lists *-owy* as a suffix for 17 of them. The actual category of meaning depends not only on the adjectival base, but also on the combination with the noun. In the same way as for N+N compounds, onomasiological coercion can result in a not fully compositional meaning. Thus, *krem orzechowy* 'peanut butter' combines *krem* 'cream' and *orzech* 'nut, walnut' to name a substance that is not really a cream and is based on a very specific type of nut.

In French, *cellulaire* in (14b) is related to *cellule* 'cell'. The status of the suffix *-aire* is less obvious than the corresponding suffix *-owy* in Polish. It corresponds to Latin *-arius*

and there are also adjectives such as *pécuniaire* ‘money_{ADJ}’ that do not correspond morphologically to a French noun, cf. *argent* ‘money’. Historically, *pécuniaire* was borrowed separately from Latin, where it does have a regular relationship to the noun *pecunia* ‘wealth, money’. The pair *argent* – *pécuniaire* is analogous in its use to *cellule* – *cellulaire*, although from the point of view of its formation it is suppletive.

For English, Levi (1978) makes the case for treating RA+N combinations in the same way as N+N compounds. The existence of synonym pairs such as *atom bomb* and *atomic bomb* gives such an analysis an initial plausibility. Many of Levi’s arguments are framed in the terminology of Generative Semantics, so that they cannot be used easily outside that framework, but a striking syntactic property is coordination as in (15):

- (15) a. *literary and musical criticism*
 b. **literary and bitter criticism*
 c. *solar and gas heating*
 d. **solar and reliable heating*

Based on Levi’s (1978: 22-23) examples, (15) gives two contrasting pairs. In (15a-b) we see that the RA *literary* can be coordinated with another RA, e.g. *musical* in (15a), but not with qualitative adjectives such as *bitter* in (15b). This means that RAs and regular, qualitative adjectives belong to different categories for coordination. In (15c-d) it is shown that the RA *solar* can be coordinated with the noun *gas* as the non-head of a compound, but not with the qualitative adjective *reliable*. This can be taken as supplementary evidence that RA+N combinations are in the same category as N+N compounds.

A point that is often mentioned as a counterargument to a compounding analysis of RA+N combinations is that the RA agrees with the N. This can be seen most clearly in Polish. An example is the contrast in (16):

- (16) a. *Bożena lubi krem orzechowy.*
 ‘Bożena likes cream_{ACC} nut_{ADJ-ACC}’, i.e. Bożena likes peanut butter.
 b. *Bożena nie lubi kremu orzechowego.*
 ‘Bożena not likes cream_{GEN} nut_{ADJ-GEN}’, i.e. Bożena does not like peanut butter.

In Polish negative sentences, the object appears in the genitive instead of the accusative. In (16b), it is not only the noun *krem* but also the RA *orzechowy* which appears in the genitive. Therefore, the N+RA combination *krem orzechowy* ‘peanut butter’ is inflected in two places. For Matthews (1974: 35), this is a reason to discard N+RA combinations from the category of compounds. He gives Latin *tribunus militaris* ‘tribune military’, i.e. a particular rank in the ancient Roman army as an example.

As with the genitive constructions in section 4, we can take two positions as to the RA+N combinations discussed here. One is to consider them as compounds, the other to analyse them as phrases. If we take them to be compounds, the notion of *compound* becomes more semantically coherent, the link between compounds and naming units is emphasized, and there is a much larger degree of cross-linguistic correspondence. If we take them to be phrases, we make compound into a more formally oriented category, thus creating the need for more phrasal naming units and a considerable set of cross-linguistic mismatches. Matthews (1974: 35) proposes that Latin N+RA combinations are idioms, treated as units in lexicography, but not in morphology. In line with such an analysis, Szymanek (2010: 219) concludes that “the number of nominal compounds is not so spectacular in Polish.” In our

data set, there is only one such compound as a translation for the 257 N+N compounds in Levi's set.

7. Prepositions

Prepositional constructions occur quite frequently among the translations of English N+N compounds in our data set, as shown in Fig. 2 above. An example is the Polish translation in the triple in (17):

- (17) a. *love song*
b. *chanson d'amour* 'song of love'
c. *piosenka o miłości* 'song about love'

Both the French (17b) and the Polish (17c) involve a preposition, but these prepositions have a very different status. In French, *de* is the preposition marking the genitive. In section 4 it was argued that the function of *de* can be seen as parallel to morphological case in Polish and 's in English. Grevisse (1980: 1146-1169) gives a wide range of uses, including one that has a striking parallel to the Polish genitive in (16b), as illustrated in (18):

- (18) a. *Charles a une voiture.*
'Charles has a car.'
b. *Charles n'a pas de voiture.*
'Charles NEG has not DE car', i.e. Charles does not have a car.

The indefinite article *une* of the positive sentence in (18a) is replaced by the genitive preposition *de* in the negative counterpart (18b).

In (17c), the preposition *o* is followed by the locative case. As such, it is analogous to the use in (19):

- (19) *Dariusz i Ewelina rozmawiają o koncercie.*
'Dariusz and Ewelina talk about concert_{LOC}'
i.e. Dariusz and Ewelina talk about the concert.

Both in (17c) and in (19), *o* can be translated as 'about'. This means that *o* in (17c) has a different role than *de* in (17b). The meaning of (17b) is derived from three factors: the meaning of *chanson* 'song', the meaning of *amour* 'love', and the concept to be named. The meaning of *de* does not play a significant role. This makes it parallel to the English compound in (17a). In (17c), however, the meaning of the whole expression is determined by the meaning of *piosenka* 'song', the meaning of *o* 'about', and the meaning of *miłość* 'love'. Whereas onomasiological coercion is instrumental in determining the relation between the two components of the compounds in English (17a) and French (17b), this relation is expressed by the preposition in Polish (17c). This means that the way (17c) gets its meaning is the regular one for syntactic expressions.

When we consider the equivalence between the three expressions in (17) in more detail, we notice that (17c) can be used as a translation of (17a) and (17b), but does not have exactly the same meaning. The use of *love song* or *chanson d'amour* implies that we consider it a genre, a type of song. This implication does not exist for (17c). Moreover, the

specification of the relation in (17c) is more explicit than in (17a) and (17b). Polish speakers I asked came up with various alternative expressions that highlight different aspects of the relation between the two components in (17a) and (17b). It seems, then, that the concept is much less established in Polish than in English and French.

In the data I collected, there were some striking examples of English compounds that did not match compounds in French and Polish. One of them is (20):

- (20) a. *abortion vote*
 b. *vote sur l'avortement* 'vote on abortion'
 c. *głosowanie za aborcją* 'vote for abortion'

Both French (20b) and Polish (20c) can be correct translations of English (20a), but they are not correct translations of each other. In fact, (20a) is ambiguous in English. In (21), two contexts are given.

- (21) a. *The abortion vote is on the agenda for Monday.*
 b. *Senator X was under huge pressure to explain his abortion vote.*

In both sentences of (21), the context is that of a Parliament. In (21a), what is on the agenda must be a vote on a law or motion pertaining to abortion. The outcome of the vote is generally not known in advance, so the relation can only be one of 'on', as in French (20b). In (21b), however, it is implied that the vote was one approving abortion. The content of the sentence indicates that Senator X must have voted and if his vote were against abortion, it would rather be an *anti-abortion vote*. Therefore, in (21b), Polish (20c) would be an appropriate translation. This is not to say, of course, that French cannot express the meaning in (21b) and Polish the one in (21a). It only shows that my informants translating (20a) assumed different readings. They could not translate (20a) without choosing one and apparently different readings came to mind more readily. Therefore, the compound (20a) does not have a translation as a compound in French and Polish, although in any specific context where it appears there is a translation that is correct in that context.

While *o* in (17c), *sur* in (20b), and *za* in (20c) are clear cases of meaningful prepositions expressing the relationship between the two nouns, there is an interesting borderline case in French with the preposition *à*. Nicoladis (2002) considers both N+*de*+N and N+*à*+N as constructions that can be analysed as compounds and notes that both *de* and *à* can be used to express possession in French (2002: 49). The discussion of *à* by Grevisse (1980: 1118-1145) is even longer than the one of *de*, but it starts by giving a much more specific range of meanings for *à* (1980: 1118), translated in (22).

- (22) The preposition *à* is used above all to mark the place, goal, time, means, manner, property.³

Although Grevisse (1980: 1118-1119) also mentions the possessive reading, it is much more constrained in its distribution than the possessive meaning of *de*. In the data set collected for this study, the French informant used *à* in 19 cases. In 7 cases the definite article was included and in 12 it appeared alone. Interestingly, there was a highly significant semantic bias among these expressions. The examples in (23) are typical.

- (23) a. *ver à soie*
 worm PREP silk
 ‘silkworm’
- b. *fer à vapeur*
 iron PREP steam
 ‘steam iron’
- c. *gâteau aux pommes*
 cake PREP_{DEF-PL} apples
 ‘apple cake’

In the type illustrated in (23a), *à* marks the right-hand element as the product, in (23b) as the means, and in (23c) as the content. All of the translations with *à* can be placed in these categories and the article occurs for all and only the expressions in the type illustrated in (23c). This suggests that the preposition *à* is meaningful in a way that *de* is not. However, it would be interesting to study this in more detail.

8. Conclusions

The questions addressed in this article are whether *compound* can be given a precise definition and if so, how such a definition can be arrived at. As was argued in section 1, the answer to the first question depends only on whether we want a definition or not. Since *compound* is a theoretical notion, we can choose whatever definition we like and impose it on the data. The problem is that if we take a definition arbitrarily, the resulting notion may not be of much use. Therefore, the method for arriving at a definition of a useful concept becomes crucial.

It is possible to approach *compound* from different perspectives. If we want to consider it a purely formal notion, we can take as a starting point the syntactic category of the components and the stress pattern. However, if we do so, it will be quite accidental if there is any semantic coherence or cross-linguistic correspondence. I would argue that *compound* is interesting above all as a category with semantic and syntactic characteristics. This view matches Jackendoff’s idea that compounding is “a possible protolinguistic ‘fossil’ in English” (2002: 249).

The method used here takes as a starting point that translation can be a good heuristic device to identify relevant cross-linguistic correspondences. Assuming that a definition of *compound* has at least some semantic aspects, we should expect that translations will tend to use corresponding realizations. However, we have to be careful to distinguish the influence of the construction, the influence of the input, and what I called *onomasiological coercion*. The distribution of these factors in determining the meaning of the resulting expression should be similar for translations if they are all considered compounds.

The outcome of applying this method is for a particular construction in a particular language to suggest whether it belongs to compounding or not. Taking as our basis the list of compounds given by Levi (1978), we found that there were three strong candidates for such constructions in French and Polish. Therefore, we considered genitive constructions, relational adjectives, and prepositional constructions in more detail. Whereas the number of N+N compounds in French and Polish among translations of English N+N compounds is

insignificant, together with the three constructions we studied here, they constitute 80% of the translations.

If we assume a concept of *compound* on this basis, the analysis of the three candidate constructions leads to the conclusion that genitive constructions and RA+N combinations are compounds, but prepositional constructions are not. In the case of genitive constructions it was noted that there is a compound variety and a syntactic variety. This contrast is reflected in a different structure in English, cf. (12). In French, the genitive is expressed by *de*. As was noted, *de* may be accompanied by the definite article. In some cases this expresses genuine definiteness, but in others genericity. In Polish, the problem is that articles are not expressed so that the ambiguity involves the equivalent of the entire spectrum of *de* with or without definite article in French.

In both French and Polish, we might use the possibility of inserting a demonstrative as a diagnostic. There are two possible effects, one of which is illustrated in (24):

- (24) a. *?*usine de ces automobiles* 'factory of these cars'
b. *?*fabryka tych samochodów* 'factory these_{GEN} cars_{GEN}'

The examples in (24) are semantically odd in the same way as the English gloss of (24a) is. The point seems to be that the genitive cannot be interpreted as sufficiently close to possession. Another possible effect is illustrated in (25):

- (25) a. *pays des ours*
land of_the bears
bear country
b. *pays de ces ours*
land of these bears

The contrast in (25) shows that with the demonstrative the interpretation of *ours* switches from generic in (25a) to specific in (25b). Demonstratives enforce an interpretation of a genitive construction as involving two concepts. In a context in which a particular species of bear has been introduced, (25b) can be used to refer to its habitat. However, (25a) cannot be used in such a context. Here *ours* is necessarily generic, so that the expression introduces a single concept rather than a connection of two separate concepts.

In the case of RA+N constructions, we also find certain borderline cases, but here the question is how to distinguish relational adjectives from other adjectives. The problem in this distinction is that many adjectives can have both roles in different contexts. Levi (1978) gives the well-known example in (26):

- (26) a. *a nervous applicant*
b. *a nervous disease*

Whereas *nervous* is a qualitative adjective in (26a) it is a relational adjective in (26b). There are various diagnostics to distinguish the two types. A well-known test is to make the adjective predicative, as in (27).

- (27) a. *the applicant is nervous*
b. ?*the disease is nervous*

As suggested by the not quite so ungrammatical (27b), this type of syntactic test has its limitations. Even nouns that are often left-hand components of compounds are sometimes reinterpreted as qualitative adjectives, as in (28).

- (28) a. *It was a fun party.*
 b. *The party was fun.*

Nowadays, (28b) is entirely grammatical, with *fun* as an adjective. The same has happened with *key*. Conversely, gradability with *very* does not exclude only relational adjectives, but also, e.g., **very superb*. Ultimately, only semantic criteria can be used to calibrate the tests.

For prepositional constructions other than with French *de*, it was concluded that the preposition assumes a semantic role that is incompatible with the analysis of these constructions as compounding. Whereas in compounding onomasiological coercion by the concept to be named plays a crucial role, in N+P+N constructions in French and Polish the preposition generally constrains the range of possible relations between the two nouns to such an extent that onomasiological coercion does not have to be appealed to. However, it would be worth exploring certain prepositions in more detail, such as French *à*.

In relation to Fig. 2, we can therefore conclude that approximately 61% of French translations and 66% of Polish translations are compounds. However, in Fig. 2 we only considered English N+N compounds as a basis. Levi (1978) does not have any examples of N's N compounds, but a third of her examples are RA+N combinations. They were excluded from the original set, because it remained to be established whether they should be classified as compounds. When we include them, we get the distribution in Fig. 3.

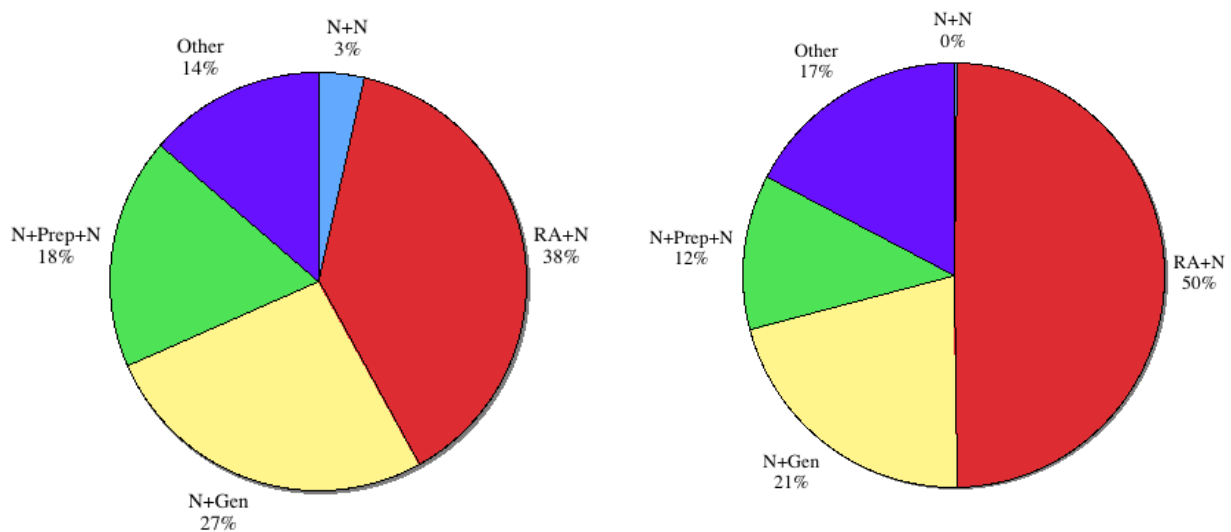


Figure 3 *Distribution of translations in French (left) and Polish (right)*

As can be inferred from the comparison of Fig. 3 with Fig. 2, the translation of RA+N from English into French and Polish results more frequently in the use of relational adjectives. As a result, 68% of French translations and 71% of Polish translations can be classified as compounds. The remaining translations, i.e. approximately 30% in each language, do not use compounding to name the same concepts. It should not surprise us that this is a significant part. After all, concepts do not come with a tag that they want to be named by a compound.

As we saw in some of the examples of N+Prep+N translations, the translational correspondence is often not complete. These expressions are often less strongly lexicalized, so that several alternative translations with a similar status exist. This effect is even stronger for paraphrases of the type we saw in (7). In such cases, translational equivalence is not achieved at word level. As mentioned in section 3, the phenomenon of missing translational equivalents at word level has been studied a lot in translation theory. Arguably, the reason for their prominence is that they are not the normal case. Very often we do find good translational equivalents at word level between different languages.

In conclusion, we have seen that it is useful and possible to define *compounding* as a term. It was argued that the definition should be based on the nature of the naming process rather than the morphophonological properties of the result. If we accept that compounding takes two lexemes as input and uses onomasiological coercion to determine the relation between them, we get a cross-linguistically useful class with a high degree of translational equivalence.

References

- ARONOFF, Mark H. 1994. *Morphology by Itself: Stems and Inflectional Classes*, Cambridge (Mass.): MIT Press, 1994..
- BAKER, Mona. 2011. *In Other Words: A coursebook on translation*, Second edition, London: Routledge, 2011.
- BLOOMFIELD, Leonard. 1933. *Language*, London: Allen & Unwin, 1933.
- BOOIJ, Geert. 2002. *The Morphology of Dutch*, Oxford: Oxford University Press, 2002.
- CHAMBERS. 1998. *The Chambers Dictionary*, Edinburgh: Chambers Harrap, 1998.
- CHOMSKY, Noam. 1981. *Lectures on Government and Binding*, Dordrecht: Foris, 1981.
- CHOMSKY, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use*, Westport (Conn.): Praeger, 1986.
- DRESSLER, Wolfgang U. 2006. Compound Types. In LIBBEN, G., JAREMA, G. (eds.), *The Representation and Processing of Compound Words*, Oxford: Oxford University Press, 2006, pp. 23-44.
- GREVISSE, Maurice. 1980. *Le bon usage: Grammaire française avec des Remarques sur la langue française d'aujourd'hui, onzième édition revue*, Paris/Gembloux: Duculot, 1980.
- ten HACKEN, Pius. 1994. *Defining Morphology: A Principled Approach to Determining the Boundaries of Compounding, Derivation, and Inflection*, Hildesheim: Olms, 1994.

- ten HACKEN, Pius. 2008. Prototypes and Discreteness in Terminology. In BERNAL GALLÉN, E., DECESARIS WARD, J. (eds.), *Proceedings of the XIII Euralex International Congress*, IULA-UPF, Barcelona, 2008, pp. 979-987.
- JACKENDOFF, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Oxford: Oxford University Press.
- KALLAS, Krystyna. 1999. Przymiotnik. In GRZEGORCZYKOWA, R., LASKOWSKI, R., WRÓBEL, H. (eds.), *Gramatyka współczesnego języka polskiego: Morfologia*, Warszawa: PWN, 1999, pp. 469-523.
- LABOV, William. 1973. The Boundaries of Words and Their Meanings. In BAILEY, C.-J.N., SHUY, R.W. (eds.), *New Ways of Analyzing Variation in English*, Washington DC: Georgetown University Press, 1973, pp. 340-373, repr. in B. AARTS, D. DENISON, E. KEIZER, G. POPOVA (eds.). 2004. *Fuzzy Grammar: A Reader*, Oxford: Oxford University Press, 2004, pp. 67-89.
- LEVI, Judith N. 1978. *The syntax and semantics of complex nominals*, New York: Academic Press, 1978.
- LIEBER, Rochelle, ŠTEKAUER, Pavol. 2009. Introduction: Status and Definition of Compounding. In LIEBER, R., ŠTEKAUER, P. (eds.), *The Oxford Handbook of Compounding*, Oxford: Oxford University Press, 2009, pp. 3-18.
- MATTHEWS, Peter H. 1974. *Morphology: An Introduction to the Theory of Word Structure*, Cambridge: Cambridge University Press, 1974.
- MUNDAY, Jeremy. 2008. *Introducing Translation Studies: Theories and Applications*, 2nd Edition, London: Routledge, 2008.
- NEWMAYER, Frederick J. 1986. *Linguistic Theory in America*, second edition, New York: Academic Press, 1986.
- NICOLADIS, Elena. 2002. When is a Preposition a Linking Element? Bilingual Children's Acquisition of French Compound Nouns. *Folia Linguistica*, 2002, 36:45-63.
- NORD, Christiane. 1997. *Translating as a Purposeful Activity: Functionalist Approaches Explained*, Manchester: St. Jerome, 1997.
- ROSCH, Eleanor. 1978. Principles of Categorization. In ROSCH, E., LLOYD, B.B. (eds.), *Cognition and Categorization*, Hillsdale NJ: Lawrence Erlbaum, 1978, pp. 27-48.
- SELKIRK, Elisabeth O. 1982. *The Syntax of Words*, Cambridge (Mass.): MIT Press, 1982.
- SZYMANEK, Bogdan. 2010. *A Panorama of Polish Word Formation*, Lublin: Wydawnictwo KUL, 2010.

VERMEER, Hans J. 2000. Skopos and Commission in Translational Action (transl. Andrew Chesterman, orig. 1989). In VENUTI, L. (ed.), *The Translation Studies Reader*, London: Routledge, 2000, pp. 221-232.

Notes

¹ This article is based on a presentation at the Second Conference on Universals and Typology in Word Formation, Košice 26-28 August 2012. I would like to thank the audience, in particular Pavol Štekauer and Bogdan Szymanek, for their questions and comments. I would also like to thank Fanny Messika and Ewelina Kwiatek for their help with the French and Polish translations, respectively. Needless to say, I am alone to blame for any remaining errors.

² The fact that (5a) and (5b) are written as one word, but (5c) is not cannot be used as an independent criterion for the boundary between compounds and phrases, but it does reflect the intuition that only (5a) and (5b) are single concepts.

³ “La préposition *à* s’emploie surtout pour marquer le lieu, le but, le temps, le moyen, la manière, la caractéristique” [my translation, PtH]

Pius ten Hacken
College of Arts and Humanities
Swansea University Singleton Park
Swansea, SA2 8PP
Wales, UK.
p.ten-hacken@swansea.ac.uk

In *SKASE Journal of Theoretical Linguistics* [online]. 2013, vol. 10, no. 1 [cit. 2013-02-04]. Available on web page <http://www.skase.sk/Volumes/JTL22/pdf_doc/06.pdf>. ISSN 1339-782X.