The University Of Sheffield.

Do Speakers Build the Categories Linguists Postulate? A Usage-Based Exploration

By:

Jaroslaw Jozefowski

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Arts and Humanities
School of Languages and Cultures

29th November 2018

Do Speakers Build the Categories Linguists Postulate?
A Usage-Based Exploration

By:

Jaroslaw Jozefowski

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

The University of Sheffield
Faculty of Arts and Humanities
School of Languages and Cultures

29th November 2018

Abstract

Linguists are naturally inclined to seek maximally general categories for the description of linguistic phenomena, e.g. *the present tense* or *the reflexive voice*. It has been taken for granted that speakers use the same categories in their daily experience with language. A few studies have indicated, however, that speakers might not be able to build some general constructions that linguists postulate (see e.g. Dąbrowska 2008a; Perek 2015). If we would like for our descriptions to reflect the linguistic knowledge of native speakers, we need to empirically investigate the cognitive reality of the categories we develop.

The main aim of this thesis was to investigate whether speakers build the categories linguists postulate and if so, how general these categories are. A number of corpus and experimental studies were conducted for Polish prefixed verbs and reflexive verbs, which explored categories of different levels of generality. The results of the studies suggest that speakers might build some general categories (e.g. the one for the Polish marker *siebie*), while they might not be able to build others (e.g. the ones for the different senses of the verbal prefix *po*-). These differences can be explained by the frequency with which the constructions occur as well as the nature of their typical contexts.

The above result underscores the importance of empirically veryfing the categories linguists postulate. Linguists must not tacitly assume that their linguistic descriptions are cognitively real because it cannot be assessed a priori whether speakers use them or not. Since speakers might not be able to construct for categories that are established in linguistics, such as verbal prefixes, some other ‘traditional’ linguistic categories might need revisiting and empirical verification.

TABLE OF CONTENTS

List of figures

List of tables

List of abbreviations

1PL:	first person plural
1SG:	first person singular
2PL:	second person plural
2SG:	second person singular
3PL:	third person plural
3SG:	third person singular
ACC:	accusative case
DAT:	dative case
FEM:	feminine
FUT:	future
GEN:	genitive case
IMPF:	imperfective
INF:	infinitive
INS:	instrumental case
MASC:	masculine
NEUT:	neuter
NOM:	nominative case
NON-VIR:	non-virile
PASTP:	past participle
PFV:	perfective
PL:	plural number
PRES:	present
PRESP:	present participle
PST:	past
REFL:	reflexive
SG:	singular number
VIR:	virile

Acknowledgements

First and foremost, I would like to thank my supervisors, Prof. Dagmar Divjak and Prof. Neil Bermel. Their exceptional patience and impressive knowledge have been instrumental in transforming this study from a few loosely coupled ideas into a coherent piece of argumentation. The greatest gift they could give me was teaching me how to question, doubt, and be sceptical – and for this gift I am most thankful. I would also like to thank all the academics I have come across during my journey with linguistics who have inspired me to study this subject, in particular my teachers at the University of Warsaw: Prof. Elżbieta Górska and Dr Agata Kochańska.

I would like to extend my deepest gratitude to my wife, Ola. This thesis would not have been completed but for her. Her unrelenting support, boundless patience, and sheer wisdom helped me persevere in what has been the most challenging endeavour of my life so far. Two more people without whom this dissertation would not have happened are my mum and my dad, who have always made sure I received sufficient education. Thank you for always having my back.

Last but not least, I am most grateful to all my friends for keeping up my spirits – even from a thousand miles away. Thank you for being out there.

Introduction

Do speakers use the same categories that linguists postulate? Linguists are naturally inclined to look for maximally general labels and neat categories that would apply to as many cases as possible, for instance, the reflexive marker. Are speakers so inclined too or do they not go along the lines set by linguists and build other, less general, categories? What is best, optimal, or most elegant from the descriptive point of view of a trained linguist does not need to be best or optimal for a native speaker of a language. In other words, linguists' language descriptions might not always be the most relevant as far as their cognitive reality is concerned (Divjak, Levshina & Klavan 2016: 451). We can never know whether speakers build the same categories as linguists postulate unless we investigate them empirically using real language data – be it corpus-derived, experimental, or observational.

The overarching aim of this thesis is to explore the relevance of general language categories for capturing the knowledge held by native speakers. To achieve that, two empirical studies based on corpus and experimental data on Polish verbs were conducted: a study on Polish reflexive verbs and a study of Polish prefixed verbs. For each of those phenomena, a corpus study first explored the question whether speakers might build maximally general constructions based on the language input they receive. Subsequently, experiments were run to investigate if speakers could build less general constructions for Polish reflexives and prefixes. The research in this thesis investigates different possible levels of language categories, progressing from maximally general categories to less general ones.

Why was Polish chosen as a source of data for the research presented in this thesis? Firstly, usage-based linguistics has demonstrated a tendency to focus predominantly on "West-European data (English in particular)" (Divjak, Levshina & Klavan 2016: 449). Empirical evidence coming from languages other than the most popular ones can help extrapolate the linguistic theory beyond the Western European domain. The data from Polish can broaden the scope of inquiry not only in terms of language families but also different aspects of language. Polish is a morphologically rich language (Gerz et al. 2018) with a relatively free word order

(Sadowska 2012: 42), which makes it different from English, whose constructions are predominantly syntactic. The inventory of English morphological markers is quite small, and, by consequence, investigations in usage-based linguistics have been biased towards English syntax. Studying Polish, with its complex morphology, offers an opportunity to obtain more data and evidence on how language categories might work by venturing beyond the domain of syntax.

Polish reflexive and prefixed verbs provide a very good case for studying the relevance of general categories for speakers of language, because linguists have described them in terms of categories of different degrees of generality. These postulated categories can be explored to assess which level of generality best reflects native speakers' knowledge of language. The Polish reflexive marker *się* has been hypothesised to express from one general meaning (Dancygier 1997; Tabakowska 2003a) up to seventeen different senses (Wilczewska 1966). Additionally, one dictionary of the Polish language[1] lists more than 7,000 verb + *się* combinations, which are essentially language categories of very low generality, each with its specific meaning. The fact that lexicographers decided to include so many reflexive verbs as separate entries indicates that speakers might hold a separate specific category for each reflexive verb, because such verbs often express much more specific meaning than the mere combination of a verb and a reflexive marker would. If we take all the above into consideration, we can see that speakers could potentially have categories for reflexives at any level of generality: from the most specific (verb + *się* pairings) to the most general (a single general meaning for the marker *się*). The same applies to Polish prefixes – they have been described in terms of general categories (see e.g. Swan 2002), and dictionaries also list many prefix + verb combinations as separate entries (e.g. 600+ entries for verbs with the prefix *przy-*[2]). In sum, both Polish language phenomena have been described in terms of categories of different levels of generality – this provides a perfect testing ground for how general the categories that speakers build can be.

Why do we need to study empirically whether speakers can build the categories linguists postulate? The primary reason is that if we investigate this question, we

[1] *Uniwersalny Słownik Języka Polskiego* (http://usjp.pwn.pl).
[2] In *Uniwersalny Słownik Języka Polskiego* (http://usjp.pwn.pl).

will be able to provide better and more cognitively accurate descriptions of linguistic structures. To provide accurate descriptions of grammatical regularities that would reflect the categories present in speakers' minds is one of the primary aims of linguistic enquiry. As Langacker put it, "[g]rammar consists of regularities … that speakers internalize and that linguists need to discover and describe" (Langacker 2003: 44). Linguists, however, have tacitly assumed that the categories they postulate for language regularities exist also in speakers' minds (Divjak, Levshina & Klavan 2016). Throughout the history of linguistics, language regularities have assumed a whole array of different names – linguists can pick and choose from *structures*, *rules*, *constructions*, *schemata* (or *schemas*), and some more. Of course, the choice of the label will depend on the theoretical commitment of a particular researcher. Thus, a Generative Grammarian might opt for **principles** or **rules** that underlie sentence construction (see e.g. Chomsky 2002 [1957]: 59); a Structural Linguist would be seeking to explain the **conventions** of language (de Saussure 1959 [1916]: 9–10); whereas a Cognitive Linguist would be looking for **schemas** (see e.g. Langacker 1987) or **constructions** (see e.g. Goldberg 2006). Paradoxically, irrespective of the different theoretical allegiances and the ontological rifts between the theories, many linguists have sought to develop categories that would achieve maximum generality. De Saussure proposed that linguistics should aim to "to determine the forces that are permanently and universally at work in all languages" (de Saussure 1959 [1916]: 6). Chomsky, even more ambitiously, proposed that the aim of linguistic study should be to "identify the specific nature of this distinctive human possession", which, in turn, is effectively an "effort to determine the genetic endowment of the faculty of language" (Chomsky 2007: 1). Both approaches call for maximum generality, because if grammatical principles applied to all languages of the world, they would inevitably have to be general enough to accommodate the immense variety in all those languages.

At the same time, many linguists do not preclude the existence of formulaic language, that is "[w]ords and word strings which appear to be processed without recourse to their lowest level of composition" (Wray 2002: 4). Formulaic sequences are expressions that speakers process and produce without the need to refer to more general categories of language. We could say that formulaic sequences are the opposite of general categories. De Saussure suggested that “when a compound concept is expressed by a succession of very common significant units, the mind gives up analysis—it takes a short-cut—and applies the concept to the whole cluster of signs, which then become a **simple unit**” (de Saussure 1959 [1916]: 177; emphasis mine). With the advent of Chomskyan Universal Grammar, the interest in formulaic sequences effectively waned, as only idiosyncratic forms were granted the status of single units of language, e.g. irregular verb forms in English or idioms. If a sequence of words could be derived from other expressions with the use of rules, it could not be granted the status of a unit – this status was reserved only for the “misbehaving” expressions. Over the last two or three decades, we have seen a resurgence of the inquiry into formulaicity in language (see Wray 2002: 7–11; Wray & Perkins 2000: 9–11). Now, not only idiosyncratic forms are seen as formulas or units – if an expression is frequent it can attain the status of a unit, even if it could also be decomposed into smaller units (Bybee 2010: 8). Perhaps, some general constructions postulated by linguists are in fact “bundles” of more specific low-level (e.g. lexical) constructions for actual speakers of language, who do not decompose them into smaller units governed by those general high-level constructions.

The two above strands of research – seeking general rules, conditions, or constructions on the one hand and investigating the existence of formulas on the other – appear to be two ends of the generality spectrum. If a sequence of words (or morphemes) becomes very frequent, it is potentially stored as unit, even if we could still derive it from a general construction. The resurgence of research into the nature of formulaic language, that is non-decomposable language categories, has not brought about much interest in the other side of the coin: the very existence of general grammatical constructions and their nature. How general can general constructions be? Does the logical possibility of developing a general grammatical category imply its existence in the minds of speakers? In other words, do speakers arrive at the same generalisations as grammarians do? If we subscribe to the major tenet of usage-based linguistics that linguistic structures arise from language use, we could investigate the above questions by studying how people use language –

be it through experiments or corpus data. Recently, experimental evidence has appeared that puts the existence of some general constructions in question, namely Polish dative case constructions with low type frequencies (Dąbrowska 2008c; Dąbrowska 2008a) and English questions with long-distance dependencies (Dąbrowska 2008c). If it has been found that speakers might not be able to build general categories quite established in the linguistic literature, perhaps some other general categories postulated by linguists, e.g. reflexives, would also not hold when subjected to the scrutiny of real-life language data. This thesis explores this question.

The thesis consists of eight chapters. Chapter 1 gives an overview of usage-based linguistics and its approach to language categories, which is fundamental for this study – it discusses the notion of a **construction**, the difference between general and specific language categories, the usage-based nature of language and the acquisition of language categories. Chapter 2 introduces the reader to the specific linguistic phenomena investigated in this thesis: Polish reflexive constructions and Polish prefixed verbs. It gives an overview of the functions of these (purported) language categories, report on the different accounts of their behaviour, and discuss the controversies and disagreements pertaining to Polish reflexives and prefixes.

Chapters 3 and 4 discuss the studies on Polish reflexives and Chapters 5 to 7 discuss the studies on Polish prefixes, which form the empirical part of this thesis. Chapter 3 presents a corpus study on Polish reflexive verbs; the behavioural profiles (Divjak & Gries 2006) methodology was employed to investigate the question whether native speakers of Polish could build a maximally general category for the reflexive marker *się*. Chapter 4 explores Polish reflexives further with a sentence-sorting experiment, which investigated whether native speakers of Polish could have (less) general categories for the **different senses** of *się*. Chapter 5 presents a corpus study on three Polish prefixes *po-*, *przy-*, and *roz-*; analogously to the corpus study on Polish reflexive verbs, the study discussed in Chapter 5 employed the behavioural profiles methodology to establish whether speakers could build maximally general categories for the prefixes *po-, przy-,* and *roz-*. Chapters 6 and 7 discuss two experimental studies – a sentence-sorting experiment and a nonce-verb experiment – designed to investigate whether speakers of Polish build categories for the different senses of prefixes *po-, przy-,* and *roz-* or whether they rather generate ad hoc categories for the different senses of these prefixes if they are forced to do so by communicative circumstances.

The empirical studies presented in Chapters 3 to 7 jointly explore the question whether speakers could build categories that linguists postulate. The studies proceeded from one maximally general category for one form (e.g. one category that would apply to all instances of *się*) to many less general categories (e.g. different senses of the prefix *po-*). Through investigating different degrees of generality, the studies provided data necessary to analyse the question of whether speakers' categories converge with any of the categories postulated by linguists and establish the highest level of generality at which speakers could build categories.

The thesis closes with Chapter 8, which presents the conclusions drawn on the basis of the results of the studies presented in Chapters 3 to 7. The implications for the study of Polish prefixed verbs and Polish reflexives are discussed first and then, methodological and theoretical conclusions follow.

Chapter 1: Usage-based linguistics and language categories

1.1. Introduction

This thesis will focus on abstract linguistic categories, their origins in the actual language use, and how we could test whether a general pattern proposed by linguists exist in the minds of "naïve" native speakers. The research presented in this dissertation is done within a usage-based linguistic framework. The main tenet of usage-based linguistics is that linguistic knowledge arises from language use – it is not innate, and each language user needs to acquire the linguistic conventions of a language from the input she receives. Consequently, in order to check whether speakers of a language develop a given abstract construction, we need to assess whether the input to which speakers are exposed – that is actual instances of language use – could allow the speakers to generalise beyond those instances to come up with more general constructions. Subsequently, we need to test whether speakers actually form those generalisations. We also need to entertain the option that speakers might not be able to abstract over the input to form more general constructions. In consequence, a general construction postulated by linguists could, in reality, exist in the minds of speakers as many smaller lexically-specific constructions.[3] In other words, the sole fact that a general construction **can** be postulated does not automatically imply that it **will** be built by speakers on the basis of input they receive.

Chapter 1 will first outline the major tenets of usage-based linguistics and then introduce the usage-based linguistic approach to linguistic constructions. Then, the chapter will present an overview of the principles of categorisation and discuss the usage-based approach to general and specific language categories. Subsequently, it will discuss how general constructions could be built from language use and whether speakers build all constructions that linguists postulate. Finally, the

[3] Assessing the overall number of constructions in a language presents a major challenge to linguists and has become a contentious issue in the field. One might, for instance, count all the verbs in a language as in collostructional analysis (Stefanowitsch & Gries 2003).

chapter will give an overview of studies on how frequency of occurrence relates to the emergence of linguistic constructions.

1.2. Usage-based linguistics and language constructions

In usage-based linguistics, "linguistic knowledge is represented by a network of form-meaning pairings called **constructions** [emphasis mine – JJ]" (Dąbrowska 2014: 618). This thesis will use the term *construction* both for very general form-meaning pairings such as the passive voice and more specific ones such as *come up with* (an English phrasal verb). One of the key tenets of usage-based linguistics says that linguistic constructions arise from real usage events, that is "actual instance[s] of language use" (Langacker 2000: 9). More technically, a usage event could be defined as "the pairing of a vocalization, in all its specificity, with a conceptualization representing its full contextual understanding" (ibid.). In the usage-based approach, a language user must learn all conventions of the language she speaks based on the linguistic input she receives using the general cognitive abilities she would use for the learning of any other "skill" (Langacker 2000: 2). This approach eschews the existence of innate grammatical structures postulated by the proponents of Universal Grammar. Universal Grammar postulates that humans are genetically endowed with a blueprint for language (syntax in particular). The acquisition of a language in Universal Grammar would consist in mapping the categories of a particular language onto the underlying Universal Grammar structures. This process would be instantaneous, and even a "single … trigger in the input [would be] … sufficient to acquire a particular linguistic category" (Diessel 2013: 348). According to usage-based linguistics, a language user cannot rely on pre-existing rules and features that she could just turn on or off. Instead, language users must abstract over many usage events in order to arrive at a grammatical construction. Simply put, "language structure emerges from language use" (Tomasello 2003: 5).

If we assume that people are not genetically endowed with linguistic knowledge and that such knowledge arises from actual usage events, the primary method of studying language would be to study the actual usage of language by its speakers. In other words, usage-based linguistics puts strong emphasis on using empirical data as evidence for any theoretical claims. Tummers, Heylen & Geeraerts went as far as to say that "you **cannot have a usage-based linguistics** unless you study actual usage – as it appears in an online and elicited form in experimental settings, or as it appears in its most natural form in corpora in the shape of spontaneous, non-elicited language data [emphasis mine]" (2005: 226). We can see that two key methods of gathering empirical language data are behavioural experiments and corpus queries. The advent of large-scale digitalised and easily searchable corpora enabled researchers to use massive amounts of non-elicited (i.e. spontaneous) usage events (Tummers, Heylen & Geeraerts 2005: 232) and analyse them statistically to discover regularities in language. As far as experimental studies are concerned, the traditional offline methods such as questionnaires or interviews have in the recent decades been supplemented by high-tech solutions that enable researchers to investigate the online processes that occur in the human brain when producing or interpreting language, e.g. EEGs or eye-tracking. The available corpora, and experimental methods and statistical techniques have (finally) rendered it possible to study whether linguistic theories stand up to the scrutiny of empirical data.

The assumption that humans do not have a universal genetic blueprint for language also means that language development cannot consist of turning some pre-defined features on or off. Instead, usage-based linguistics postulates that speakers acquire language with general cognitive abilities – "[a] usage-based model … takes as its null hypothesis the view that language is an extension of other cognitive domains." (Bybee & Beckner 2010: 829). The most important of those abilities are: categorisation and statistical and sequential learning. As one of the primary functions of language consists in helping humans to categorise the world around them, research on categorisation has an important role in usage-based linguistics. Especially exemplar-based models of human categorisation are particularly suited to the assumption that language arises from usage, because exemplar-based models postulate that categories are built based on many instances (exemplars) of a particular category that a person encounters over time. In other words, in exemplar-based models, categories arise from experience – just like language arises from usage. Different models of categorisation and how they relate to linguistics will be discussed in Section 1.3. The frequencies of occurrence of different patterns must be stored in order to determine whether a linguistic pattern is frequent and prominent enough to become a linguistic category. Frequency and its relation to usage-based linguistics will be discussed in Section 1.7.

In usage-based linguistics, a construction is the basic unit of linguistic organisation. Goldberg (1995: 4) defined constructions as pairings of form and **meaning**, or, later, as "learned pairings of form with semantic or discursive **function**" (Goldberg 2006: 5). For Langacker, a construction is "either an expression (of any size), or else a schema abstracted from expressions to capture their commonality (at any level of specificity)" (2003: 43). These two definitions complement each other, as they emphasise different aspects of constructions – the former focuses on the functional nature of constructions, whereas the latter highlights the usage-based origins of constructions by stating that some of them are abstracted from expressions that are more concrete. According to Goldberg, "[a]ny linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist"; that notwithstanding, if a pattern is fully predictable from other established patterns, it should be considered an individual construction, provided that it occurs with "sufficient frequency" (2006: 5). Contrary to the approach advocated by generative grammar, in usage-based linguistics, language structures (i.e. constructions) are not abstract combinatory rules devoid of meaning whose 'slots' are filled with lexical times that carry all the meaning. In usage-based linguistics, each construction must have a meaning (or a function) – the difference between lexical and grammatical (e.g. syntactic or morphological) items is that the meaning of the latter is more abstract and goes beyond concrete lexical items. Language forms a continuum of categories from the most concrete (i.e. lexical) to the most abstract (syntactic) items, with multiple levels possible in between. Before we explore how linguists approach general and specific language categories, we need to discuss how categories are structured, and how they are formed in people's minds. The next section will present an overview of the main approaches to categories and categorisation.

1.3. Categories and categorisation

Established linguistic constructions are essentially mental categories. In order for a construction to be conventionalised in a linguistic community and for speakers to re-use the construction, the construction must be recognised as a category so that we know that for a given type of communicative situation, we must use this particular construction. The situation is analogous to using the category DOG – to use this category, we must first recognise a particular type of creatures as a coherent category DOG.

"Concepts are the glue that holds our mental world together" (Murphy 2004: 1). The concepts that we build are mental representations of categories, i.e. classes of objects or events in the world – be it 'real' world or abstract world (ibid). If we did not form mental concepts for 'similar' objects, we probably would not be able to interact with the world, as every single event and object would be entirely unique. In other words, "[t]he capacity to classify stimuli into a limited number of categories, i.e. to organize and structure objects in the world around us, is one of the most fundamental abilities in cognitive functioning: categorizing stimuli is one of the cognitive operations … that make the world more predictable, because many unknown properties of newly encountered stimuli can be induced with sufficient certainty as soon as the stimulus is recognized as a member of a certain category" (Divjak & Arppe 2013: 222). The categories that we use in our daily lives feel entirely natural, and because of that, we tend to think that it is not "a great intellectual achievement to identify" (Murphy 2004: 1). Nevertheless, the process of learning (or building) and using categories is far from simple, and a number of attempts have been made to explain the nature of categorisation. Those attempts can be broken down into three major strains: the classical view, prototype theories, and exemplar theories. This section will briefly introduce the three approaches, with the greatest emphasis on the latter two, since they have been extensively used in usage-based linguistics.

The classical view stipulates that categories are represented in human minds as **definitions** (Murphy 2004: 11). A good definition should enable people to only classify as members of a given category the entities that actually belong to this category. In the classical view, a good definition consists of a set of **necessary** and **sufficient** conditions. Let us take the category SQUARE as an example. The following definition of a square consists of a set of necessary and sufficient conditions: “[a square is] a closed flat figure, with four sides of equal length, and all interior angles equal” (Aitchison 2004: 2). Each of the conditions in the definition is necessary, while all of them combined form a set that is sufficient for a figure to be a square. This definition will “capture” all instantiations of the category SQUARE, regardless of whether they are, for instance, black or blue, and each instantiation will be as good an example of the category as any other. The world of classical categories is a world of neatly packaged and clearly delineated objects whose category memberships are precisely defined.

While the world of classical categories is clearly delineated, the categories in the ‘real’ world appear to defy definitions. Such categories as SQUARE can be precisely defined, but as soon as one tries to define a less artificial category – BIRD, for instance – some difficulties emerge. Let us define a bird as an animal that has two wings, a beak, a tail, is covered in feathers, walks upright, and flies. What about ostriches, kiwis, or chickens? Should they not be included in the category BIRD because they cannot fly? It is difficult to find a set of criteria that would pick out all members of a category and exclude every other object – some category members will not have **all** the features deemed necessary to be a member of the category in the classical view. A kiwi would not be a perfect example of the category BIRD, and it might take people more time to decide whether a kiwi is a bird or not. This phenomenon is called typicality effects, and it constitutes a problem that the classical approach cannot reliably account for.

The problems of classical approaches with accounting for typicality effects and borderline cases (such as kiwis) made researchers look for other theories of categorisation. Rosch (see e.g. 1975; 1978) suggested that categories are not definitions made of necessary and sufficient conditions, instead, categories are based on prototypes. Prototypes could be defined as summary representations, that is, sets of “features that are **usually** found in the category members, [where] some features are more important than others” (Murphy 2004: 42, emphasis mine). In contrast to the classical approach, not all features are of the same importance, and an item may even lack some features, but might still be classified as a member of a given category – the prototype approach will have no problem classifying a one-winged bird as a bird. Real-world categories have inherently unclear boundaries (they are "fuzzy"), because they need to allow for classifying the less obvious examples. The reason for this is that we have only a limited number of category concepts but a virtually unlimited number of objects to classify (Murphy 2004: 21) – if our categories were not flexible enough, we would end up with many very detailed categories whose definitions would be very restrictive. The prototype approach can also account for typicality effects: the typical examples of the category will have most of the crucial features present, while the less typical examples might miss one or more important features, for instance, a kiwi bird will not have the "can fly" feature of a typical bird. The importance of each feature is expressed as a weight: wings might be a very important feature of a bird, and they will be given a large weight, while the attribute "can sing" might be less important and thus will be given lesser weight. Categorising an item in this approach consists in comparing the sum of an item's features multiplied by their weights with the sum of prototype's weighted features and then calculating a similarity rate.

What if human beings did not need any summary representations with features at all to effectively categorise objects? According to the **exemplar view**, instead of "extracting features" from all examples (or exemplars) of a category to form prototypes, humans represent categories as "detailed memory traces of all the individually encountered exemplars of the concept; in the most extreme version of exemplar theory **no abstraction** would take place across these stored exemplars" (Divjak & Arppe 2013: 224, emphasis mine). For example, the representation of the category BIRD would be tantamount to all birds we have encountered and can remember – no abstract features would be extracted from the exemplars. Of course, the memories of some exemplars will have faded away, while others will be incomplete, but still, all available exemplars will be accessed when deciding whether an item belongs to a given category or not (Murphy 2004: 49). In the prototype approach, to categorise an item, one needs to calculate how similar the item is to the prototype and then decide whether the similarity is sufficient enough for the item to belong to the category. The **exemplar view** stipulates that one needs to compare the item being categorised to all remembered exemplars of a given category and assess whether it is similar enough to those exemplars to be classified as a member of the category. In other words, the prototype view suggests that during categorisation, we compare the item being categorised to the features of an abstract summary representation (the prototype), while the exemplar view suggests that we compare the item being categorised to **all members** (exemplars) of a category we can recall. Exemplars can also explain typicality effects – the most typical members of the category will be similar to the largest number of exemplars of a given category

The two constructs – prototypes and exemplars – might seem to be incompatible with each other and have usually been treated as competitive explanations of categorisation phenomena (Divjak & Arppe 2013: 224). One way of looking at prototypes and exemplars however, is not to regard them as two conflicting theoretical constructs but rather as two ends of an abstractness continuum (Verbeemen et al. 2007: 540). Exemplars, being concrete instances of categories, would come in at the least abstract end of the continuum. Prototypes, on the other hand, are summary representations outlining the characteristic **features** that result from the abstraction over all encountered members of a category – consequently, they would occupy the most abstract end of the continuum (see Figure 1.1.). In other words, "[t]he exemplar representation corresponds to minimal abstraction, and the prototype representation corresponds to maximal abstraction" (Vanpaemel & Storms 2008: 733).

Figure 1.1. A continuum of abstractness

Some evidence exists that categories intermediate between exemplars and prototypes might carry the optimal degree of abstraction. Vanpaemel & Storms (2008; 2010) revisited many datasets used for the investigation of the exemplar theory experiments (e.g. Nosofsky 1986; Nosofsky 1987; Nosofsky 1992), and tested whether a model based on intermediate levels of abstraction could explain the data better. In most cases, the re-analyses showed that a categorisation solution that involved **some abstraction** – that is, more abstraction than in “pure” exemplars and less than in “pure” prototypes – was the optimal solution and explained the data better than extreme models (either exemplar or prototype). For only a handful of datasets, the fully concrete exemplar solution involving no abstraction at all was the optimal one. Interestingly, the prototype-based explanation (involving full abstraction) did not constitute the optimal solution in any of the datasets. In other words, Vanpaemel & Storms (2008; 2010) found strong evidence for the existence of categories whose abstractness is intermediate between that of fully concrete exemplars and fully abstract prototypes.

One more aspect of categorisation begs attention – the strength, or rather the permanence, of categories. So far in this section, we have tacitly assumed that people either ‘have’ a category or not, and when they have the category, it remains ‘permanently’ in their minds. Let us take the category DOG for instance – once someone develops the category, they will be able to categorise dogs as instances of the category DOG. These types of categories, also called *common categories*, are well-established in people’s memory (Barsalou 1983: 211) and they can be spontaneously activated upon encountering an example of a category. What about categories such as THINGS TO TAKE ON A TRIP TO LISBON? Such categories are called *ad hoc categories*, and they are created to achieve particular goals (Barsalou 1983) – in this case, it would be packing oneself for a trip to Lisbon. It would be quite unlikely that one would develop a ‘permanent’ category for situations like this, unless one took a trip to Lisbon on a regular basis – the category THINGS TO TAKE ON A TRIP TO LISBON will be established for the fulfilment for one particular goal. This category will most likely not come to one’s mind upon seeing a T-shirt without the context of preparing for a trip to Lisbon, because the representation of such a category in memory is too weak or non-existent (Barsalou 1983: 224) – one will much more likely categorise the T-shirt as a member of the categories T-SHIRT or CLOTHING. If people can create ad hoc categories to achieve particular goals, we must also entertain the possibility that speakers could create ad hoc categories to achieve particular communicative goals, for instance, to understand unfamiliar linguistic input.

To sum up, the two dominant modern theories of categorisation, the prototype theory and the exemplar theory, propose two opposite accounts of the categorisation process – the former postulates that people categorise objects and events in the world using fully abstract summary representations (prototypes), while the latter postulates that categorisation takes place via comparing objects or events to previously encountered exemplars, which represents no abstraction at all. These two approaches might not be entirely incompatible, as they might actually represent two extremes on a spectrum of abstractness with some intermediate types of categories in between – and these "in-between" categories might sometimes constitute the optimal solution for categorisation. The following subsection (1.2.2.) will provide an overview of how linguists approach abstract and specific language categories, with an emphasis on usage-based approaches.

1.4. General vs specific constructions

"[G]enerality is a virtue" and linguists are right in that they "seek general rules and universal principles" (Langacker 1987: 45). At the same time, however, Langacker warns against positing rules that are overly general and thus eschewing or ignoring idiosyncrasy in the quest for maximum generality. An apparent paradox emerges when one tries to come up with increasingly general categories – the more abstract a generalisation becomes the more exceptions and ad-hoc sub-rules one needs to postulate to accommodate the linguistic phenomena the original generalisation was meant to account for. As we could see in the above definitions of constructions, Construction Grammar (Goldberg 1995; 2006) and Cognitive Grammar (Langacker 1987; 1990; 1991 and later), as well as usage-based linguistics in general, allow for a large degree of redundancy. If linguistic patterns are frequent (or entrenched) enough, they are likely to be stored as units in the minds of native speakers. A pattern might be stored as an independent unit (or construction) even if it could potentially be generated by a more general construction.

Some approaches in linguistics, however, eschew any sort of redundancy. In Generative Grammar, for instance, the two principles: *economy of representation* and *economy of derivation* state that language(s) avoid any "'superfluous elements' in representations and derivations" (Chomsky 1995: 130). In other words, such approaches as Generative Grammar aim to remove from grammar all elements for which more general rules could be posited. Strangely enough, the economy of derivation and representation is not tantamount to "computational efficiency". Chomsky suggests that "language design … appears to be in many respects 'dysfunctional'" (1995: 162) and that we need not expect that language is designed in a way that optimises the efficiency of use. In sum, language in Generative Grammar is economical in terms of storage (or representation) but uneconomical in terms of processing (computation). Usage-based linguistics postulates a different kind of economy: *economy of processing* or *economy of use*. Economy of processing manifests itself in automatisation processes in language (Langacker 2008: 16–17). Automatisation takes place when a pattern is used so often that it becomes entirely automated and thus achieves the status of a unit of language, defined as a pairing between form and meaning that is sufficiently entrenched in the minds of a language community. If a speaker has an entrenched and automated unit at her disposal, e.g. *moonless night*, she does not need to retrieve all the individual pieces that make the unit (i.e. *moonless* and *night*) and assemble them in order to be able to convey the meaning. Instead, the speaker must only retrieve the unit as a whole, and she will be immediately ready to use it in linguistic interaction. Langacker compares this process to how tying a shoe becomes "thoroughly mastered" (ibid.) when one has repeated the action multiple times. Automatisation obviates the need for "conscious monitoring" (ibid.), hence, we could say that it reduces the amount of processing resources required for performing a given action, e.g. tying shoelaces or uttering an expression. At the same time, Langacker (2008: 17) warns that "unit status does not entail the absence or unimportance of components, merely the routinized nature of their execution". In the case of *moonless night,* the unit status of this particular multi-word pattern does not mean that its components do not exist as separate units. As a result, we have three different expressions automatised and stored as units of language. What follows is that automatisation, as conceived by Langacker, requires a lot of redundancy to work efficiently because many entrenched (i.e. automatised) expressions become units, even though they could also be generated from a more general construction.

In Construction Grammar, Goldberg (e.g. 2006; 2009) has also recognised the importance of lower-level generalisations in addition to more general constructions and proposed a continuum of constructions of varying levels of schematicity. As we can see in Table 1.1., Goldberg's classification includes very general constructions such as the Ditransitive Passive on the one hand, and very local generalisations such as idioms or complex words on the other.

CONSTRUCTION TYPE	EXAMPLE
Word	e.g. tentacle, gangster, the
Word (partially filled)	e.g., post-N, V-ing
Complex word	e.g. textbook, drive-in
Idiom (filled)	e.g. like a bat out of hell
Idiom (partially filled)	e.g. believe <one's> ears/eyes
Covariational Conditional	The Xer the Yer (e.g. *The more you watch the less you know*)
Ditransitive	Subj V Obj1 Obj2 (e.g. *She gave him a kiss; He fixed her some fish tacos*)
Passive	Subj aux VPpp (PPby) (e.g. *The cell phone tower was struck by lightning*)

Table 1.1. Constructions of different levels of schematicity (after Goldberg 2009: 94)

Partially filled words, such as N-book (*textbook, coursebook, notebook)*, provide a good example of how usage-based linguistics deals with more local generalisations. Even though *textbook* could potentially be derived from a very general N(oun)-N(oun) construction, a more local construction is proposed, namely N-book. The reason for positing such a specific low-level construction is most likely the relatively high frequency of such word combinations (Goldberg 2006: 5). Since speakers use some unit combinations frequently, they are produced automatically without conscious regard for their internal composition. Thus, they effectively become similar to tying shoes or other everyday actions – we carry them out without thinking what the next step in the sequence will be. This alleviates the pressure on working memory and frees up attentional resources, which can then be used to attend to other stimuli or perform other tasks (see e.g. Beilock et al. 2002: 6). Langacker suggests that less general constructions (and even specific expressions), "expressing regularities of only limited scope, may on balance be more essential to language structure than high-level schemas representing the broadest generalizations" (Langacker 2000: 29).

Less general constructions are hypothesised to "be more essential to language structure" (Langacker 2000: 29) and reduce the amount of processing resources needed to produce language (Langacker 2008: 16–17). A crucial question is in order here: how are more specific constructions – and, consequently, constructions in general – acquired? Do children form larger units by stringing together, or chunking, smaller elements according to pre-established and regular patterns, e.g. do they build an N-book construction from the more general N-N construction? Or, perhaps, children start off with very concrete expressions only to abstract more general constructions from them? The following section will cover the most important findings on the acquisition of linguistic categories in the paradigm of usage-based linguistics.

1.5. Acquiring constructions

The question from the closing paragraph of the previous section was: how are constructions acquired? Tomasello (2003) suggests that children acquire language by learning concrete expressions they can use in very specific situations and only then do they formulate more general constructions, building upon those concrete expressions. Let us now analyse this process in more detail. First, a child starts producing simple one-word utterances that express a "holistic undifferentiated communicative intention", e.g. *bath* to describe an entire event of bathing (Tomasello 2003: 37); those utterances are called **holophrases**. Sometimes, a holophrase can consist of a few words that have a clear communicative function, e.g. *lemme-see*.

Next, children begin to put two holophrases together to describe a communicative situation, e.g. *ball table* to say that a ball is on or under a table. Those expressions are still very concrete in the sense that neither part of them belongs to a larger (more general) category (Tomasello 2003: 114). In other words, *ball* in the expression *ball table* is not yet part of the category of nouns or subjects, or, for instance, “moveable objects”, etc. More or less at the same time as they start to produce two-word utterances, children also begin to use expressions with a more systematic structure: **pivot schemas**. The expressions in question are composed of an item that defines the communicative intention, e.g. *more*, and a slot that can be filled with a number of different words. Thus, children can formulate many expressions such as *more juice, more milk* or *more cereal*, which, however, always rely on one concrete item. What is important here is that those expressions have no internal structure, or syntax, so *milk more* would be practically equivalent to *more milk* (Tomasello 2009: 76).

Shortly afterwards, **item-based schemas**, that is expressions with the first signs of syntactic marking, are produced. In contrast to pivot schemas, in item-based schemas, children must use grammatical devices to structure the utterance. For instance, in a sentence *John hit Mary*, the English word-order for transitive events determines which participant was the agent and which one was the patient. Despite their being more abstract than pivot schemas, item-based schemas remain entirely reliant on the item they are based on, e.g. the verb *hit*. At this stage, children cannot yet extrapolate the knowledge about a limited number of items to an abstract transitive construction (Tomasello 2009: 77–78) .

Only between two and three years of age, children start to use their first more **general constructions** such as datives, locatives or passives in English. Young speakers’ constructions gradually increase in generality, but even at five or six years of age, they can still rely on more concrete patterns, including fixed items. Diessel & Tomasello (2000) investigated the development of relative clauses in children between the ages of 1;9 and 5;2, and they found that most instances of relative clauses were based on a few very similar recurring patterns. In one of those patterns, the sentences always began with a fixed presentational phrase such as *Here’s the…* or *That’s the…* , as in *Here’s the tiger that’s gonna scare him* (Diessel & Tomasello 2000: 137).

In summary, during language development, speakers arrive at more general constructions gradually – through exposure to linguistic input. They begin with concrete expressions with no internal structure and slowly generalise over those expressions to build categories that increase in generality. Children's utterances that seemingly use general grammatical constructions can, in fact, long remain based on more concrete patterns (see e.g. Diessel & Tomasello 2000). Some research suggests that language acquisition is an endeavour that never stops, and continues well into adulthood. Not only does the process of language acquisition continue into adulthood, but it can also give different results. In other words, contrary to what other theories (e.g. Generative Grammar) say, adults speaking the same language might build different grammars of different levels of generality – they will not always develop the most general constructions possible. The grammars that (at least some) speakers arrive at might diverge from the grammars proposed by theoretical linguists. The next section will report some studies that explore the issue of speakers' grammars vs linguists' grammars.

1.6. Linguists' grammars vs speakers' grammars

Even though they recognise and emphasise the importance of lower-level, local, or more concrete, constructions, usage-based linguists seem to focus mainly on the more general constructions. The area of lower-level constructions has been explored many times (see e.g. Siyanova-Chanturia 2014 for an overview). Those studies have predominantly made an assumption that lower-level constructions co-exist with the general constructions they can be derived from. However, the empirical reality of general constructions has hardly been explored at all.

As mentioned earlier, Diessel & Tomasello (2000) found that children as old as six years of age continue to use more local constructions, or templates, to produce sentences which have been interpreted in other linguistic paradigms as produced with the use of a more general construction. It turns out that even adult language production can rely heavily on a limited number of templates (Dąbrowska 2008c; Dąbrowska, Rowland & Theakston 2009). Evidence has also been found that adults might not develop some general categories at all (Dąbrowska 2008a). Let us now briefly discuss these studies below.

Dąbrowska (2008c) and Dąbrowska, Rowland & Theakston (2009) studied English questions with long-distance dependencies (LDDs), such as the following:

(1.1.) *When did you say you would come back?*

This type of questions contains a dependency between the question word (e.g. *when*) and a 'gap' at the end of the sentence. Technically, any number of clauses could be placed between the question word and the gap, and examples such as the following one have been proposed:

(1.2.) *Which problem$_1$ do you think (that) Jane believes (that) Bill claims (that) Mary solved ________$_1$?* (Ouhalla 1994: 71 in Dąbrowska, Rowland & Theakston 2009: 572).

In (1.2.), the gap has been marked with a 'blank' and a subscript 1, while the question words (*which problem*) related to the gap have been marked with a subscript 1.

A query in the spoken part of the British National Corpus by Dąbrowska (2008c: 573) demonstrated, however, that questions with LDDs hardly ever come in the form similar to the question in (1.2) and are usually very stereotypical – most of them can be generated with two templates: *WH do you think S-GAP?* or *WH did you say S-GAP?* (see the example in 1.1.). Most sentences that were different from the template contained only minor changes, and only 4% substantially departed from either of the prototypical for. In other words, real usage examples exhibit much less variety than examples put forward by formal linguists (Dąbrowska 2008c: 393). Dąbrowska (2008c) conducted a grammaticality judgment experiment in which participants had to rate sentences that either conformed to the two prototypical templates (*WH do you think S-GAP?* or *WH did you say S-GAP?)* or departed from them in various ways, e.g. a the verb was changed (*WH do you* ***believe*** *S-GAP?*) or the question or more than one clause intervened between the WH-word and the gap (*What do you think [Jo believes]$_1$ [he said at the court hearing]$_2$?*). The results showed a significant prototypicality effect for questions with LDDs – the more a question departed from the prototype the less grammatical it was judged by the participants. Dąbrowska (2008c) conducted the same kind of experiment but for declaratives with verb complement clauses, e.g. *But you think the witness will say something if they don't intervene.* In contrast to questions with LDDs, declaratives did not exhibit any significant prototypicality effects. This result goes in line with the results of the corpus query, which showed that declarative sentences are much more varied than their LDD counterparts.

To explore how LDD questions are acquired, Dąbrowska, Rowland & Theakston (2009) conducted a series of experiments with English children aged from 4;6 to 6;9 and adult controls. The participants were asked to repeat different types of questions with LDDs and their declarative counterparts: prototypical (*What do you think the boys will really like?*), unprototypical (*What does the man really hope they will like?*), and deeply embedded (*What do you think he said they will like?*) (all examples from Dąbrowska, Rowland & Theakston 2009: 577). It was significantly easier for children to repeat the more prototypical instances of both LDDs and declaratives, which suggests that children rely heavily on lexically specific templates for both questions with LDDs and declaratives with verb complement clauses. In adults, a similar result was obtained only for questions, not for declaratives – which is in line with the previous experiment by Dąbrowska (2008c). Overall, Dąbrowska, Rowland & Theakston's (2009) experiment results suggest that children as old as almost seven years of age do not yet exhibit full linguistic proficiency, and that grammatical development continues well beyond childhood. It also provides further evidence that even adults might not take full advantage of very general constructions, relying on lexically specific templates instead.

In the area of morphology, Dąbrowska (2008a) investigated Polish dative inflections. Polish datives are a good ground for testing specific vs general constructions, because they are quite regular. There are only four Polish dative endings: *-owi, -u*, *-i/y,* and *-ie* (Swan 2002: 45, 67, 112). The ending for a given noun is chosen based on the noun's gender: masculine nouns usually take *-owi,* neuter nouns take *-u*, whereas feminine nouns take either *-i/y,* or *-ie* (Dąbrowska 2008a: 935). Polish nouns (of any gender) can be divided into a limited number of phonological "neighbourhoods", that is clusters which comprise phonologically similar nouns. The number of nouns in each neighbourhood varies, which means that some neighbourhoods consist of many nouns (e.g. feminine nouns ending in -*arka*), whereas others are populated by a low number of items (e.g. masculine nouns ending in *-olog*). Dąbrowska (2008a) conducted a series of experiments to see whether native speakers of Polish supply endings equally accurately to nouns from densely and sparsely populated neighbourhoods. The participants in the experiments were given nonce nouns resembling nouns from different phonological neighbourhoods and were asked to inflect them in a meaningful context (which was also supplied). In the final experiment, participants were tested on their ability to supply endings to real nouns, which was the control condition. As far as the results are concerned, feminine and neuter nouns showed significant neighbourhood density effects in the nonce-noun studies – the endings for nouns resembling real nouns from high-density neighbourhoods were supplied much more accurately. As far as nouns from low-density neighbourhoods are concerned, participants often did not supply the "correct" ending or did not supply any ending at all. Masculine nouns showed only a slight effect. It must be emphasised that nouns of one gender from both densely and sparsely populated neighbourhoods take **the same endings**, so it was not the case that the participants did not have a general construction from which they could infer the ending. Dąbrowska (2008a: 944) sees this result as evidence that speakers prefer to rely on lower-level constructions, even though an overarching general construction could theoretically be arrived at. Interestingly, between a third to a half of participants could not supply the target ending to **any** of the **neuter nouns** from low-density neighbourhoods. This suggests that some speakers might not form general constructions for some categories at all.

The number of errors in Dąbrowska's (2008a) study correlated strongly with participants' education: less-educated participants made significantly more errors in the nonce-noun inflection task. Dąbrowska hypothesised that the differences are attributable to participants' overall experience with various kinds of language (2008a: 947), which translates into the amount of different nouns they had experienced in the dative case. The dative case is usually used with animate participants, which are predominantly masculine or feminine in Polish – neuter nouns are normally inanimate. Neuter nouns can be used in the dative, however, it is mostly in a limited number of prepositional constructions, which are "decidedly high-register or even archaic, and therefore they are found primarily in formal written texts" (ibid.). The low frequency of occurrence of neuter nouns in the dative might have caused that less-educated participants experienced far fewer instances of neuter nouns inflected in the dative, and the exposure might not have been enough for them to form a general construction that would allow them to inflect nonce nouns.

When interpreted together, Dąbrowska's (2008a; Dąbrowska 2008c) and Dąbrowska, Rowland & Theakston's (2009) studies indicate strongly that limits might exist as to how general some constructions are in actual speakers of a given language. The prototypicality effects in questions with LDDs suggest that speakers do not need very general constructions to be able to produce and comprehend even complex structures. Moreover, "LDD questions are fully acceptable only with particular lexical content[, which] suggests that they are more like a **constructional idiom** than a fully general construction" (Dąbrowska 2008c: 418; emphasis mine). The term 'constructional idiom' seems particularly apt, because it puts on a par constructions that have long been regarded to belong to separate modules of language: idioms (lexicon) and grammatical or syntactic constructions. Dąbrowska's (2008a) experiments on the Polish dative provided further evidence that speakers might prefer to rely on lower-level constructions. Additionally, some speakers (e.g. those with lesser linguistic experience) might not form certain general categories at all. Of course, general constructions can still exist – declarative sentences with verb complement clauses are a case in point. Dąbrowska nevertheless, suggests that "lexically specific variants have a privileged status, in that they are ontogenetically earlier, apparently easier to access, and preferred by speakers" (2008c: 593).

As rare as they are, studies on lower-level vs more general constructions are not limited to experimental methods. Perek (2015) took a collostructional[4] approach to investigate alternations in English argument structure constructions using corpus data, specifically the conative construction. The conative alternation can be seen in sentences of the following kind:

(1.3.) a. *Bill kicked the ball.*

[4] For more information on collostructional analysis see Stefanowitsch & Gries (2003); Gries & Stefanowitsch (2004); Stefanowitsch & Gries (2005).

b. *Bill kicked **at** the ball*

The conative construction presented in sentence (1.3b) consists in inserting the preposition *at* before the direct object. When it comes to its semantics, the conative construction "can be broadly described as a 'detransitivizing' construction" (Perek 2015: 90) that indicates some sort of a "directed action" (Perek 2015: 94), which means that the purported construction should have a very general meaning. The conative construction can indicate missed contact, lack of affectedness, or repetition, among others, which indicates that the construction's "semantic contribution … is highly variable, and, if anything, difficult to grasp with a single generalization." (Perek 2015: 94). If the meaning of the conative construction cannot be summarised with a single generalisation, we could pose the question whether one general conative construction as a whole exists in speakers' minds.

According to Perek, "the meaning of a construction can be largely traced back to its verbal distribution", that is "a construction comes to be associated with the meaning of verbs most frequently occurring in it." (2015: 11) To put it differently, the verbs that occur most frequently with the construction should match this construction's semantic meaning closely). Perek (2015) used a collostructional analysis – which measures the association between a construction and the verbs that instantiate it – to study the conative construction as a whole. The results of the analysis did not show a direct relationship between the construction and the verbal distribution that one might have expected. The verbs that were most strongly associated with the conative construction did not belong to any particular class – they did not share any specific type of context or meaning. What is more, some of the most frequent verbs associated with the construction belonged to two classes with conflicting meanings: allative verbs (e.g. *kick*) and ablative verbs (e.g. *pull*), which means that the conative construction would have to accommodate potentially incompatible contexts. Based on this result, Perek hypothesised that the conative construction (see Example 1.3b above) can be "better described as a set of verb-class-specific [i.e. lower-level – JJ] constructions" (2015: 114). Perek studied four separate lower-level verb-class-specific constructions that could be described as sub-constructions of the conative construction (INGESTION, CUTTING, PULLING, and HITTING) to see whether their verbal distribution corresponded with their constructional meanings. To simplify, Perek aimed to see whether the four lower-level constructions were more likely to appear with verbs that are compatible with those constructions, i.e. with verbs that share some meaning with the given construction. The results of the study showed that this was indeed the case (Perek 2015: 140). The verb-class-specific constructions were associated strongly with verbs that shared prominent aspects of meaning with the given construction, while repelling (i.e. occurring very rarely with) verbs whose semantics was at odds with the proposed constructional meaning (ibid.). In other words, in contrast to the conative construction as a whole, the verbal distribution of verb-class-specific constructions correlated with the semantics of those constructions. Perek takes this result as an indication that "the usage basis of argument structure constructions might be more visible at lower levels of generalization" (2015: 140). We can say that Perek's (2015) study of the conative construction provides evidence that lower-level verb-class-specific constructions exist, but does not provide such evidence for the conative construction as a whole. Similarly to Dąbrowska (2008c: 593), Perek does not see the results of the (2015) study as evidence that more general high-level constructions do not exist at all; however, lower-level constructions might be more aligned with real language usage and might more accurately reflect speakers' knowledge of language.

The studies presented in this section provide data which show that the different levels of generalisation must be systematically investigated and tested if we want to discover how language is stored in and produced by the human mind – any postulated constructions (especially the most general ones) must be subjected to the scrutiny of empirical data. Speakers of a language might not always use the categories that linguists postulate. One of the reasons that speakers might not build some general categories is that their frequency in the input they receive is too low for speakers to be able to formulate generalisations over the more specific categories. The next section will discuss the relationship between language constructions and their frequency of occurrence.

1.7. Constructions and frequency

Following the assumption that language is usage-based, the frequency of use of certain linguistic patterns should determine whether they can become conventionalised units of language and correlate with how entrenched those patterns are in the minds of language speakers. Let us take an extreme case as an example: a pattern used once by one speaker is unlikely to become conventionalised and thus used by the wide language community. On the other hand, patterns used often by large numbers of speakers could become conventionalised ways of expressing certain meanings or functions.

We can measure the frequency of a word or construction in many ways, but the two most important indicators in usage-based linguistics are **token frequency** and **type frequency**. Token frequency is a tally of the occurrences of **the same** linguistic form (a word or a phrase) in a given corpus, e.g. the word *cows* or the phrase *I kicked a ball*. Type frequency, on the other hand, counts "how many different lexical items a certain pattern or a construction is applicable to" (Bybee & Thompson 1997: 378). Using Bybee & Thompson's (1997) example, the English regular past tense construction marked by *-ed* will have a very high type frequency because it applies to a large number of different verbs, while the type frequency of the vowel change in words such as *h**a**ng-h**u**ng* will be markedly lower because this pattern applies only to a handful of verbs.

The two types of frequency defined above can have different impact on the productivity of a construction and the manner in which a construction is acquired. According to Bybee & Thompson (1997), the higher the type frequency of a construction, the more productive the construction is and the more likely to be extended to new items. If a pattern occurs with many **different** lexical items, speakers will be more likely to build a more general category, because the pattern will not be restricted to a few items. Consequently, "[t]he more items a category must cover, the more general will be its criterial features and the more likely it will be to extend to new items" (Bybee & Thompson 1997: 384). The meaning or a function of a linguistic pattern (a construction) must be general enough to apply to many lexical items and accommodate many potential contexts that will differ from one another. Lastly, high type frequency means that speakers will use a given construction more often, which will eventually "strengthen its representational schema, making it more accessible for further use, possibly with new items" (Bybee & Thompson 1997: 384). In sum, if a representation of a pattern is strong in the memory of the speakers of a language, they will be more predisposed to apply it to items with which they have not heard it yet, as long as the pattern is general enough to be used in many different contexts.

Token frequencies bear a somewhat different impact on the productivity and acquisition of constructions. Diachronically, tokens of very high frequency are resistant to change, e.g. the irregular English past tense forms. Even though the suffix *–ed* is the primary marker of past events in present-day English, a number of verbs form the past tense differently; for instance, the past tense form of *get* is *got*. *Get* is presumably a verb of very high frequency of occurrence and because of the frequency, it has retained its 'old' past tense marking – Bybee & Thompson (1997: 380–381) deemed such "conservative" behaviour of high-frequency tokens the Conserving Effect. Bybee (1985: 117) suggested that high-frequency forms tend to remain unchanged over different stages of a language's development due to their **lexical strength**. Each time a speaker of a language uses or hears a given token, it strengthens the token's representation – the token's "exemplar cluster" (Bybee 2010: 75) – in the speaker's memory. The stronger the representation grows, the easier it becomes for a speaker to access this particular form, and, thus, this form becomes more likely to be accessed than a potential more compositional, yet less entrenched, form such as **gotted*.

On the other hand, high-frequency tokens can also have a facilitatory effect on the acquisition of the "more compositional" constructions, that is constructions with high type frequencies such as the English Verb-Locative construction (*He went into the park*). In constructions with a Zipfian[5] distribution of tokens, one or more high-frequency exemplars (tokens) facilitate the acquisition of the construction by providing a prototype around which the construction will subsequently be built (Goldberg 2006: chap. 4). Additionally, in order to facilitate the acquisition, those high-frequency exemplars must carry a sufficiently general meaning for the

[5] A Zipfian distribution is a distribution in which a relatively small number of highly frequent words account for most tokens, while the rest of words occur with relatively low frequencies (Zipf 1935).

speakers to be able to generalise beyond them and build the construction (Ellis & Ferreira-Junior 2009: 379).

Ellis & Ferreira-Junior (2009) investigated the relation between the frequency distribution of verbs in constructions and the acquisition of those constructions by second language (L2) learners. The study involved three English argument structure constructions: verb locative (VL), verb object locative (VOL) and ditransitive (or verb object object; VOO). Some examples of the constructions are, respectively: *He went home*, *She put the book on the shelf*, and *They gave the parcel to the driver.* Ellis & Ferreira-Junior (2009) aimed to test whether the distribution of the input that L2 learners receive is Zipfian, and whether they first use the most frequent, general, and prototypical exemplars. The study consisted of two parts: a longitudinal analysis of a corpus of interviews with L2 learners, and a questionnaire in which native-English participants were asked to rate the prototypicality of the verbs used in each of the three constructions.

In the corpus analysis, Ellis & Ferreira -Junior (2009) studied the language of seven non-native English learners who were interviewed by native English speakers every four to six weeks over a period of 30 months. They calculated the frequencies of verbs used in each of the studied constructions – both by the interviewers and the interviewees. The corpus analysis showed that the distribution of each of the three constructions in the input produced by interviewers was Zipfian. The verbs in the language produced by interviewees also exhibited Zipfian distributions, albeit somewhat "amplified" ones – the most frequent verb used in the construction accounted for more uses of the construction than it was the case in interviewers' language. Moreover, the frequency of lemma use by an interviewee correlated strongly with that of the interviewer, which demonstrates that frequency distribution conditions construction learning. Lastly, Ellis & Ferreira-Junior found that one “pathbreaking verb ... seeds the construction and leads its development” (2009: 375), that is one verb is used as the first verb in a given construction, and later, it is also used much more frequently than other verbs. Such a verb must be semantically prototypical and also sufficiently generic to function as the prototype of an entire construction (Ellis & Ferreira-Junior 2009: 379); the ‘pathbreaking’ verbs for the constructions were *go* (VL), *put* (VOL), and *give* (VOO).

Ellis & Ferreira-Junior (2009) conducted a questionnaire among native English speakers to assess the prototypicality of the verbs used in the studied constructions, and, in particular, to verify whether the pathbreaking verbs are also the most prototypical verbs for the constructions. The respondents were asked to rate 80 verbs "for the degree they followed the meaning" of each of the constructions. In the results, the three pathbreaking verbs from the longitudinal study ranked very high on the scale of prototypicality, but only one of them (*give* for the VOO construction) reached the top of the classification. The highest-rating verbs for the constructions were: *walk, move,* and *run* (VL); *bring, move,* and *send* (VOL); and *send* and *give* (VOO). Ellis & Ferreira-Junior argue that *go* (for VL) and *put* (for VOL) were used the most frequently despite not being the most prototypical verbs for their constructions, because they are the most generic out of the verbs that appear in the constructions (2009: 379).

As a concluding remark, it must be added that the effects of frequency pertain to human experience in general, not only to language, and their existence is supported by psychological research – "[t]he more times we experience something, the stronger our memory for it, and the more fluently it is accessed" (Ellis 2012: 7).

1.8. Conclusions

In usage-based linguistics, language structures arise from usage – speakers build those structures by abstracting over many usage events. Language does not require any special innate capabilities (e.g. a genetic imprint for grammar), but uses human general cognitive capacities, such as categorisation. If we take these assumptions to be true, general linguistic constructions, such as the (light) reflexive marker or verbal prefixes would arise as a result of abstracting over many reflexive verbs or prefixed verbs. So far, the existence of those general constructions has mostly been taken for granted (with the few exceptions of Dąbrowska 2008a and 2008b, and Perek 2015, for example), and linguists tacitly assumed that if a general structure can be posited, speakers would necessarily possess those constructions. In other words, linguists' grammars would be equal to speakers' grammars. If language structures do arise from usage, however, linguists need to provide evidence supporting the existence of general constructions. The mere fact that a structure **can** be used for **describing** a language does not automatically imply that speakers use this structure in daily communication – a structure might be too general to be useful. Perhaps speakers use only the specific constructions for the concrete instances of the postulated structure and never attain the abstractions postulated by linguists. This thesis will explore the above issues on the basis of data on Polish reflexive verbs and prefixed verbs – the next chapter will introduce the two Polish language phenomena.

Chapter 2: Polish prefixed verbs and reflexive verbs

2.1. Introduction

Do speakers build the general constructions that linguists postulate? This question will be investigated in this thesis based on data on Polish reflexive verbs and prefixed verbs. The two types of Polish verbs appear to be a good testing ground for the linguists' categories vs speakers' categories. Categories of different levels of generality have been postulated for the reflexive markers and verbal prefixes – they have been described either as having one invariant meaning (i.e. as maximally general constructions) or as having a number of distinct senses (i.e. as constructions of a lesser degree of generality). Consequently, we can say that we have a range of constructions of different degrees of generality that speakers could potentially build and use in their daily linguistic interactions – and we can investigate empirically whether they could actually build them.

This chapter will begin with an overview of the different approaches to the study of Polish reflexive verbs and prefixed verbs respectively; subsequently, it will provide more details about the construction of the empirical studies. The section on reflexives will first review the multifunctional approaches to reflexive verbs, which propose that the reflexive marker *się* (an inherent part of reflexive verbs) has a number of different senses. Subsequently, the section will introduce the reader to the monofunctional accounts of the reflexive marker (with the greatest emphasis placed on usage-based accounts), which postulate that the 'light' reflexive marker *się* and the 'heavy' reflexive marker *siebie* each have one invariant meaning. The section on prefixed verbs will begin with an overview of the descriptive approaches to Polish verbal prefixes in general, including a list of the different meanings or senses that the prefixes can carry. Then, the section will review the usage-based approaches to Polish verbal prefixes. Finally, the chapter will conclude with a description of how reflexive verbs and prefixed verbs will be investigated in this thesis.

2.2. Reflexive verbs

The first case study investigating general vs specific constructions will be based on data about Polish reflexive verbs. Two markers have traditionally been regarded as reflexive in Polish linguistics: *się* and *siebie*.

(2.1)	a.	*Tata*	***się***	*ubiera.*
		dad	**REFL**	dress
		'Dad is getting dressed'		
	b.	*Uważa*	*na*	***siebie****!*
		watch	on	yourself$_{\text{ACC}}$
		'You'd better watch yourself!'		

The grammatical status of reflexive markers (and reflexive verbs) has been notoriously difficult to define, not only in Polish. Researchers have used operational definitions based on an array of different criteria: syntactic, semantic, and functional. The issue becomes even more complicated in languages with multiple candidates for the reflexive marker, such as Polish (Frajzyngier & Curl 2000: viii). Polish linguistic literature offers a number of different approaches to reflexive verbs. Most studies make the implicit assumption that the pronoun *siebie* is the 'true' reflexive marker (see e.g. Swan 2002: 159). The status of *się*, on the other hand, has always been problematic, as it defies traditional grammatical categorisations. So far, *się* has been defined as a syllable (Bogusławski 1977: 103), as a derivational morpheme (Wilczewska 1966: 19), an enclitic and 'defective' form of the pronoun *siebie*$_{\text{ACC}}$ (Nagórko 2007: 155). *Się* does not behave like a 'proper' pronoun – it cannot be part of prepositional phrases, nor can it be co-ordinated with other pronouns or nouns. Only *się*, constitutes an inherent part of verbs classified as **reflexive verbs** in Polish – *się* must be present along a verb in order for the verb to function as and be categorised as a reflexive verb. The position of *się* within a clause is also fairly constrained, as it can occur only in the nearest vicinity of the verb; only adverbials can intervene between *się* and the verb it is attached to. Many studies do not explicitly define or establish the status of *się* as a whole and instead present an extensive taxonomy of different functions of *się* (e.g. Kubiński 1982; Niedzielski 1976). Overall, regardless of their theoretical affiliations, linguists have always tried to search either for one overarching function of the marker – e.g. Wilczewska's (1966) derivational morpheme – or multiple “more specific” functions, e.g. Kubiński (1982), Niedzielski (1976), Tabakowska (2003a), but also Wilczewska (1966). In other words, *się* was accounted for in terms of a single maximally general construction (a monofunctional approach) or multiple constructions of intermediate degree of generality (a multifunctional approach).

To the best of my knowledge, it has not yet been investigated empirically whether speakers could arrive at the maximally general single-function categories for *się* or even the intermediately general multifunctional categories. This thesis will attempt to investigate whether *się* is an exponent of one general category or a number of less general intermediate categories. The third, extreme, option would be that *się* does not mark any categories of even moderate generality and should be regarded as an element of many low-level reflexive verbs, that is, each reflexive verb would constitute a separate category. The present section will give an overview of the different approaches to the study of *się* and Polish reflexive verbs – beginning with the multifunctional approaches and then following with monofunctional approaches – and put them in the context of general/less general/specific categories.

2.2.1. Multifunctional approaches to *się*

As was mentioned previously, some researchers treat *się* as a marker of multiple (albeit related) functional categories. In other words, some researchers postulate that *się* exhibits extensive polysemy – the marker has been hypothesised to have as many as six (Kubiński 1982), seven (Niedzielski 1976) or even seventeen (Wilczewska 1966) different functions. The different meanings of *się* proposed by Polish linguists seem to form three overarching groups in terms of their formal properties: reflexive-type, reciprocal, and impersonal. In the first group (reflexive-type), the verbs paired with *się* can occur with any type of subject noun (regardless of its gender, person or number) and can also be freely inflected. In its impersonal function, *się* “imposes” limits on the verb it pairs with: the verb can only appear in its singular neuter form; the clause containing a verb with impersonal *się* must not contain an overt subject – the clause must be subjectless. The three following sections will briefly introduce each group and provide examples from Polish.

2.2.1.1. Reflexive-type functions of *się*

The reflexive-type group is the most heterogeneous and possibly the most controversial group out of all three groups. Since *się* can occur with a vast number of verbs, many different functions of the marker have been proposed. There has been no consensus as to the exact number of functions, and often, verb + *się* combinations given as an example of one function in one study appear in a different group in another study, e.g. Niedzielski (1976: 178) categorised *bać się* 'be afraid' as passive, whereas Kubiński (1982: 57) put it in reflexiva tantum[6]. Yet another issue in the description and taxonomy of reflexive-type functions of *się* is that the different subgroups proposed within this group have been based on semantic distinctions rather than formal characteristics. In contrast, the impersonal *się* and the reciprocal *się* are classified as separate groups based on **both** their semantics **and** their formal properties. We can already see a discrepancy here that begs investigation – it would be interesting to test whether native speakers of Polish are sensitive both to the formal and the semantic distinctions or the formal or semantic ones exclusively.

2.2.1.1.1. True reflexive verbs

The most important category present in many studies are true reflexive verbs (Kubiński 1982) or direct reflexive verbs (Wilczewska 1966). The sentence in (2.2.) would be an example of this construction:

(2.2)	a.	*Janek*	*myje*	***się***
		Janek$_{\text{NOM}}$	wash	się
		'Janek is washing (himself)'		

(Kubiński 1982: 56)

Niedzielski (Niedzielski 1976: 171) suggests that in true reflexive verbs (as opposed to pseudo-reflexive verbs), *się* can be substituted with *siebie*, as in (2.2b):

[6] For more information about reflexiva tantum, see section 2.2.1.4.

(2.2) b. *Janek* *myje* ***siebie***

Janek$_{NOM}$ wash siebie

'Janek is washing himself'

(Kubiński 1982: 56)

Both (2.2a) and (2.2b) are indeed grammatical sentences in Polish, but they differ slightly in their meaning. Kubiński suggests that (2.2a) is "perceived by some native speakers as having larger integrity than" (2.2b) and in (2.2a) "the agentive function of the subject NP does not seem to be as stressed as in" (2.2b) (1982: 58). What it means is that some native speakers perceive the subject in (2.2b) as split to some extent. The impression might be that – even though they are the same person – the washer and the person being washed are separate to some extent. In Wilczewska's (1966: 29) view, apart from the difference in transitivity, directly reflexive verbs (i.e. true reflexive verbs) do not differ substantially in meaning from their non-reflexive counterparts. What would follow is that those reflexive verbs are directly derived from their non-reflexive counterparts.

Many scholars assume that this function is the closest to the "main" function of reflexives: expressing a situation in which the subject performs the action on itself. The subject is always the agent in the given situation, which naturally confines the group of possible subjects to humans or, sometimes, animate beings (Wilczewska 1966: 29).

2.2.1.1.2. Extensions of true reflexive verbs

If we define true reflexive verbs as reflexive verbs whose semantics does not differ significantly from the semantics of their non-reflexive counterparts, a group of verbs exists that could be regarded as an extension of true reflexive verbs. Those verbs would not permit the substitution of *się* with *siebie*, however, their semantics is still similar to that of their non-reflexive counterparts.

Extensions of reflexive verbs can perform three functions. Firstly, they can express actions whose recipient is not exactly the subject of the clause but some property that belongs to the subject (Wilczewska 1966: 35), e.g. *zapiąć się* ('fasten, zip, button') or *spakować się* ('pack'). The subjects of such verbs do not pack or fasten themselves – they button their shirts and pack their bags. Secondly, verbs in this group can express actions whose object is not the verb's subject as a whole but one of its parts (Wilczewska 1966: 37), e.g. *zmarszczyć się* ('frown', lit. 'crease oneself') or *skrzywić się* ('wince', lit. 'contort oneself'). The last function of reflexive verbs in this group is to select only one (or a few) possible referents out of the whole range of referents that could potentially serve as the object of its non-reflexive counterpart, usually when referring to emotions or mental states (Wilczewska 1966: 41). In other words, transitive verbs such as *opanować* ('control') can take a range of potential objects, however, their reflexive counterparts lexicalise only one object – the verb *opanować się* ('calm down, control oneself') refers only to situations in which a human subject controls his/her emotions.

All extensions of true reflexive verbs have one property in common: they exhibit metonymic qualities. The subject of a given verb + *się* construction grants access to one lexicalised object in its (i.e. the subject's) dominion – be it a property that belongs to the subject, a part of the subject, or only one referent out of many possible ones. What differentiates true reflexive verbs from extensions of reflexive verbs is that in the former, the subject performs the action on itself, while in the latter, the subject performs the action on a metonymically related entity. The difference seems to be entirely semantic, as there are no significant formal differences between the two groups.

2.2.1.1.3. Inchoative/resultative verbs

Inchoatives "express the inception [of] or a change in a process" (Niedzielski 1976: 178). The classic example of an inchoative reflexive meaning is:

(2.3.)	*Drzwi*	*otworzyły*	*się*
	door	open	*się*
	'The door opened'		

The sentence conveys only a change of state – the door was closed but now it is open. Nevertheless, one might also regard such sense of the verb *otworzyć się* as resultative, since the change of state must have been a result of some action or a process. Niedzielski (1976: 178) subsumed inchoatives under passive reflexive verbs, because the cause of the action/process/change is usually either unknown or unspecified in such situations. Wilczewska (1966) postulated a category of passive-resultative-inchoative verbs, which appears to be quite an accurate label, because demarcating clear boundaries between passive, resultative and inchoative verbs might often be very difficult. One might claim, for instance, that sentence (2.3.) should not be categorised as inchoative but as passive instead, because some external force must have opened the door – a door cannot just open by itself.

Inchoative/resultative verbs constitute a category that is further away from the extensions of true reflexive verbs. In contrast to the latter, inchoative/resultative verbs do not imply that the action or change of state was instigated or caused by the subject itself.

2.2.1.2. Reciprocals

The second main group of *się* functions is the reciprocal function. Some situations, e.g. kissing or fighting, typically involve two (or more) participants performing the same kind of action simultaneously on each other. In English, reciprocal situations of this sort are either expressed by a clause in which the verb does not take an overt object (*They kissed*) or by *each other* (*They argued with each other*). In Polish, the reciprocal meaning is conveyed by *się*, hence, the equivalents of both English examples would be expressed by a VERB + *się* construction:

(2.4)	a.	Marek	i	Ola	*pocałowa-l-i*	*się*
		Marek	and	Ola	kiss-PST-3PL.VIR	się
		'They kissed'				
	b.	*Pokłóci-l-i*	*się*			
		argue-PST-3PL.VIR	się			
		'They had a row'				

The verb with the reciprocal *się* always appears inflected in the plural number. In contrast to reflexive-type *się*, which typically involved one entity serving both as the agent and the patient of an action, in reciprocals, we usually have two (or more) entities. The two entities remain in a specific relation with each other: each entity is simultaneously an agent of an action and a patient of a similar action performed by the other entity (Wilczewska 1966: 88).

The mere occurrence of a verb in its plural form will not always entail the interpretation of *się* as reciprocal. For example, the verb *obudzić się* ('wake up') will rather encourage the reflexive-type reading:

(2.5) *Obudzi-l-i* *się*

wake.up-PST-3PL.VIR *się*

'They woke up'

Sentence (2.5) above would most likely be interpreted to mean that two people woke up spontaneously and simultaneously, not that they woke each other up. According to Frajzyngier, "[t]he reciprocal meaning is a result not only of the combination of plural participants and the reflexive marker but also of the inherent properties of the verb. If the verb allows a reciprocal situation, then the clause is interpreted as reciprocal" (2000: 181). If the reciprocal meaning is the result of also "the inherent properties of the verb", it means that the reciprocal *się* might actually be encoded as many concrete verb + *się* constructions instead of a relatively general RECIPROCAL *SIĘ* construction.

Other authors consider the reciprocal *się* as a construction of "wide [i.e. significant – JJ] productivity" (Wilczewska 1966: 91) or as a syntactically motivated construction (Niedzielski 1976: 185). If we try to interpret it in terms of specific/general constructions, Wilczewska (1966) and Niedzielski (1976) seem to consider the reciprocal *się* a general construction that exists above and beyond the concrete verb + *się* combinations. To the best of my knowledge, no empirical studies to date have investigated whether the reciprocal *się* exists as a general construction in the minds of native speakers of Polish or it only exists as part of many concrete verb + *się* constructions. This thesis will seek to shed light on this issue.

2.2.1.3. Impersonals

In Polish, if one wants to speak about a situation without mentioning the instigator or the cause, one can use *się*. The purpose of the impersonal *się* is to demote the agent so as to put the emphasis on the action, not the agent that performs it. The agent in the impersonal *się* is maximally diffuse and the agent could be anyone or anything (see example 2.6a). The impersonal *się* belongs to a larger 'family' of agent-demoting constructions, such as passives, middles, and setting-subject constructions (Słoń 2007: fn 9).

(2.6) a. *W weekend nie chodz-i się*
in weekend no walk-3SG.PRES się
do pracy
to work
'One does not go to work at weekends'

b. *W weekendy (oni) nie chodz-ą do pracy*
in weekends (they) no walk-3PL.PRES to work
'She does not go to work at weekends'

Impersonal constructions with *się* exhibit prominent formal characteristics that differentiate them from any other construction with *się*. Firstly, the verb always occurs in third person singular neuter. Secondly, the clause containing the verb and *się* in this function is always subjectless. That said, "the absence of the subject nominal in Polish is not sufficient to make a sentence impersonal" (Słoń 2007: 262) because "subjects are frequently left unelaborated in Polish" (Słoń 2007: 261), even in 'standard' active-voice SVO sentences. In clauses with the impersonal *się,* not only is the subject unelaborated, it cannot be elaborated at all – there is "no grammatical possibility" of doing that (Słoń 2007: 263). To put it differently, clauses with the impersonal *się* cannot contain a subject nominal at all – there is not even any potential of inserting a subject. If we compare the sentence with the impersonal *się* in (2.6a) to a 'standard' SV(O) sentence in (2.6b), we can see that a subject can be inserted in the latter sentence, as indicated by *oni* ('they') included in parentheses. The obligatory lack of subject in the clause with the impersonal *się,* most likely derives from its main purpose – the agent is demoted and maximally diffuse, so the clause does not contain any subject. The final formal characteristic of the impersonal *się* is that the impersonal *się* can be attached to any verb (Wilczewska 1966: 151), be it transitive or intransitive – even to verbs that would not accept *się* in the reflexive-type function. In comparison to the reflexive-type kinds of *się*, the impersonal *się* is part of a language pattern that is very distinct formally (verbs only in 3SG NEUTER, subjectless) and has a clear function: demote the agent. Due to the pattern's formal distinctness and its relatively uniform function, it is likely that native speakers of Polish store and use the impersonal *się* as a general construction. That notwithstanding, similarly to the reciprocal *się*, no empirical studies have been conducted to my best knowledge that would investigate in what form the impersonal *się* exists in native speakers' minds.

2.2.1.4. Reflexiva tantum

Apart from the three main groups of *się* functions, one more group exists, whose nature is quite special: *reflexiva tantum*. Reflexiva tantum, or deponents, constitute a very heterogeneous group, defined only by the fact that the verbs in this category do not have non-reflexive counterparts. Probably, the most obvious example of a *reflexivum tantum* verb is *bać się* 'be afraid, be scared'. We could easily separate the reflexive marker (*się*) and the verb, but *bać* without the marker does not exist, and the verb denoting 'scare' is *przestraszyć*. Sawicki (1988: 85) suggests that the reflexive marker in reflexiva tantum cannot be ascribed any function, **because** it does not supply contrast with any unmarked form. Consequently, the marker must have fused with the verb, and the verb and the marker should be considered one item. According to Tabakowska (2003a: 8), deponent verbs can be traced back to transitive verbs in the earlier diachronic stages of the development of the Polish language, even though their derivational composition has lost its transparency.

Wilczewska (1966: 114–115) objects against subsuming all verbs that do not have transitive counterparts under a single umbrella category of reflexiva tantum. Instead, Wilczewska (1966: 114–146) proposes three categories: non-transparent reflexiva tantum, deverbal reflexiva tantum, and denominal reflexiva tantum[7]. In the case of non-transparent reflexiva tantum, native speakers of contemporary Polish would not be able to trace any active etymological affinity to other words in the language (Wilczewska 1966: 115). This category cannot be analysed in terms of synchronic semantic derivation and must be studied with regards to its diachronic development – *bać się*, mentioned earlier, is a case in point.

Deverbal reflexiva tantum, another category singled out by Wilczewska (1966), derive from other verbs, as the name suggests. In contrast to "ordinary" reflexive verbs, however, the derivation does not proceed simply as verb + *się*. Deverbal reflexiva tantum usually come about as a result of two changes: prefixing a simple verb and adding the marker *się*. They often carry meanings of intensified or completed actions – the degree of expressed intensification depends on the prefix that a given verb contains (Wilczewska 1966: 121). Let us consider the verb *nawąchać się* ('smell/inhale something for a long time and saturate oneself with the smell'), which could be broken down into: *na*[8] + ***wąchać*** + *się.* A bare verb *nawąchać* does not exist, while *wąchać się* (a reflexive verb without the prefix) means 'smell each other' (or 'smell oneself'), and it would be difficult to say that *nawąchać się* was derived by means of a 'simple' prefixation from *wąchać się*. The of aspect smelling oneself is absent from *nawąchać się,* and if we wanted to derive *nawąchać się* from *wąchać się*, it would probably mean 'saturate oneself with the smell of oneself'. The meaning of the whole verb rests on both "additions" (*na* and *się*) and cannot be arrived at through stepwise derivation. As far as denominal reflexiva tantum are concerned, the category contains reflexive verbs derived from either nouns or adjectives. Similarly to deverbal reflexiva tantum, most denominal formations also require a prefix (though this does not always have to be the case). Additionally, because the derivation bases are not verbs, denominal reflexiva tantum must always receive a verbal suffix.

Non-transparent reflexiva tantum most likely exist as single units in the minds of native speakers, because their etymological and morphological transparency is

[7] The Polish names are, respectively: *reflexiwa tantum etymologicznie izolowane* (lit. 'etymologically isolated reflexiva tantum'), odczasownikowe *reflexiwa tantum*, and odimienne *reflexiwa tantum*.
[8] *Na* is a verbal prefix related to the preposition *na* 'on'.

very low (see Sawicki 1988; Tabakowska 2003a). Deverbal and denominal reflexiva tantum display some transparency, and it seems possible that speakers can draw analogies between forms such as *wąchać* ('smell') and *nawąchać się* ('smell/inhale something for a long time').

2.2.1.5. Multifunctional approaches: summary

In the multifunctional approaches, the functions of *się* described by researchers belong to three major groups: reflexive-type, reciprocal, and impersonal. These groups are distinguished based on different criteria: functional (semantic) and formal (syntactic). The first group (and its subgroups) has been established based on exclusively functional criteria. This group is the most heterogeneous of the them all and, when investigated closely, the reflexive-type *się* seems to express situations in which either the agent and the patient are the same entity, or when the patient is in a metonymic relation to the agent – it is either an inherent part of the agent or an object in its dominion. The lack of distinctive formal features as well as a vaguely defined function (or set of related functions) do not make the reflexive-type *się* a very likely candidate for a general construction that would be stored and used by native speakers of Polish.

The characteristic formal property of the second group, the reciprocal *się,* is that the verb is always inflected for the plural number. This property, however, is not a very strong cue, because verbs with the reflexive-type *się* can also appear in the plural number. The reciprocal *się* conveys a quite clearly delineated meaning – it marks situations in which entity A performs an action on entity B and, at the same time, entity B performs an identical action on entity A. Its formal properties and a clearly delineated function make the reciprocal *się* a better candidate for a construction that would be stored and used productively by native speakers.

The impersonal *się* has both a distinctive formal pattern and a clear and unique function. The sentences with the impersonal *się* do not contain an elaborated (or overt) subject, and the sentence would become unacceptable if a subject were inserted. The impersonal *się* demotes the agent, makes it maximally diffuse and puts the focus only on the action conveyed by the verb. These two properties of the impersonal *się* make it a very likely candidate for a general construction stored and used productively by native speakers.

Apart from the three groups, many researchers also mention a group of 'outliers' – reflexiva tantum. The verbs in this group do not have their non-*się* counterparts (although they might have had them in the past), and *się* does not appear to carry one coherent function. Reflexiva tantum are idiosyncratic, and native speakers most likely encode them only as specific (lexical) constructions.

2.2.2. Monofunctional approaches: reflexive-type markers and grammatical voice

Some authors have analysed reflexive-type markers in terms of grammatical voice and suggested that they should belong in the system of voices, along with the active and the passive voice. Thus, they ultimately postulated that reflexive-type markers carry high-level general meanings, being the exponents of grammatical voices.

Klemensiewicz (1946: 79) defined the reflexive voice as a device that codes situations in which an activity is performed on the entity that performs the activity (i.e. the agent also happens to be the patient). More recently, Nagórko (2007: 105–106) also postulated that reflexive verbs should be considered a voice. In the reflexive voice, according to Nagórko, the agent assumes the object slot of a clause, hence it serves as both the subject and the object, or, in other words, the subject and the object are co-referential. The reflexive voice is a construction with a limited scope, and it cannot apply to any type of agent – the agent must exhibit animacy, i.e. it must be animal or human (Nagórko 2007: 106). Beside the active, passive and reflexive voice, Nagórko recognised the middle voice. The middle would refer to situations such as the one presented in sentence (2.7):

(2.7) *Szkoła* *się* *buduje*

school$_{\text{NOM}}$ się build

'The school is being built' (lit. 'The school is building')

Nagórko (2007: 106) defined the middle voice as "total obliteration of the syntactic slot for the agent". In Polish, sentences of this type would be used in situations when the speaker wants to avoid any reference to the agent / instigator of an action.

Situations of the middle type seem to be happening "by themselves", without any assistance from external actors. In Nagórko's (2007) view, the middle voice has only one exponent: *się*, whereas the reflexive voice can be coded by two markers: *się* and *siebie.*

Both Nagórko (2007) and Klemensiewicz (1946) regard grammatical voices as a purely syntactic phenomenon. A syntactic transformation of a sentence, e.g. from the active to the passive voice, does not entail a change in meaning – "the [semantic] content of the clause in its new [alternative – JJ] form remains essentially the same, thus grammatical voice has no semantic value" (Nagórko 2007: 104).

In Cognitive Linguistics, Tabakowska (2003a) also suggested that reflexive-type markers in Polish code grammatical voice. Differently to Nagórko (2007) and Klemensiewicz (1946), however, Tabakowska perceives voices in semantic terms. Tabakowska, extending Kemmer's (1993) theory of the middle voice to Polish, suggested that the middle voice should be recognised as a much wider phenomenon in Polish. In Tabakowska's (2003a) account, *się* and *siebie* are exponents of two different voices: the middle and the reflexive voice respectively. Both voices code "real-world" coreference between the agent and the patient; however, they differ in the way they portray the participants of an event. In the reflexive voice, the roles of agent and patient are still distinguishable to some degree, whereas in the middle voice, both roles are conceived as a "single holistic entity" (Tabakowska 2003a: 15).

The following subsections will present the Cognitive Linguistic approach to voice and Polish reflexive verbs-type markers in more detail. First, a brief overview will be given of voice in Cognitive Linguistics. Subsequently, Kemmer's (1993) theory of the middle and reflexive voice will be introduced. The final subsection will present Tabakowska's (2003a) and Dancygier's (1997) Cognitive Linguistic accounts of Polish reflexive-type markers.

2.2.2.1. Grammatical voice: reflexives and middles

Before we proceed to analysing reflexive-type markers in terms of voice, let us define what grammatical voice is. Langacker (2004: 65) regards grammatical voices as “alternate ways of construing situations and presenting them linguistically”. The most prominent example of a voice difference is the opposition between the active and the passive voice:

(2.8)	a.	*Robotnicy*	*zbudowali*	*szkołę*		
		workers	built	school$_{ACC}$		
		‘Labourers built the school’				
	b.	*Szkoła*	*została*	*zbudowana*	*przez*	*robotników*
		school$_{NOM}$	became	built	by	workers$_{ACC}$
		‘The school was built by labourers’				

As we can see, the same real-world event is liable to be construed in various ways. In the active voice (2.8a), both the agent and the patient are present in the sentence. By virtue of being the subject of the sentence, the agent enjoys the greatest prominence in (2.8a). In contrast, a sentence in the passive voice (2.8b) presents the patient as the most prominent entity; the agent does not even need to be expressed. We can say that the primary function of the passive voice is that of defocusing the agent (Shibatani 1985: 830).

According to Langacker (1999: 24–25; 2008: 355–360), transitive active voice sentences are the grammatical correlate of the canonical transitive event, that is, transitive events are usually expressed by a transitive sentence in the active voice. In a canonical transitive event (Figure 2.1.), an agent performs an action on a patient, an interaction between the two entities (marked as an arrow) takes place and, as a result of the interaction, the patient changes its state (marked with a squiggly line).

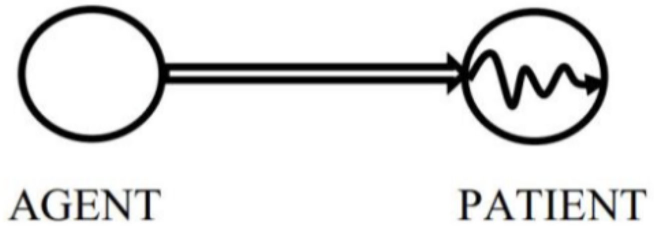


Figure 2.1. Canonical event model (based on Langacker 1991: 285; simplified)

Reflexive verbs diverge from the canonical event, where one entity interacts with another entity, and the state of the patient changes. In events of the reflexive type, an interaction also occurs but in this case, the entity interacts with itself– the event involves only one entity.

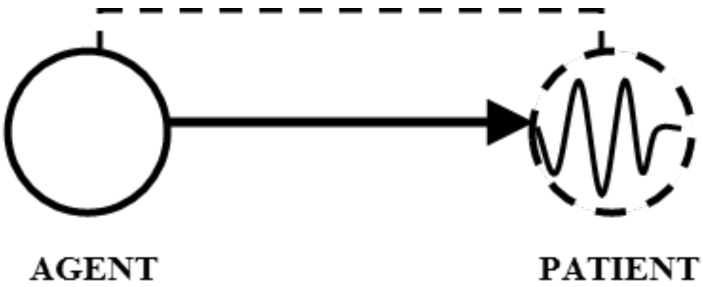


Figure 2.2. Reflexive event (based on Kemmer 1993: 71)

Figure 2.1. appears quite similar to Figure 2.2.; they both have two semantic roles: the agent and the patient, an interaction occurs between them, and the patient changes its state. In Figure 2.2., however, the agent and the patient happen to be the same entity (they are co-referential), which is indicated by the broken line connecting the agent and the patient.

(2.9) a. *Jacek* *uderzył* *kolegę*

Jacek$_{\text{NOM}}$ hit friend$_{\text{ACC}}$

'Jacek hit his friend'

b. *Jacek* *uderzy-ł* *siebie* *zamiast* piłki

Jacek$_{\text{NOM}}$ hit siebie$_{\text{ACC}}$ instead-of ball$_{\text{GEN}}$

'Jacek hit himself instead of hitting the ball'

The sentence in (2.9a) is an active transitive sentence – a realisation of the canonical event. A transfer of energy occurred between Jacek (the agent) and his friend (the patient) and, as a result, the friend's state changed – he probably got bruised. In contrast, the event in (2.9b) involves only one participant who happens to be both the agent and the patient in this interaction. Since this constitutes a special situation, it calls for a special type of construction: the reflexive.

In typically transitive actions, such as hitting someone or something, we normally expect the agent and the patient to be two different entities. Only very rarely does it happen that in such situations both roles are assumed by one entity. Nevertheless, there are also situations in which we expect the agent and the patient to be the same entity, such as in example (2.10).

(2.10)	*Tata*	*się*	*ogolił*
	dad_{NOM}	się	shaved
	'Dad shaved'		

One might argue that in normal circumstances people shave (and perform other grooming activities) without any external help. Interestingly, both English and Polish require a special type of marking in (2.10): in Polish, the marker *się* is used, whereas in English, the object is omitted altogether.

Kemmer (1993: 66) suggested that situations of the type exhibited in (2.10) should be considered middle, as opposed to 'true' reflexive verbs in (2.9b). Since we would usually expect people to perform those actions themselves (i.e. without any external help), we do not perceive the agent and the patient roles as distinct. Kemmer argues that in middle situations the participant fulfilling the role of agent and the patient is conceptualised as a "single holistic entity" (1993: 66). The reflexive, on the other hand, signals that the agent and the patient – which would otherwise be two different participants of a transitive event – happen to be the same entity. What follows is that "some separation of initiating and endpoint entities is maintained, despite the coreference of the participants" (Kemmer 1993: 66). Thus, the middle and the reflexive differ in **relative distinguishability of participants**, that is in how much mental separation there is between the participant as the initiating entity and the participant as the endpoint entity.

Types of events form a continuum whose items differ in the degree of distinguishability of participants (Figure 2.3.).

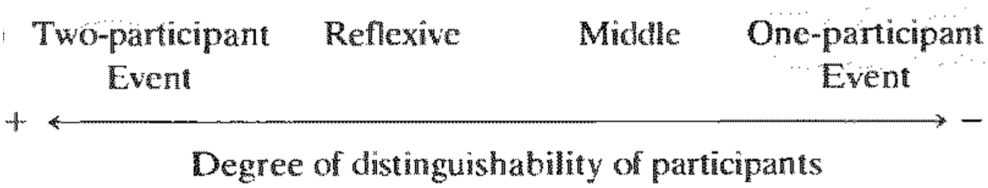


Figure 2.3. Kemmer's (1993: 73) event type continuum

The participants in two-participant events attain the highest distinguishability because they are essentially two separate entities. One-participant events are at the opposite end of the spectrum – the event involves only one participant, and the action does not have an object. Reflexive and middle events are intermediate between two-participant and one-participant events in terms of the distinguishability of their participants. The grammatical correlates of two-participant events and one-participant events are transitive and intransitive active clauses respectively.

Kemmer (1993: 19–22) also compiled a list of situation types commonly coded as middle cross-linguistically. The list in Table 2.1. presents those situation types; the list has been supplemented with examples from Polish (my own). For the sake of convenience and readability, the middle marker is printed in bold:

	SITUATION TYPE	POLISH EXAMPLE	TRANSLATION
1	non-translational motion	*ukłonić* ***się***	bow
2	translational motion	*czołgać* ***się***	crawl
3	change in body posture	*wyprostować* ***się***	straighten up
4	emotion middle	*cieszyć* ***się***	be happy
5	emotive speech actions	*skarżyć* ***się***	complain
6	other speech actions	*spowiadać* ***się***	confess
7	cognition middle	*zastanawiać* ***się***	think, ponder
8	spontaneous events	*skończyć* ***się***	run out, end

Table 2.1. Kemmer's (1993: 19–22) middle situation types

As we can see, the list contains verbs that could be said to belong to three major groups: motion (1-3), emotion/communication/cognition (4-7), and spontaneous events (8). The three groups are disparate when it comes to their semantic nature and the types of situations they express; the only thing they have in common is that they do not take direct objects (i.e. they are intransitive) and many of them have transitive counterparts. If the middle marker indeed has a single overarching function, and speakers store and use verb+MIDDLE MARKER as an established construction, they would need to abstract from all these disparate types of situations.

2.2.2.2. Middle and reflexive voice in Polish

Inspired by Kemmer's (1993) work on the middle voice, Tabakowska (2003a: 15) suggested that mere "real-world coreference" is insufficient for an accurate description of *się*. As we saw in Section 2.2.1., *się* can express situations in which there can be no coreference (i.e. the function characteristic for "true reflexive verbs"); it would be difficult, for instance, to say that inchoatives/resultatives (Section 2.2.1.1.3) indicate real-world coreference of participants. In such constructions, *się* cannot be exchanged for *siebie*, which motivated Tabakowska (2003a: 12) to suggest that *się* should be treated as the exponent of the middle voice and *siebie* as the exponent of the reflexive voice, in Kemmer's (1993) terms.

Tabakowska (2003a: 15–16) postulated that the middle in Polish, as opposed to the reflexive, is characterised by the following features: (a) low degree of distinguishability of participants in the event; (b) low degree of agent identifiability; (c) higher degree of the endpoint entity affectedness. The first characteristic of the middle voice comes directly from Kemmer's (1993) work and has been elaborated at length in the previous section. The two remaining characteristics deserve more explanation at this point. Low degree of agent identifiability is directly relevant to inchoative and resultative uses of *się* (see section 2.2.1.1.3.), in which the actual agent does not bear any importance in the conveyed situation or cannot even be identified at all:

(2.11)	*Konkurs*	***się***	*rozpoczął*
	competition	**się**	started
	'The competition started'		

In reality, the competition cannot commence by itself – it is always someone that needs to declare the competition open. In sentence (2.11) above, the agent that actually commenced the competition is irrelevant to the situation, and it would even be difficult to identify that agent. When it comes to the higher endpoint entity affectedness, Kemmer (1993: 50) described this feature using an example of the action of getting dressed. Getting dressed is often expressed with a middle voice marker, and in Polish, it is expressed with *się* (*ubierać się* 'get dressed'). When one is getting dressed "the affected parts of one's own body are not merely passively affected, as affected body parts are in an other-directed action; they actually in most cases participate to some degree in the action" Kemmer (1993: 50), which, consequently, makes them more involved and, thus, more affected. The more affected the endpoint entity is in a given situation, the more likely it is "to be viewed as one that cannot be directed towards others" Kemmer (1993: 51), and the more likely for the agent/patient to be conceptualised as a "single holistic entity" (Kemmer 1993: 66), which is one of the defining characteristics of the middle voice.

Dancygier (1997) also analysed *się* and *siebie* from a cognitive-linguistic standpoint. In contrast to Tabakowska (2003a), Dancygier did not explicitly analyse *się* in terms of the middle voice, however, the two studies seem to converge on most issues. Dancygier (1997: 325) saw *siebie* as the exponent of the true reflexive, "representing two distinct semantic roles as filled by one entity". *Się*, in contrast, carries a role-neutralising function: "it reduces the number of expressed participant roles, by eliminating syntactic expression of some of those roles in central syntactic slots, and/or by diminishing the distinction between roles held by the central participant" (Dancygier 1997: 325). The differences between *się* and *siebie* can essentially be interpreted in terms of relative distinguishability of participants (see Kemmer 1993; Tabakowska 2003a). In *się*, the agent and patient roles become neutralised, hence, the distinguishability of participants is very low. This corresponds to the situation from sentence (2.10), where a person shaved – people usually perform this action by themselves and on themselves, and there is hardly any conceptual distinction between the agent and the patient. When it comes to *siebie,* in contrast, the distinguishability between the participants is high, because these participants represent "two distinct semantic roles as filled by one entity" (Dancygier 1997: 325) – as in the situation when someone hit themselves instead of a ball (see sentence 2.8b).

Się can indicate reduced "participant responsibility and volitional involvement" (Dancygier 1997: 314), because with the participant roles neutralised, there is no distinct role of agent, and only an agent can act volitionally and be responsible for an action. Let us consider some of Dancygier's examples:

(2.12)	a.	*Uderzyłam*	*się*	*o*	*róg*	*stołu*
		hit	się	about	corner	table

‘I hit myself on the corner of the table’

b.	*Wzięłam*	*magiczną*	*różdżkę*	*i*	uderzyłam
	took	magic	wand	and	hit

nią	*siebie,*	*a*	*potem*	*dzieci*
she$_{\text{INS}}$	siebie$_{\text{ACC}}$	and	then	children

‘I took the magic wand and I hit myself and then my children’

(Dancygier 1997: 315)

The event in (2.12a) happened entirely unintentionally, without any control or volition involved. Consequently, *się* in (2.12a) cannot be substituted with *siebie*. On the other hand, in (2.12b) the subject consciously took a magic wand and hit herself and then her children, hence, *siebie* is not only perfectly acceptable but even the only grammatical option (Dancygier 1997: 315).

2.2.2.3. Monofunctional approaches: summing up

Monofunctional approaches either see *się* as a 'degenerate' version of the reflexive pronoun in the accusative (*siebie*) or as the exponent of the middle voice. In the former case, *się* would essentially fulfil the same function as *siebie*: it would indicate the co-reference of participants, albeit only in the accusative case. Essentially, the two items – *się* and *siebie* would be exponents of the same construction. In the latter approach, *się* (being the exponent of the middle voice) also encodes the co-reference of participants, however, the events it is used for are events in which the distinction between the agent and the patient role is blurred, i.e. the distinguishability of participants is low. *Się* would then apply to situations such as shaving, etc., where we expect only one participant by default; while *siebie* would encode situations in which the co-reference of participants occurs as a result of a coincidence and is by no means the default configuration.

Monofunctional approaches postulate very general constructions. The difference between *się* and *siebie* (in Tabakowska's (2003a) approach) rests in the degree of conceptual separation of the same real-world entity co-referenced in the roles of both agent and patient. In order to arrive at a construction of such a general meaning, speakers would have to abstract from many events of the many different situation types that the middle voice is postulated to apply to. Additionally, the difference in meaning between the reflexive marker and the middle marker appears to be very subtle, and speakers of Polish would probably need many different contexts to reliably establish constructions for both markers. No empirical evidence has been gathered so far as to whether native speakers of Polish develop the general constructions with *się* and *siebie* – this thesis will seek to bridge this gap.

2.2.3. Reflexive-type markers: interim conclusions

We can distinguish two primary approaches to reflexive-type verbs: the multifunctional approach (Wilczewska 1966; Niedzielski 1976; Kubiński 1982) and the monofunctional approach (Dancygier 1997; Tabakowska 2003a). The difference between the former and the latter boils down to the degree of generality in the categories they postulate. The monofunctional approach aims at explaining reflexive verbs in terms of maximally general categories, while the aim of the multifunctional approach is to develop functionally coherent categories of lower generality. Overall, despite the differences, the studies in both groups postulate categories of the degree of generality higher than what we would often call lexical in linguistics, i.e. specific categories. In both types of approach, native speakers would need to abstract from many instances of different reflexive-type verbs to arrive at categories of more general nature, such as the middle voice. Specific categories are recognised only in the case of deponent verbs, or *reflexiva tantum*, which cannot be broken down into the base non-reflexive verb and the reflexive marker – they constitute a kind of a fallback category for all reflexive verbs that could not be explained with higher-level categories. Linguists have so far tacitly presupposed that if higher-level categories for reflexive-verbs can be postulated, native speakers of Polish possess them. There exists a possibility, however, that native speakers do not arrive at those general categories and, consequently, that reflexive verbs (or at least some of them) are not instances of any higher-level categories (e.g. instances of the middle voice). None of the studies described in the present section investigated whether the general constructions they postulated would hold when compared with empirical usage data – they all had an exclusively introspective nature. This thesis seeks to bridge the gap between the theory and empirical evidence by using experimental and corpus data to explore the empirical validity of the general categories for reflexives.

2.3. Prefixes

Polish verbal prefixes will be another linguistic phenomenon studied in this thesis in order to investigate whether speakers can build the general constructions proposed by linguists. Just like in the case of Polish reflexives, no consensus has been reached as for that status of Polish verbal prefixes. The prefixes have been described in terms of constructions of varying generality – similarly to Polish reflexive-type markers. Linguists working on the Polish language have attempted at "distilling" a single overarching general function for all prefixes as well as accounting for the different functions of different prefixes. Whether native speakers of Polish store and use these constructions for prefixes (either the maximally general or the slightly more concrete ones) has not yet been investigated empirically to my best knowledge. This thesis will seek to fill this gap.

Let us start with a brief overview of what Polish verbal prefixes and prefixed verbs are. Polish prefixed verbs are verbs derived from other verbs by means of adding a prefix, e.g. *jechać* 'go$_{\mathrm{IMPF}}$, drive$_{\mathrm{IMPF}}$' –> *po* + *jechać* = ***po**jechać* 'go, drive$_{\mathrm{PFV}}$'. The dominant 'function' of all prefixes often postulated in the literature is that of a perfectivising device – when a prefix is added to an imperfective verb, the verb becomes perfective (see e.g. Dąbrowska 1997: 467; Christensen 2011: 1; Śmiech 1986: 7). Linguists have postulated sixteen verbal prefixes for Polish: *do-, na-, nad(e)-, o(b)-, od-, po-, pod-, prze-, przy-, roz-, u-, w-, wy-, wz-, z-/s-/ś-* and *za-*. Diachronically, most prefixes grammaticalised from prepositions, or "adnominal elements, which then later turned into lexemes defined as prepositions" (Tabakowska 2003b: 157). In other words, Polish verbal prefixes and Polish spatial prepositions are related diachronically, and some authors (e.g. Dąbrowska 1997; Tabakowska 2003b) suggest that prefixes and their corresponding prepositions still display some semantic affinity. The prefixes, apart from their perfectivising function, carry more "specific" meanings – they usually alter the spatial or temporal characteristics conveyed by the base verb. For example, when we supplement the verb *budować* 'build$_{\mathrm{IMPF}}$' with the prefix *do-*, the verb ***do**budować* will mean 'add a new part to a building$_{\mathrm{PFV}}$' as in the following sentence:

(2.13)	*Uniwersytet*	*dobudował*	*nowe*	*skrzydło*	*do*	*budynku.*
	University	built-to	new	wing	to	building.

‘The university added a new wing to the building’

Each Polish verbal prefix tends to have a few functions, but, while “the semantic contribution of the prefix may often be understandable ex post facto, it is not possible to provide rules for predicting the semantic result of adding a given prefix to a given verb” (Swan 2002: 281). What it means is that even though some regularities exist as to how the prefixes are used with verbs, the ultimate meaning of a given prefixed verb might be lexicalised, or, in other words, its meaning might be idiosyncratic. This property of Polish prefixed verbs makes them a very suitable source of data for the analysis of specific vs general constructions. It can be investigated empirically whether speakers of Polish can arrive at the more abstract generalisations proposed by linguists and use them in the comprehension and production of prefixed verbs.

2.3.1. Describing prefixes

The research on prefixed verbs in Polish linguistics has proceeded in two directions: descriptive and structuralist. In descriptive linguistics, researchers have attempted to find and describe the different meanings prefixed verbs can assume. Since covering all prefixes in a single monograph would be a formidable endeavour, most studies focused either on a single prefix (Sokolova & Lewandowski 2010; Tabakowska 2003b; Christensen 2011), a single function (Krupianka 1969), or a specific type of verbs, usually motion verbs (Giermak-Zielińska 1979; Striekałowa 1962). The only comprehensive descriptive study of all Polish prefixed verbs so far is Śmiech’s (1986) monograph. Prefixed verbs have also been given ample space in grammars of Polish – both those aimed at Polish readers (see e.g. Nagórko 2007) as well as grammars written in other languages and aimed at foreign language students and scholars (see e.g. Swan 2002 or Sadowska 2012). Different researchers have postulated different meanings of prefixes. It is beyond the scope of this study to compare all accounts, and only Swan’s (2002) classification will be presented here as an illustration, since it is the most comprehensive. Table 2.2. presents the different meanings of three chosen prefixes, but Swan (2002) listed the postulated meanings for all sixteen prefixes.

PREFIX	MEANING	EXAMPLES	
		POLISH	TRANSLATION
po-	for a while	*poczekać*	wait for a while
	a bit	*posłuchać*	listen a bit
	a single execution	*pokryć*	cover$_{\text{PFV}}$
	enter a state	*poczuć*	come to feel
przy-	arrival at destination	*przyjść*	arrive
	levelling	*przygnieść*	press down, flatten
	attachment	*przykręcić*	screw down
	somewhat	*przypalić*	burn slightly
roz-	in various directions	*rozrzucić*	scatter around
	destruction	*rozgnieść*	turn to mush
	expansion	*rozszerzyć*	expand
	dissipation	*rozładować*	discharge
	into bits	*rozerwać*	rip to shreds
	throughout or all-encompassing	*rozgniewać*	enrage
	undo	*rozkręcić*	unscrew

Table 2.2. Meanings of the prefixes *po-, przy-,* and *roz-*; based on Swan (2002: 283–284)

As we can see, the number of different meanings for each prefix in Table 2.2. is large (from four in *po-* to seven in *roz-*); the number of meanings for other prefixes ranges from two to nine. The very number of prefixes and their meanings gives an idea of how intricate and rich the postulated system is. Let us assume that the prefix *przy-* is a general construction, and the speaker, in order to produce the verb *przypalić* 'burn (slightly)$_{\text{PFV}}$, ignite$_{\text{PFV}}$' would need to compose the verb out of *przy-* and *palić*$_{\text{IMPF}}$ It seems that if a speaker were to face the choice of seven possible meanings of the prefix *przy-* and at least three meanings of the verb *palić* (if dictionary listings are any indication) when attempting to produce or process the verb *przy+palić –> przypalić*, the processing or producing of the verb would take a lot of cognitive resources due to the sheer number of different possible combinations. This entails a question: do speakers actually arrive at those meanings and use these constructions in their daily communication?

2.3.2. Prefixes: grammar or lexis?

The other influential avenue in research on Polish prefixed verbs, apart from the descriptive approaches, is the issue of whether prefixes belong to grammar (syntax) or lexis (word formation). Krupianka (1969) suggests that prefixes are mainly word-formation devices, but some of them also perform a syntactic function – they modify the verb's case government or alter its transitivity. For instance, *biec* 'run$_{\mathrm{IMPF}}$' is intransitive, whereas its *o-* derivate, *obiec* 'run around$_{\mathrm{PFV}}$' requires the accusative:

(2.14) a. *Biegł szybko.*

ran$_{\mathrm{IMPF}}$ fast

'He was running fast'

b. *Obiegł budynek.*

ran-around$_{\mathrm{PFV}}$ building

'He ran around the building'

Krupianka (1969), however, could not decide whether all prefixes as a category governed case or transitivity, or only some prefixed verbs were capable of performing this function. If we were to translate this into usage-based terminology, we could say that the former option would correspond to general constructions, whereas the latter would be roughly equivalent to specific constructions.

Giermak-Zielińska (1979), similarly to Krupianka (1969), saw prefixes as belonging to the domain of semantics rather than syntax, however, Giermak-Zielińska suggested that prefixes are lexical elements bordering on the domain of grammar. In other words, In Giermak-Zielińska's (1979) account, prefixes as a category are not as independent as nouns or verbs, because they cannot exist without the verbs they attach to. As we can see, despite postulating a lexical nature of prefixes, Giermak-Zielińska (1979: 11) still suggested that prefixes are coherent categories – they are regular paradigms, which **speakers take advantage of** when interpreting existing prefixed verbs or coining new ones. If speakers can take advantage of such paradigms, or constructions, it means that they must be cognitively real(istic). Giermak-Zielińska (1979) takes the cognitive realism of such categories for granted and does not provide any evidence supporting this account. Besides suggesting that prefixes form a separate category, Giermak-Zielińska recognised that prefixed verbs contain a “surplus of meaning” in comparison to their unprefixed counterparts, and the surplus is not directly related to the meaning of the prefix itself (1979: 11). Again, in terms of usage-based linguistics, we could say that the “surplus” of idiosyncratic meaning might rest in lower-level constructions.

Prefixed verbs of motion in Polish diverge from the general trend in which prefixes transform an imperfective verb into a perfective one. In motion verbs, a prefixed derivate from an unprefixed imperfective motion verb does not always have to be perfective, e.g. *pływać* ‘swim$_{IMPF}$, sail$_{IMPF}$’ – ***do****pływać* ‘approach swimming$_{IMPF}$’. This feature of motion verbs is related to their very special property: Polish verbs of motion often have two imperfective versions Striekałowa (1962). The Polish verb ‘swim’, for instance, can either assume the form of *pływać* or *płynąć*. Even though both are imperfective, the verbs differ in the meaning they convey – the former refers to durative situations, while the latter is iterative:

(2.15)	a.	*Płynął*	*z*	*prądem.*
		swam$_{IMPF}$	with	current
		‘He swam with the current’		
	b.	*Pływał*	*co*	*rano.*
		swam$_{IMPF}$	what	morning.
		‘He would go swimming every morning’		

In relation to this phenomenon, Striekałowa (1962) encountered another difficulty in her study on Polish prefixed verbs of motion. A structural linguist, Striekałowa attempted to find the exact route of derivation for Polish prefixed motion verbs. Some verbs, such as ***po****wstawać* ‘emerge’, turned out to be problematic. The verb *powstawać* is an imperfective verb that could be related to either another imperfective verb *wstawać* ‘rise, get up$_{IMPF}$’ or a perfective verb *wstać* ‘rise, get up$_{PFV}$’. According to Striekałowa (1962), *powstawać* could have been derived in two ways: (1) by prefixing *wstać* to obtain ***po****wstać* ‘emerge$_{PFV}$, rise$_{PFV}$’ and then infixing *powstać* to obtain *powst**aw**ać* ‘emerge$_{IMPF}$’; or (2) by prefixing *wstawać* ‘rise$_{IMPF}$, get up$_{IMPF}$’ to obtain *powst**aw**ać* ‘emerge$_{IMPF}$’. Striekałowa could not decide which derivational route was more plausible. If a speaker were to derive the meaning of the prefixed verb each time when processing or producing language, she would be facing the same dilemma. Speaking in terms of specific vs more general constructions, the issue described in Striekałowa (1962) gives reasons to suspect that, at least in the case of *powstawać*, it would be necessary for a speaker to have the meaning of the verb stored for quick retrieval in order to avoid unnecessary pressure on cognitive resources.

2.3.3. Prefixal networks

In usage-based linguistics, no comprehensive study of Polish prefixes as a whole has been published yet to my best knowledge. That notwithstanding, a number of studies have appeared that investigated some selected prefixes, e.g. Dąbrowska (1997), Tabakowska (2003b), Lewandowski (2014; 2016), Sokolova & Lewandowski (2010). The two major trends emerging from those studies were either to look for cognitive motivations for particular meanings of chosen prefixes (Dąbrowska 1997; Tabakowska 2003b) or to investigate how prefixes relate to other constructions, e.g. the locative alternation (Lewandowski 2014; Lewandowski 2016; Sokolova & Lewandowski 2010).

Dąbrowska (1997) aimed to elucidate how five Polish prefixes – *prze-, do-, od-, po-* and *za-* change the meaning of the verb they attach to, and how the properties of the object delimit the action conveyed by the verb. Using the apparatus of Langacker’s (1987; 1990; 1991 and later) Cognitive Grammar, Dąbrowska (1997) sought the underlying principles behind the semantics and functions of Polish verbal prefixes, emphasising the link between the prefixes and the corresponding spatial prepositions they are related to. According to Dąbrowska (1997), the differences between different prefixes, and also the different meanings of one prefix, boil down to different configurations of **trajector** and **landmark**. For example, one meaning of the prefix *prze-* involves a trajector moving “from one edge of a band-shaped landmark to the other” (Dąbrowska 1997: 469), as in the following sentence:

(2.16) Chłopiec przepłynął rzekę

boy swam-across$_{PFV}$ river

'The boy swam across the river'

The other sense of *prze-* expresses a trajector crossing a boundary, that is a landmark whose "width … equals zero" (Dąbrowska 1997: 472), as in example (2.17):

(2.17) Przemytnik przekroczył granicę

smuggler stepped-over$_{PFV}$ border

'The smuggler crossed the border'

The difference between the meanings lies in the fact that in (2.16), the landmark is a band-shaped object, whereas in (2.17), the landmark is a boundary that does not physically exist.

Prefixes and verbs which they can attach to exhibit limited compatibility. Not every prefix can attach to every verb, because some trajector-landmark configurations of prefixes and verbs can be at odds with each other. If they are compatible, however, "each prefix reinforces a different aspect of the meaning of the simplex verb, and thus modifies its meaning in different ways" (Dąbrowska 1997: 483).

Dąbrowska's (1997) study establishes a network of different senses of prefixes and links both spatial and temporal meanings of prefixes to their prepositional counterparts. This sort of approach aims at achieving maximum generality and unifying seemingly unrelated phenomena. Dąbrowska indicated that "it is beyond doubt that many [prefix-verb] combinations are … lexicalised" and suggested, in opposition to Śmiech (1986), that "this does not mean that they are arbitrary" (1997: 479), because the motivations for different lexicalisations derive from the underlying principles of prefixal meanings.

Tabakowska (2003), similarly to Dąbrowska (1997), sought to establish the general principles of verbal prefixation in Polish by comparing verbal prefixes to their prepositional counterparts. Tabakowska (2003) developed a Lakoff-style (see e.g. Lakoff 1987 and Section 1.2.2. for more information) radial semantic network for one prefix: *za-* in an attempt at creating a classification that on the one hand would be granular enough so as to accommodate all prefixed verbs and, at the same time, sufficiently general so as to prevent positing small subclasses that would contain only a few verbs.

Overall, Tabakowska's network established the systemic relations between the different meanings of *za-* and the meanings of the preposition *za-*, which had been previously treated as unrelated. This does not entail, however, that Tabakowska's descriptions departed entirely from previous work on Polish prepositions. On the contrary, Tabakowska suggested that the network for *za* offers "underlying unifying principles to old well-grounded insights, thus uniting individual fragmentary descriptions within a coherent overall model" (2003b: 157–158).

Both Dąbrowska's (1997) and Tabakowska's (2003b) studies are purely theoretical attempts, and no claims are made as to the existence of the postulated constructions in the human mind. Tabakowska (2003b) remained agnostic as to whether the derivational processes take place on-line in the speaker's minds or whether they are a feature of the linguistic system of Polish. That notwithstanding, Tabakowska (2003b: 174–175) recognised that some prefix+verb combinations have lexicalised and thus are "no longer felt (by present-day speakers of Polish) to include a prefix", which indicates that lower-level constructions might exist at least for some verbs. Tabakowska (2003b) treated lower-level constructions rather as an exception to a rule (i.e. general construction) than a norm and did not explore the issue of the cognitive reality of the postulated construction.

2.3.4. Prefixes vs the locative alternation

Lewandowski (2016) and Sokolova & Lewandowski (2010) explored the relationship between prefixed verbs and the locative alternation. In a nutshell, the locative alternation stands for the two possible ways of expressing locative relations:

(2.18) a. Jack sprayed the wall with paint. **(Goal-Object construction)**

b. Jack sprayed paint onto the wall. **(Theme-Object construction)**

In (2.18a), the theme (*paint*) is the direct object, whereas in (2.18b) it is the goal (*the wall*). The locative alternation exists also in Polish:

(2.19) a. Jan załadował ciężarówkę złotem.

Jan *za*-loaded lorry$_{ACC}$ gold$_{INS}$

'Jan loaded the lorry with gold'

b. Jan załadował złoto na ciężarówk-ę.

Jan *za*-loaded gold$_{ACC}$ on lorry$_{ACC}$

'Jan loaded the gold onto the lorry'

In the above sentences, (2.19a) is an example of a Goal-Object construction, whereas (2.19b) is the Theme-Object construction[9].

Sokolova & Lewandowski (2010) built constructional profiles of the Russian prefix *за-* [za-] and the Polish prefix *za-*. Both prefixes have been suggested to carry similar meanings: COVER/FILL, PLACE or REACH A NATURAL ENDING, and Sokolova & Lewandowski (2010) aimed to investigate the differences and similarities between the prefixes and to establish which version of the locative construction each one prefers. The verbs they studied were divided into three broad categories: manner verbs, path verbs and hybrid verbs. The first two categories have been widely recognised in usage-based linguistics (see e.g. Talmy 1985; Talmy 1991; Talmy 2000; Slobin 2004): path verbs express detailed information regarding the trajectory of motion, e.g. *enter*, whereas manner verbs carry detailed information about the specifics of motion, while leaving the path underspecified, e.g. *hobble*. The results of Sokolova & Lewandowski's (2010) study suggested that different semantic categories of verbs prefer different locative constructions[10]. Manner verbs go mainly with the Goal-Object construction, but hybrid verbs and path verbs do not display a strong preference for either construction, and Sokolova & Lewandowski (2010) proposed that "[t]he proportion between the Theme-Object and the Goal-Object constructions for 'hybrid' verbs depends on the idiosyncratic

[9] For a usage-based linguistics analysis of the locative alternation see Iwata (2008).
[10] For a similar comparative study on the locative constructions in Polish and Spanish see Lewandowski (2014).

properties of individual verbs" (2010: 379). This, in turn, suggests that prefixed verbs, at least those with the prefix *za-*, carry detailed meaning that goes beyond the sum of prefix + verb, and that speakers might need to rely on lower-level construction even for the choice of locative constructions.

Lewandowski (2016) studied the interactions between verbs, prefixes and argument structure constructions in Polish. The study entertained the hypothesis that "it is not the case that prefixation gives rise to a change-of-state verb with a different argument structure, but rather that it is the verb that is brought into conformity with the abstract change-of-state construction headed by a resultative prefix" (Lewandowski 2016: 178). In other words, verbs do not specify the argument structure themselves but rather combine with the one they are the most compatible with. When paired with a resultative prefix, the given verb can be coerced into another argument structure construction as long as this conforms to the Semantic Coherence Principle (see Goldberg 2006: 39–40), e.g. the verb *pryskać* 'spray' combines with the Goal-Object construction, whereas its prefixed version, *spryskać* 'spray' goes with the Theme-Object construction. Lewandowski's (2016) study proposes an intricate system of interactions between prefixal meanings, concrete verbs and abstract argument structure constructions. It appears to presuppose tacitly that speakers arrive at those very abstract generalisations, and they use them in the processing and the production of language.

2.3.5. Prefixes: interim conclusions

Just like in the case of reflexive verbs, the different approaches to prefixed verbs in Polish essentially differ with respect to the amount of generality they postulate. The levels of generality can come under various guises, depending on the theoretical affiliation of a particular researcher – in studies based in structuralist linguistics, it would be the difference between belonging to the domain of lexis or to the domain of syntax (Krupianka 1969; Giermak-Zielińska 1979). Usage-based linguistics does not postulate separate modules for different aspects of language – instead, it postulates a continuum of constructions from the most specific to fully general. When described in terms of usage-based linguistics, lexis would correspond to low-level specific constructions, while syntax would be equivalent to high-level general constructions. More descriptive accounts (e.g. Swan 2002; Śmiech 1986) aimed at providing unified descriptions of all possible meanings that prefixes can carry. If each of those meanings were to be a construction in native speakers' minds, they would require speakers to abstract from many different verbs with the given prefix in a given meaning. Tabakowska's (2003b) and Dąbrowska's (1997) studies on prefixes took a more usage-based approach to the matter, but, similarly to their descriptive counterparts, they also aimed at establishing different, general, meanings for prefixes. The difference between these studies and the descriptive studies rests in the fact that Tabakowska (2003b) and Dąbrowska (1997) attempted to account for the different meanings of prefixes in terms of general human cognitive abilities.

Some authors, even those who postulate general categories for prefixes, say that some prefixed verbs are "no longer felt (by present-day speakers of Polish) to include a prefix" (Tabakowska 2003b: 174–175) or that prefixed verbs contain a "surplus of meaning" (Giermak-Zielińska 1979: 11) which cannot be a result of combining an unprefixed verb and a prefix. In other words, some prefixed verbs are very likely to constitute strong lower-level categories. If some prefixed verbs are "no longer felt ... to include a prefix", perhaps, by extension, (at least some) prefixes do not form general categories at all, and speakers rely exclusively on lower-level constructions when using prefixed verbs? None of the studies described in this section provided real usage data that would indicate that general constructions for prefixes exist. This thesis will attempt to rectify this issue and explore the question empirically.

2.4. Composition of study and converging evidence

Polish reflexive verbs and prefixed verbs will be studied in this thesis, and a mixed corpus and experimental approach will be adopted. Both types of verbs will be studied in a similar manner. First, behavioural profiles – a corpus technique – will be used to investigate whether the contexts that reflexive verbs and prefixed verbs occur in could enable speakers of Polish to build general constructions for the reflexive markers and verbal prefixes *po-*, *przy-*, and *roz-*. As the second step, experiments will be conducted to investigate whether speakers of Polish could build and use constructions for the different senses of the 'light' reflexive marker *się* and prefixes *po-*, *przy-*, and *roz-* postulated in the relevant literature. One experiment (a sentence-sorting study) will be conducted for *się*, and two experiments (a sentence-sorting and a nonce-verb task) will be conducted for the prefixes.

Why has a combined corpus and experimental approach been selected for the studies discussed in this thesis? The overarching aim of such a composition was to gather converging evidence for the issue studied. The search for converging evidence has been long recognised in usage-based linguistics (see e.g. Lakoff & Johnson 1999: 79–81) and forms one of the methodological foundations of the field. Essentially, to look for converging evidence is to look for evidence coming from different sources and methods. It is assumed that "the skewing effects of any one method will be canceled out by the other methods" (Lakoff & Johnson 1999: 80).

In terms of "cancelling the skewing effects", experimental methods can balance out the deficiencies of corpus research and vice versa. Modern linguistic corpora afford access to previously inconceivable amounts of data, which enables the researcher to depart from her own intuitions and base the research on real usage data. On the other hand, one might say that the data one can find in corpora could be biased, because the texts they come from have often been written by professional writers and, most likely, edited a number of times (Gilquin & Gries 2009: 7). The non-spontaneous nature of corpus data might prevent us from studying the mechanisms that lie behind the online processing (or interpretation) and production of language. Some spoken corpora – which offer more spontaneous data – do exist, however, they are usually limited in size and thus severely restrict the power of statistical analyses one could perform. Corpora also offer only very limited options of controlling for confounding variables, e.g. context or demographic factors such as age, education, gender. Experimental methods can alleviate the two problems mentioned above. Data obtained experimentally are spontaneous by definition and allow the researcher to tap into selected cognitive processes related to language; one can also control for many confounding factors in experiments. On the other hand, experiments have one major drawback: more often than not, they require subjects to perform tasks that they would not normally perform in real communicative situations – they have low ecological validity. According to Gilquin & Gries (2009: 5), experimental data are usually less naturalistic than corpus data. In sum, combining corpus and experimental approaches allows us to balance out the drawbacks of each method and better tap into what might be happening in the minds of language users.

Secondly, by using both corpora and experiments, the different methods can not only balance out the deficiencies of one another, but we can also investigate **different aspects** of a given phenomenon – this study is a case in point. A corpus technique – behavioural profiles – was chosen as the method for investigating the maximally general constructions for the reflexive markers and the prefixes. As far as the less general constructions for the reflexive marker *się* and the prefixes are concerned, two experimental methods (sentence-sorting and a task involving nonce verbs) will be used to establish whether speakers could build and use those constructions.

The corpus studies and experiments discussed in this thesis aim at obtaining converging evidence. What if the evidence will not converge, though? Diverging evidence can also have a very valuable empirical contribution. It can expose the limitations of different research methods, e.g. show that a given method is not suited for a particular type of data or that it fails to take some relevant phenomena into account which speakers are sensitive to (see e.g. Mos, van den Bosch & Berck 2012). Diverging evidence can also potentially show that the object of a given study (e.g. a grammatical category) is not a homogenous phenomenon and thus compel the researcher to rethink the nature of the studied category.

2.5. Conclusions

The chief question of this thesis – whether speakers can build and learn the general categories proposed by linguists – will be explored with the use of data on Polish reflexive verbs and prefixed verbs. This chapter presented an overview of the different approaches to the study of Polish reflexive verbs and prefixed verbs as well as the categories of different degrees of generality that have been proposed by researchers for each of the phenomena. The chapter also outlined the combined corpus and experimental approach that will be adopted for the studies discussed in this thesis. The following chapter (Chapter 3) will discuss the first of the empirical studies: a corpus study on Polish reflexive verbs.

Chapter 3: Reflexives corpus study

3.1. Introduction

In the literature on Polish reflexive verbs, researchers have postulated constructions of varying levels of generality. One approach in the research on Polish reflexive verbs was to look for the overarching meaning (or function) of the Polish reflexive marker *się*, usually in opposition to its heavy, pronominal counterpart *siebie*. The other approach was to look for the different possible functions of the light reflexive marker *się* and to classify the different reflexive verbs accordingly, e.g. if one of the functions of *się* is RECIPROCAL, the verb *całować się* would belong to the reciprocal category. If we look at these approaches from the perspective of usage-based linguistics, the first approach essentially means searching for one **maximally general construction** for *się*, while the second approach would be equivalent to searching multiple constructions of lower generality. Since linguists have postulated constructions of different levels of generality, we can investigate empirically how general are the constructions that native speakers can attain. In other words, using corpus and experimental data, we can investigate whether native speakers of Polish generalise over many instances of different Polish reflexive verbs and build the general categories for Polish reflexive marker *się* – be it one general construction for the marker as a whole or a few (less) general constructions for the different senses/functions of the marker.

What is more, Polish reflexive verbs exhibit high levels of idiosyncrasy, which is (somewhat indirectly) evidenced by the fact that *Uniwersalny Słownik Języka Polskiego* (‘The Universal Dictionary of Polish’; http://usjp.pwn.pl) includes approximately 6,500 separate entries for combinations of verb + *się*. Such a vast number of entries indicates that the author of the dictionary found it insufficient to only list *się* and ‘bare’ verbs (i.e. verbs without the marker *się*). The most likely rationale behind including so many reflexive verbs as separate entries was that the meaning of a particular reflexive verb is more than the sum of its parts – that is, the sum of the meanings of the verb and *się*. It is likely that the users of the dictionary would not be able to infer the meaning of a verb + *się* combination by simply putting the two together. So many Polish verb + *się* combinations listed as separate items might indicate that speakers store reflexive verbs as specific constructions and take advantage of those specific constructions for the production and processing of language in everyday communicative situations. The question that immediately comes to mind when we consider the above is: do the meanings expressed by reflexive verbs have enough in common for the speakers to abstract from them and build a general construction (or constructions) for the reflexive marker *się*?

Even though theoretical studies on Polish reflexives abound, there have been relatively few empirically-minded studies, e.g. studies using experimental or corpus data. Without empirical data, we cannot know whether the general constructions for Polish reflexives postulated by linguists are present in native speakers' minds. The present thesis seeks to bridge this gap. This chapter will discuss the results of a corpus study ran on data on Polish reflexives. The corpus study will aim to see whether the two reflexive-type markers – *się* and *siebie* – form coherent usage-based constructions. For this purpose, behavioural profiles (Divjak & Gries 2006) of both markers will be built based on data from the plTenTen & NKJP corpora.

3.2. Corpus study: behavioural profiles of *się* and *siebie*

In the corpus study, behavioural profiles of the two reflexive-type markers – *się* and *siebie* – will be constructed and explored statistically with correspondence analysis and conditional inference trees. Behavioural profiles are a technique that enables the researcher to explore the syntactic, morphological, and semantic properties of contexts that a given construction appears in. In other words, behavioural profiles facilitate investigating the distribution of a construction at a level higher than the individual words it co-occurs with.

The study aims to see whether stable and coherent behavioural profiles of Polish reflexives – *się* and *siebie* – could be constructed and investigate what kinds of variables govern the behaviour of either marker. In particular, the study is designed to investigate whether there were general properties in corpus data that could distinctly characterise each reflexive-type marker. If such general properties are found, we could surmise that native speakers could build general categories for the reflexive markers based on the input. These properties would indicate that the contexts where a given marker appears have something in common that will remain when the specific contexts are abstracted from. The chief hypothesis the corpus study will pursue is that the two markers have two distinct behavioural profiles and the behaviour of each marker is governed by high-level properties: *siebie* marks events where participants are more distinguishable (Tabakowska 2003a) and the action is more volitional (Dancygier 1997); *się*, on the other hand, is more likely to mark events with less distinguishable participants and less volitional actions. The next section will present an overview of the corpus study method: behavioural profiles.

3.2.1. The method: behavioural profiles

Behavioural profiles is a technique pioneered by Divjak (2006) and Gries (2006) (for a general overview see Divjak & Gries 2009; Gries 2010). It has proven very effective in many studies to date, e.g. Berez & Gries (2009) on the polysemy of *get*, Divjak & Gries (2008) on the near synonymy of Russian 'try'-verbs, Gries & Otani (2010) on the synonymy and antonymy of English *small-* and *large-*adjectives, and Jansegers, Vanderschueren & Enghels (2015) on the polysemy of the Spanish verb *sentir*. At the basis of the approach lies the assumption that the distribution of a word or expression reflects its semantic, pragmatic and functional characteristics (Gries 2010) – an assumption inspired by Harris' (1954) distributional hypothesis that words occurring in similar contexts should have similar meanings and Firth's (Firth 1957: 11) famous saying: "You shall know a word by the company it keeps". Unsurprisingly, behavioural profiles have been predominantly used in studies of lower-level constructions, i.e. studies on verbal synonymy and polysemy – the distributional properties of a given word (or words) can be used to predict the choice of a word in context (synonymy) or to discover which meaning clusters form coherent wholes (polysemy). This study will venture beyond verbal polysemy/synonymy and utilise behavioural profiles to explore constructions that have usually been claimed to be grammatical, and thus more general: Polish reflexive verbs.

The essence of behavioural profiles lies in annotating multiple corpus examples for morphological, syntactic and semantic characteristics (ID tags) and, subsequently, analysing them with the use of statistical methods such as cluster analysis, correspondence analysis, logistic regression or conditional inference trees. Behavioural profiles thus differ greatly from other corpus techniques that take into consideration only the distribution of words, e.g. collocations and colligations. This means that behavioural profiles consider much more information than the surface co-occurrence of words (collocations) or the presence of a word in a particular grammatical construction (colligations). Moreover, in contrast to many purely descriptive methods, behavioural profiles are deeply couched in linguistic theory – the properties for annotation always stem from theoretical considerations (Gries 2010: 325). Owing to this, the insights derived from behavioural profiles can be further explored in experimental paradigms and as such are potentially compatible with findings in psycholinguistics and cognitive neuroscience.

Behavioural profiles are also different from the early studies in usage-based linguistics. One of the main theoretical principles of Langacker's (1987) Cognitive Grammar was that language is usage-based, that is linguistic constructions emerge as a result of speakers' generalising over multiple usage events. Despite having adopted the usage-based thesis as the theoretical foundation, many studies in usage-based linguistics did not base their conclusions on actual usage data and did not venture beyond introspective analyses (Divjak & Gries 2009: 59). Introspective research is essential for the discipline as a whole in that it supports the development of the cognitive-linguistic theory and acts as a source of new hypotheses. That notwithstanding, introspective investigations inherently suffer from the biases of a particular researcher (Gibbs 2007: 4). Behavioural profiles offer a way to alleviate those biases because judgments are based on many contexts and parameters as opposed to a number of carefully selected (or contrived) sentences, as is the case with introspective research. Moreover, each ID tag in a given study should have an operational definition, which further increases the replicability and objectivity of behavioural profiles.

A behavioural profile analysis consists of four steps:

1. the retrieval of (a representative random sample of) all instances of a word's lemma from a corpus;
2. a (so far largely) manual analysis of many properties of the word forms (i.e. the annotation of the ID tags);
3. the generation of a co-occurrence table;
4. the evaluation of the table by means of exploratory and other statistical techniques (Divjak & Gries 2009: 61).

First, one must extract examples of the object of study from the corpus. In behavioural profiles, examples must correspond to roughly "'natural' unit[s] of expression" (Divjak 2015: 46). A "natural unit of expression" might be a sentence, a clause, or even the context of a whole utterance. The whole set of extracted examples must also form a representative and random sample, which prevents the sample from being biased in favour of a particular genre or single author's style. In practice, it is best if each example comes from a different text by a different author, e.g. different websites, different books, or different conversations in the case of a spoken corpus. The condition of representativeness and randomness also helps satisfy the assumption of independence of observations, which forms the basis of inferential statistics (see e.g. Freedman 2010, Chapter 2).

Step two involves inspecting the extracted examples closely and annotating them for many different properties and thus assigning ID tags to the examples. The ID tags can correspond to morphosyntactic, semantic, and pragmatic categories, e.g. tense, the semantic category of a verb, or politeness respectively. The annotation, in most cases, must be done manually and is thus dependent on the judgment of a particular researcher. In other words, the annotation is somewhat introspective – that is, it exhibits a characteristic it was designed to avoid. In order to reduce the amount of possible subjective judgment within the analysis, it is crucial that the properties that the data are tagged for lend themselves to operational definitions. Operationalising notions such as tense or number should not be particularly difficult, because they are encoded by specific affixes in most languages. Semantic properties, however, will always require precise definitions before they can be operationalised. One can also use the already available resources such as dictionaries or semantic databases (see WordNet, for example) to objectify the process of assigning semantic ID tags.

Besides some amounts of introspection, behavioural profiles as a method share the common disadvantage of corpus methods in general, that is, their significant reliance on the corpus the analysis is based on. In most official corpora, a large majority of data comes from carefully edited written sources, usually produced by highly educated authors (i.e. books or newspapers). If an analysis is based exclusively on such corpora, its findings cannot be extrapolated beyond the high registers of language. This study will seek to offset this problem by including two different corpora in the analysis: the official corpus of Polish (NKJP) and a corpus of Internet language (plTenTen); see section 3.2.3. for more details about the corpora and a discussion on how the inclusion of two different corpora can make the conclusions of an analysis more robust.

After the data have been annotated, the researcher must generate a co-occurrence table, which shows how many times each feature occurs in the dataset and with which example the feature is associated. This step quantifies the frequencies of qualitative properties and thus enables statistical techniques to work – without the quantification, no numerical analysis would be possible. Statistical analysis constitutes the final step in behavioural profiling. The annotation usually produces too many datapoints for a researcher to be able to make conclusions just by eyeballing them, and the use of statistical methods becomes not only advisable but crucial for the technique to be applied successfully.

3.2.2. Predictions

The literature on the subject makes it possible to formulate some more detailed predictions about the possible behavioural profiles of the two Polish reflexive-type markers, *się* and *siebie*. All the predictions are displayed in Table 3.1.

properties predicted to co-occur with *się*	properties predicted to co-occur with *siebie*
reflexiva tantum	*sam*
reciprocal and impersonal situation types	emphasis
preverbal clause position	co-ordination
overt subjects	volitional actions
verbs of motion and bodily care	verbs of communication, perception, and cognition
	animate subjects

Table 3.1. Properties of contexts predicted to co-occur either *się* or *siebie.*

Let us now look at the predictions in more detail. Firstly, *siebie* is more likely to appear in situations when an action is performed volitionally, and *się* in situations where the volitionality is reduced. Dancygier suggested that *się* carries a role-neutralising function: "it reduces the number of expressed participant roles, by eliminating syntactic expression of some of those roles in central syntactic slots, and/or by diminishing the distinction between roles held by the central participant" (Dancygier 1997: 325). If participant roles are indeed neutralised in cases when *się* is used, it entails that the agent role is made less prominent. Once the agent's salience is reduced, we should also expect a reduction in the volitionality of the action/process being expressed. Dancygier (1997: 315) observed that the use of *się* "often results in a less volitional or controlled interpretation of the activity referred to". Conversely, *siebie* should not entail any reduction of volitionality. Volition can only be attributed to animate subjects (humans in particular), who can consciously instigate actions. The type of subject, therefore, constitutes a very good benchmark of volition(ality). Consequently, *siebie* should be more likely to appear in clauses where the subject is animate. Importantly, the type of subject lends itself to straightforward and reliable operationalisation, since the judgment of whether a subject is a human or a different entity is not usually subjective.

Another property of *siebie* usually appears in contexts where co-reference is unexpected – it draws attention to the fact that the subject also happens to be the object of a transitive verb, which would otherwise normally require two different entities. In other words, *siebie* emphasises the unusual situation in which the subject is also the object of a transitive verb. By extension, *siebie* might be more likely to occur in emphatic contexts in general. The typical emphatic marker in Polish is *sam,* which conveys that the action was performed only by the subject and without any external help; it also indicates high volitionality of an action. Dancygier (1997: 313) suggests that "[t]he emphatic marker is independent from the reflexive, and has a full paradigm of forms marked for number and case". Even though the emphatic marker is independent from the reflexive-type marker, it does co-occur with reflexive-type markers, especially with *siebie*. A query in NKJP (*The National Corpus of Polish*) indicated that the conditional probability that *się* occurs with *sam* is 0.0024, whereas, in the case of *siebie,* it grows to 0.0625 – this means that we are 26 times more likely to find *sam* co-occur with *siebie* than *się*. It is, therefore, reasonable to hypothesise that *sam* will predict the occurrence of *siebie*.

There are also semantic classes of verbs that might correlate more strongly with either *się* or *siebie. Siebie* might be more likely to occur with verbs of communication, perception, cognition, and *się* is more likely to appear with verbs of motion or bodily care. Verbs of communication, perception, and cognition convey situations in which participants are the least likely to be co-referential. In contrast to verbs denoting grooming actions, e.g. shaving, we do not normally expect people to talk to themselves, see themselves or assess themselves. Since "the prototypical function of reflexive markers is to signal co-referentiality of participants for events in which participants are normally distinct entities" (Kemmer 1993: 66), we can predict that *siebie* would more frequently occur in such situations. In contrast to verbs of communication, perception and cognition, verbs of motion and bodily care verbs denote actions where we would often expect both the acted and the acted upon participant to be co-referential. Such actions "tend to be conceived as unary or atomic" (Kemmer 1993: 58) and have low distinguishability of participants. Hence, we would expect *się* to occur more frequently with these classes of verbs.

3.2.3. Choice of corpora

The behavioural profiles for *się* and *siebie* will be built upon data extracted from two corpora: The National Corpus of Polish (NKJP) and plTenTen. The ultimate aim of such a choice of corpora was for the composition of data sources to reflect to some extent the real input that the native speakers are exposed to in their daily interactions with language. NKJP was supplemented with data from plTenTen because the former contains hardly any texts extracted from the Internet – a mere 7% of all texts in NKJP came from Internet sources. A study conducted in 2013 by Ofcom (the UK's communications regulator) concluded that Internet news was the main source of information for ~55% people aged 18-35. The dominance of the Internet at the expense of the printed media makes the 7% of Internet-derived texts included in NKJP an amount too small to reflect the reading habits of modern-day native speakers of Polish. The two subsections below (3.2.3.1 and 3.2.3.2) will give a brief overview of both corpora used in the study and outline their advantages and disadvantages as sources of quantitative data.

3.2.3.1. The National Corpus of Polish

The National Corpus of Polish (NKJP) was built by a consortium of three Polish academic institutions: IPI PAN, IJP PAN, and The University of Łódź, and PWN, an academic publisher. NKJP incorporated parts of previous projects conducted by these institutions: the IPI PAN corpus and the PWN corpus. The consortium independently acquired more documents later. The final balanced version of the corpus consists of 300M segments[11], which makes it the largest 'traditional' corpus of the Polish language (Lewandowska-Tomaszczyk et al. 2012).

The two principles that guided the construction of the corpus were representativeness and balance. According to (Górski & Łaziński 2012: 26), a corpus is representative when it reflects some reality external to the corpus itself, that is when it reflects the structure of the language or the dialect on which it is based. As far as balance is concerned, a corpus is balanced when none of its elements dominates over any other; in practice, it meant that none of the sources should comprise more than 50% of the corpus (Górski & Łaziński 2012: 30). The proportions of texts in the written section of NKJP were based on readership and circulation statistics in Poland; the proportions were subsequently adjusted to make

[11] In most cases one segment was equal to one orthographic word (a string of characters stretching from one space to another). Some segments, however, were shorter than an orthographic word, for instance: *-że* and *-li* particles, parts of double names and surnames (*Jean-Pierre* or *Nowak-Jeziorański* would both be interpreted as two segments by NKJP), etc. (Szałkiewicz and Przepiórkowski 2012: 61).

the corpus more balanced. The final composition of the corpus is presented in Table 3.2.:

Source	Corpus share (%)
Newspapers	50
Fiction	16
Non-fiction	14
Spoken	10
Internet	**7**
Miscellaneous (official documents, etc.)	3

Table 3.2. Composition of NKJP

As we can see, written texts comprise 90% of the corpus. This comes as no surprise when we consider the fact that the process of collecting and transcribing spoken language data is costly and time-consuming (Pęzik 2012: 39). The corpus relies heavily on newspaper texts, which can be seen as a disadvantage. Newspaper texts are produced in a very deliberate way, and they usually undergo profound editing before they are sent for publication. As a consequence, newspaper language is not very 'natural', which can prevent a linguist from tapping into more general distributions.

The whole NKJP was automatically annotated for morphosyntactic features, such as part of speech, gender, number, or tense. Although its 93-per cent accuracy (Szałkiewicz & Przepiórkowski 2012: 96) does not attain perfection, it is remarkably high for a such a morphologically complex language as Polish and entirely sufficient for most purposes, especially for such small-scale queries as the two case studies that are part of this thesis.

3.2.3.2. plTenTen

In contrast to NKJP, the TenTen corpora were not developed by an academic institution, but a private company. The creators of the TenTen family aimed at constructing large-scale corpora for major languages – the size of the corpora was to be in the order of 10 billion words (Jakubíček et al. 2013). Presently, the TenTen family offers corpora for 31 languages, including Polish. The plTenTen corpus was extracted from the World Wide Web in June 2012 using the SpiderLing crawler and its final version contained ~7.7 billion words. It is available as part of TheSketchEngine framework.

PlTenTen is a web corpus, but unlike provisional corpora of the Web-as-a-Corpus type, it is a snapshot corpus. Thus, the content of the corpus does not change over time, and the queries conducted in TheSketchEngine are entirely replicable. As previously mentioned, the makers of NKJP were primarily concerned with balance and representativeness. In contrast, the primary guiding principle of plTenTen was its size — even if we consider the unbalanced version of NKJP, plTenTen is still more than four times larger. Certainly, an automatically crawled corpus cannot be representative in the sense of NKJP. PlTenTen does not contain carefully selected printed texts, but at the same time, the share of online texts in NKJP amounts to 7% (Górski & Łaziński 2012: 33). If figures about news readership are suggestive of readership in general, people spend more and more time reading online materials — online news platforms are the main source of news for ~55% of respondents aged between 18 and 35 (Newman & Levy 2013: 26). One could say that plTenTen is representative in its own way because it represents the most common source of linguistic input (besides conversation) for many people. The issue of representativeness cannot be easily settled – both corpora are (un)representative in their own way.

PlTenTen has one feature that puts it at an advantage to NKJP: it contains more spontaneously generated language, e.g. blogs or online diaries. If the language in a corpus is less deliberate, we are more likely to tap into actual usage patterns, as opposed to language use mediated through stylistic norms. We are also more likely to come across words or expressions that are not used in more official registers.

Besides the advantages mentioned above, plTenTen has its own share of issues. Since the crawling algorithm is automatic, the crawler cannot easily distinguish between strings of real human language and strings of machine-generated gibberish. Some website designers automatically generate nonsensical content full of relevant keywords in order to boost their webpage's rank in search engines (Gyongyi & Garcia-Molina 2005). The amount of such pages that were incorporated into the plTenTen corpus is unknown, but, as we will see later, the problem surfaced in the present study. The other issue is that, even though the corpus is tagged morphosyntactically, the authors do not provide any accuracy statistics, so we do not know how reliable their automatic annotation is.

Nevertheless, this problem was not relevant to the studies included in this thesis, because the annotation of the examples was done manually and any errors in tagging were immediately identified.

3.2.4. Data

The data for the study came from two corpora: the National Corpus of Polish (NKJP) and plTenTen – due to the reasons outlined in the previous section. For the purpose of the study, 250 occurrences of *się* and *siebie* each were extracted from each corpus (1000 contexts in total). The syntactic position of the marker was not limited. Because *się* and *siebie* can occur either before or after a verb, the query included *się* and *siebie* both in a pre-verbal and post-verbal position. The amount of extended context for each example was as wide as either of the corpora would yield. The extracted examples were then inspected visually, and any unsuitable examples[12] were pruned. Finally, the first 250 examples (for each marker) were selected from the pruned dataset, and they constituted the final dataset that the behavioural profile analysis was conducted on. The following subsections (3.2.4.1. and 3.2.4.2) will give an overview of how the data were annotated and discuss the fundamental properties of the obtained data.

3.2.4.1. Annotation scheme

The 1,000 extracted examples were subsequently tagged for 16 variables, which are reported in Table 3.3.:

Type of ID tag	ID tag	ID tag levels	Example	Translation
morphological	tense	past	*ogolił się*	He shaved.
		present	*goli się*	He is shaving.
		future	*ogoli się*	He will shave.
	aspect	perfective	*ogolił się*	He shaved.
		imperfective	*golił się*	He was shaving.
	mood	indicative	*ogolił się*	He shaved.
		subjunctive	*ogoliłby się*	He would shave.
		imperative	*ogól się!*	Shave!
	gender	masculine	*obraził się*	He got offended.

[12] plTenTen, being a corpus of Internet language, sometimes contains content from websites that are computer-generated gibberish, aimed at optimising a website's position in search engines. Examples of this sort were removed from the database.

		feminine	*obraziła się*	She got offended.
		neuter	*obraziło się*	It got offended.
	person & number	1st person singular	*ogolę się*	I will shave.
		2nd person singular	*ogolisz się*	You will shave.
		3rd person singular	*ogoli się*	He will shave.
		1st person plural	*ogolimy się*	We will shave.
		2nd person plural	*ogolicie się*	You will shave.
		3rd person plural	*ogolą się*	They will shave.
	verb type	finite	*ogoliliśmy się*	We shaved.
		infinitive	*ogolić się*	shave
		past participle	*ogoliwszy się*	having shaved
		present participle	*golący się*	shaving
syntactic	clause position	pre-verbal	*Musimy **się** zobaczyć.*	We must see each other.
		post-verbal	*Skup **się**!*	Focus!
	coordination	present	*Zobaczyła w lustrze siebie I swoje dziecko.*	She saw herself and her child in the mirror.
		absent	*Po śniadaniu się ogolił.*	He shaved after breakfast.
	overt subject	present	***Ja** się wcale nie chwalę.*	No, I’m not bragging at all!
		absent	*Chwalili się swoim bogactwem.*	They bragged about how rich they were.
	V-INF construction	present	*Nie **mogłam** się **zapisać** na ten kurs.*	I couldn’t enrol in this course.
		absent	*Zapisałam się na kurs.*	I enrolled in this course.
	sam	present	*Oszukał **sam** siebie.*	He cheated himself.
		absent	*Zobaczyła siebie w telewizji.*	She saw herself on TV.
semantic	*tantum*	present	*Dzieci bały się burzy.*	The children were afraid of the storm.
		absent	*Zanurzył się w wodzie.*	He dipped into the water.
	volition	present	*Zaakceptowali wynik.*	They accepted the result.

		absent	*Potknęła się.*	She tripped.
	semantic verb class	motion	*poruszać się*	move
		perception	*przyglądać się*	observe closely
		contact	*dotykać się*	touch oneself/one another
		communication	*porozumiewać się*	communicate
		competition	*ścigać się*	race
		change	*przekształcić się*	transform
		cognition	*namyślać się*	ponder/ruminate
		consumption	*najeść się*	eat until full
		creation	*wytworzyć się*	emerge/appear
		emotion	*złościć się*	be angry
		social	*spotykać się*	meet (with someone)
		bodily	*myć się*	wash
	subject type	human	*lekarz*	doctor
		(other animate)	*kot*	cat
		inanimate	*trawa*	grass
		abstract	*uwaga*	attention
	emphasis	present	***Niechże** już przestanie!*	Oh, why won't she stop!
		absent	*Samochód porusza się po drodze.*	The car is moving down the road.
	situation type	reflexive	*Najadła się.*	She ate until full.
		reciprocal	*Pobili się.*	They went into a fist fight.
		passive	*Szkoła się buduje.*	The school is being built.
		impersonal	*Tak się nie robi!*	One doesn't do that!

Table 3.3. ID tags used in the study

The first set of ID tags (*tense, aspect, mood, gender, person & number, verb type*), i.e. the morphological variables, are properties typical for all Polish verbs. The morphological properties of a verb convey rich semantics pertaining to how a situation is portrayed. Aspect, for instance, can convey that an action or event happened once or, just the opposite, that it happens habitually. The number and person convey information about who the subject of a sentence is. Humans often speak about what they did themselves; consequently, the reflexive-type markers could correlate with the first person singular or plural, because one of their postulated functions is signalling the coreference of subject and object.

Two syntactic variables – *clause position* and the presence of *overt subject* – are linked together. The clause position might depend to some extent on whether an overt subject appears in the clause. *Się* seems to occur before the verb more frequently than siebie, and when it occurs before the verb, an overt subject is usually present. A variable that could also predict *siebie* to some extent is *co-ordination*. In this study, coordination is used in reference to clausal objects and means that the object consists of two (or more) entities, e.g. *He did it to protect **himself and others***; in such contexts, only *siebie* can be used in Polish. The presence of *sam* ('on one's own') might also be a good predictor for *siebie*, because *siebie* frequently occurs in the phrase *samego siebie* (lit. 'him himself').

The semantic variables pertain mostly to the nature of the action encoded by a given verb. *Subject type* and *volition* jointly allow assessing how likely it is that the subject instigated the action expressed in the clause. If the subject of the analysed sentence is human (or at least animate), it is much more likely that it instigated the action, as opposed to inanimate and abstract objects, whose causal powers are rather limited. The other property, i.e. volition, was annotated more subjectively – I assessed whether the context suggests that the subject acted deliberately. An example of a volitional context might be the following sentence:

(3.1)	*Reżyser*	*obsadził*	*siebie*	*w*	*roli*	*głównej*.
	director	cast$_{\mathrm{PST}}$	siebie$_{\mathrm{ACC}}$	in	role	main

'The director cast himself in the lead role.'

Casting oneself (or anyone else) in a role is (at least usually) an intended, that is volitional, action – such a context would be labelled as volitional. As a contrast, the context in (3.2) would be labelled as non-volitional:

(3.2)	*Nie*	*poznały*	*sam-e*	*siebie.*
	not	recognised	sam-PL.NON-VIR	siebie$_{\mathrm{GEN}}$

'They didn't recognise themselves.'

In the above context, the impression of not recognising oneself does not constitute an action that depends on one's will – it happens whether we intend it or not. Hence, the context would be tagged as non-volitional.

The variable *semantic verb class* stands for the semantic class of the analysed verb. The semantic classes used in this study were taken directly from the Polish equivalent of WordNet, Słowosieć[13]. WordNet's classification consists of 15 semantic classes of verbs: communication verbs (e.g. *discuss*), verbs denoting cognitive actions (e.g. *ponder*), etc. In the process of annotation, every verb was queried in Słowosieć and, subsequently, coded based on the information obtained from the system. Nevertheless, in a small number of verbs, either the semantic class in Słowosieć did not correspond to the context or Słowosieć did not give any class for this particular verb. In such cases, the semantic class of the verb in question was adjusted (or assigned) based on the annotator's judgment. The adjusted verbs can be found in Appendix 6.

Reflexives can also perform emphatic functions (Kemmer 1995), so the data were also coded for *emphasis*. Kemmer defined the emphatic function of *-self* in English as "a grammatical device for accessing referents of some degree of prominence in the discourse"; the referents are chosen "in contrast to other potential entities that the speaker assumes might have been more likely to be referred to instead at that point in the discourse" (1995: 60). This definition seems to apply to the function of Polish *sam*, but it does not seem accurate for Polish reflexives. Moreover, emphasis thus defined would be difficult to operationalise. In this study, the context was classified as emphatic when it contained one or more expressions of the following types: (1) comparatives, superlatives or diminutives; (2) intensifiers (e.g. *niesamowicie* 'incredibly' or *okropnie* 'terribly'); (3) evidential discourse markers indicating certainty (e.g. *oczywiście* 'obviously').

Finally, two more variables were annotated for: *tantum* and *situation type*. *Tantum* is a binary yes/no variable that stands for whether a verb is a reflexiva tantum verb or not. A reflexiva tantum verb is a verb that only has a reflexive version – it does not have a non-reflexive counterpart. The verb *bać się*, for instance, is a reflexiva tantum verb, because there exists no 'bare' verb *bać* without *się*. Reflexiva tantum

[13] http://plwordnet.pwr.wroc.pl [Accessed 14 Oct 2018].

verbs do not have any semantic property in common – they come from diverse semantic classes, for instance, communication (e.g. *odezwać się* ’speak out’), cognition (*zastanawiać się* ’wonder’), or social events (*rozstawać się* ’part’). Even though it does not stand for any semantic property, the reflexiva tantum variable can indicate that a given reflexive verb is lexicalised, or, speaking in terms of constructions, that a given reflexive verb is a specific construction that achieved a unit status in the minds of speakers. The variable *situation type* corresponds to some extent to the different types of *się* postulated by researchers. The four different types (reflexive, reciprocal, impersonal, and passive) constitute a general classification and do not exhaust every option postulated in the literature. The reason for such a decision was that those four categories lend themselves to clear operationalisation. The context was classified as reciprocal if more than one entity was involved in the same mutual action (e.g. people hugging each other). In passive contexts, the subject could not be the agent of the action taking place (e.g. *Szkoła się buduje* ‘The school is being built’). The ‘impersonal’ label was given to contexts in which agents could not be determined and the sentence of subjectless (e.g. *W Polsce je się obiad o piętnastej* ‘One has lunch at 3 pm in Poland’). Finally, ‘reflexive’ was an umbrella category that classified all other situations (true reflexives, extensions of true reflexives, and inchoatives/resultatives in terms of the categories discussed in Chapter 2).

The annotation process was manual; no cross-validation with other researchers was performed. The risk that errors are made in the process of manual annotation was reduced through strict operationalisation of variables. The syntactic and morphological tags were inferred directly from the formal properties of a given sentence (e.g. the presence or absence of an overt subject). When it comes to semantic variables, either an external source was consulted (WordNet) or an operationalisation was developed that reduced the likelihood of arbitrary judgments, e.g. the context was classified as emphatic only if certain words or constructions occurred in the context. The only variable that allowed for subjective judgment was the volitionality of the action taking place in the context.

3.2.4.2. Dataset structure

After inspecting the structure of the encoded sample, it appeared that the data exhibited some problems, which could impact on the subsequent analysis. Firstly, the NKJP dataset included 48 contexts in which target verbs were neither infinitives nor finite verbs: 37 deverbal nouns, 10 present participles, and 1 past participle. The issue might have arisen because of erroneous morphosyntactic tagging in the NKJP. As the dataset had not been inspected prior to the annotation to avoid biasing the annotation, the problem had not been detected until the annotation process was completed. The problematic contexts constituted only 4.8% of the entire data, and thus they remained in the dataset so that no data were lost[14].

Another problem was the low frequencies of some variables. Low frequencies can yield unreliable estimates in regression analysis, with large standard errors. In CA, low frequencies can create outliers in the plot and distort the visualisation. Animal subjects occurred only five times in the sample, and since such a number might pose problems for regression analysis, this category was merged with animate human subjects into one category *animate* in both samples. Semantic verb class contained three verb classes whose frequency was lower than 10: *competition* (4 occurrences), *consumption* (4 occurrences), and *bodily care* verbs (4 occurrences).

The data were inspected again, and the verbs were put in classes closest to the classes selected in the initial coding. Eventually, the semantic verb class variable consisted of 11 levels: *change, cognition, communication, contact, creation, emotion, motion, perception, possession, social, stative*. The last problematic variable was *number/person* – the second person plural category occurred only four times in the sample. The category was merged with second person singular to obtain a single *2sgpl*[15] category.

Low cell counts were not the only issue present in the data. The variable *mood* contained hardly any occurrences of levels other than *indicative*[16]; the variable also exhibited a very high number of NAs (320 missing values). Such a distribution

[14] As a check, the final model in the regression analysis was also run for the dataset without the problematic contexts. The results were virtually identical, which implies that the problematic contexts do not distort the analysis (see Section 3.2.5.2.3. for more details).
[15] Second person singular/plural.
[16] The imperative mood occurred only once and the subjunctive mood did not occur at all.

renders a variable virtually unusable since a regression cannot actually 'access' all the different levels when predicting the marker. As a consequence, the variable *mood* was not used in any of the statistical analyses.

3.2.5. Analysis

The corpus data will be analysed using three statistical techniques: correspondence analysis, logistic regression, and conditional inference trees. The first technique, correspondence analysis, will help visualise and explore the data, see which properties correlate with one another, and conduct a preliminary verification of the predictions outlined in Section 3.2.2. Binomial logistic regression is a "standard" confirmatory technique that allows one to model a binary choice – such as the choice between *się* and *siebie.* It can tell which properties impact the choice between *się* and *siebie* and how much they impact the choice. For instance, it can tell that *siebie* is *n* times more likely to occur in contexts containing the word *sam*. What it cannot tell, however, is which **combinations** of properties will make *się* or *siebie* occur in a given context – conditional inference trees allow for doing exactly that. For instance, a conditional inference tree is capable of showing that if a context contains a perception verb, an animate subject and the word *sam*, and the verb is not a *reflexiva tantum*, *siebie* will most likely occur in this context.

If the results of the three types of analyses converge, we will obtain very robust evidence in support of (or against) the hypothesis that speakers of Polish could build coherent general usage-based categories for *się* and *siebie.* It would mean that the conclusions we can draw on the basis of the data are robust and consistent regardless of the type of statistical method used to analyse the data. The three following sections will present the results of the correspondence analysis (Section 3.2.5.1), logistic regression (Section 3.2.5.2), and conditional inference trees and forests (Section 3.2.5.3.).

3.2.5.1. Correspondence analysis

As the first step, the transformed data were explored by means of correspondence analysis – a technique that allows for visualising multivariate data. Correspondence analysis (CA) "is a method of data analysis for representing tabular data graphically" (Greenacre 2007: 1). In other words, correspondence analysis is an **exploratory** statistical technique that allows the researcher to "simplify" the graphical presentation of the data and look for correlations between multiple variables. As a result, CA allows the researcher to explore the data and see whether the properties expected to correlate with one another do exhibit such correlation. CA can also elucidate relationships that were not previously predicted to occur but nevertheless exist in the data. Overall, correspondence analysis can aid in making a previously formulated hypothesis more precise and formulating new ones.

3.2.5.1.1. Method

In essence, correspondence analysis allows for visualising the relationships between many variables that could not be visualised otherwise. Thanks to such a visualisation, we can see which variables (or properties) are strongly related to one another and which ones are not. Visualising the relationships between two or three variables graphically is not difficult – humans can see and interpret up to three dimensions, but the task becomes much more complicated as the number of variables increases. Correspondence analysis is well-suited for the analysis of behavioural profile data, as behavioural profiles usually operate on multitudes of variables. The corpus data extracted for the study on Polish reflexive-type markers was tagged for 17 variables. The overall number of possible values those variables can assume is 57, so a perfect visualisation of such data would require 56 dimensions – the number of dimensions is calculated as the number of possible values minus one. Such a visualisation, although 'perfect', would have no analytic value because humans cannot perceive or think in 56 dimensions. To alleviate this problem, CA reduces the number of dimensions and visualises all variables on a two- or three-dimensional plot and presents the relationships between them as distances – the smaller the distance between two features, the more related they are.

Since CA is essentially a dimension-reducing technique, some information will be lost in the process (Greenacre 2007: 43). Dimension reduction is analogical to the process of drawing a cube (or any other solid figure) on a piece of paper – we are unable to reflect every single property of a figure faithfully, e.g. the angles become distorted. Unfaithful as it is, the two-dimensional representation of a solid figure remains informative and serves its purpose. Similarly, CA also serves its exploratory purpose, even though some information is lost in the process. We have to bear in mind that due to the information loss, some part of the variation in the data will remain unexplained. In CA, the variation is called inertia, so the explained variation will be displayed as explained inertia. As long as the explained inertia remains high – and, consequently, the unexplained inertia is low – the analysis is interpretable and gives robust insights. One can artificially inflate the amount of inertia explained by adding more variables with multiple levels. An analysis with, say, 50 different variables would then be likely to have higher explained inertia than an analysis with 5 variables, but it would have no explanatory power, because the very purpose of doing research is to pinpoint the very few variables that crucially influence the behaviour of the phenomenon under investigation. Simply speaking, one must not include in the analysis variables whose inclusion cannot be justified on theoretical grounds. Another reason not to include too many variables in the analysis is that it would increase the likelihood of finding correlations by sheer chance (Glynn 2014: 134).

In CA, an analysis begins with a table that contains the frequencies of co-occurring features. Each feature has a profile that is expressed as a column containing the frequencies of how often a feature co-occurs with other features. Each profile is also given a mass, which is calculated as the marginal frequency (i.e. the column total) of the profile. Mass indicates the importance of a given profile within the analysis – the higher the mass is, the more important the profile is and, consequently, the greater the influence of this profile is on the outcome of the analysis (Greenacre 2010: 631). Subsequently, CA calculates the distances between profiles to assess how ‘similar’ or ‘dissimilar’ particular features are. The profiles and the distances between them are visualised in a two-dimensional or three-dimensional plot. The greater the distance between particular features, the greater the difference between them and the less likely they are to co-occur together.

To sum up, CA is an exploratory technique for categorical data, which enables the researcher to reduce the number of dimensions and visualise correlations between multiple variables. The technique constitutes an efficient tool for analysing complex datasets. One must be aware, however, that it should be used solely for exploratory purposes; as such it will never provide information as to whether the relationships found are statistically significant. In order to assess the statistical significance of relationships between factors, one needs to use a confirmatory technique such as regression analysis.

3.2.5.1.2. Procedure

The type of correspondence analysis employed in this study is multiple joint correspondence analysis using a Burt matrix to correct for low explained inertia caused by the inclusion of many variables (Greenacre 2007: 145–146); the correspondence analysis was conducted in R 3.4.0 (R Core Team 2017) with the package FactoMineR (Lê, Josse & Husson 2008). The procedure was as follows: at first, all possible variables were included in the model; subsequently, two other models were built – one for formal properties only (i.e. morphological and syntactic variables), the other for semantic and pragmatic variables. Glynn (2014: 141) warns that including too many variables (with multiple levels) may cause problems with interpretation and significantly decrease the explained inertia. Technically, however, there is no limit on how many variables (and levels of variables) can be analysed by multiple joint correspondence analysis (Greenacre 2007: 145–152). The full dataset analysis will serve a particular purpose in this study, which will be clarified in the following subsections.

Each analysis includes reflexive markers (i.e. *się* and *siebie*) as supplementary points. If we treat a point (in this analysis, a variable level) as supplementary in a correspondence analysis, we assign no mass[17] to it and "their contribution to [the overall] inertia is zero" (Greenacre 2007: 89). In other words, a supplementary point is passive – it will appear on the plot, but its presence will not change the position of other variables, which would have happened if the variable had not been given the supplementary status. Technically speaking, supplementary variables are

[17] Mass is the marginal frequency of a row or column in the input matrix for the correspondence analysis (Greenacre 2010: 631). Mass is an indicator of a point's importance in the analysis – the larger the mass, the more important a point is. Mass equal to zero would indicate that a point is of no importance to the analysis, and, consequently, it will not influence the outcome of the analysis in any way. See Section 3.2.5.1.1. for more information.

“project[ed] … on an already existing solution configuration” and thus “have no influence on the geometric orientation of the axes” (Greenacre & Blasius 2006: 3–40). In other words, supplementary variables are added to the plot only **after** the entire analysis has been conducted. The analyses were conducted to explain the behaviour of the reflexive marker, hence, the variable *marker* could not be used to explain the reflexives’ own behaviour, and as such had to be ‘eliminated’ from the analysis. On the other hand, the markers still needed to appear in the plots of the analyses, because we need to see which variables correlate with either *się* or *siebie*. Adding the markers as supplementary points prevented them from having an impact on the position of the remaining variable levels but allowed them to remain on the plot at the same time.

On a correspondence analysis plot, the variables which are likely to occur together form clusters – the smaller the distance between two variables is, the more associated they are. If variables appear in different quadrants of the plot, the association between them is low. If variables appear in two diagonally opposite quadrants (e.g. the upper right quadrant and the lower left quadrant), they are negatively associated with each other.

3.2.5.1.3. Results and discussion

First, an analysis of the full set of variables was conducted. We can see that the explained inertia summed over the two dimensions displayed in the plot amounts to 28.34 %, which is not a very high result. Since the explained inertia is quite low, the analysis should be treated with caution. That notwithstanding, the dataset used for the analysis includes a large number of variables, out of which not all must be correlated with the reflexive markers investigated (*się* and *siebie*). For this reason, the dataset is bound to contain much variation, and we cannot expect the explained inertia to reach very high amounts.

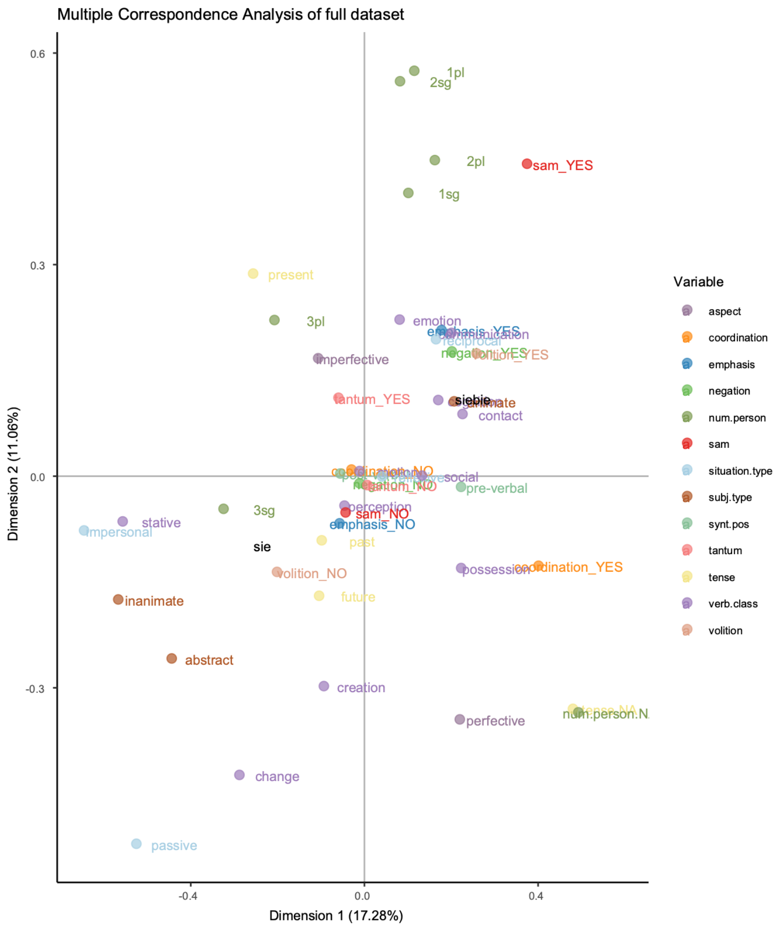


Figure 3.1. Correspondence analysis: all variables

In the plot of the analysis (Figure 3.1.), *się* and *siebie* lie in two different quadrants located on a diagonal, so we can say that they are negatively associated with each other. *Siebie* co-occurs with the following features: verbs of cognition, verbs of contact, verbs of communication, verbs of emotion, animate subjects, presence of negation, reciprocal situations, and volitional actions. *Się* is associated with non-volitional actions, third person singular, and the future tense. Chi-squared tests were run to see whether the differences in the distribution of variables between *się* and *siebie* are statistically significant. The results of the chi-squared tests showed that the difference in the distribution of a variable between the markers is statistically significant in six cases: verbs of communication ($\chi 2 = 24.73529$, df $= 1$, $p < 0.00001$), presence of emphasis ($\chi 2 = 33.37931$, df $= 1$, $p < 0.00001$), animate subjects ($\chi 2 = 46.33132$, df $= 1$, $p < 0.00001$), volitional actions ($\chi 2 =$

42.41966, df = 1, p < 0.00001), third person singular ($\chi 2$ = 72.13613, df = 1, p < 0.00001), non-volitional actions ($\chi 2$ = 68.18879, df = 1, p < 0.00001); it was not statistically significant also in six cases: verbs of cognition ($\chi 2$ = 2.60215, df = 1, p = 0.10672), verbs of contact ($\chi 2$ = 4.16667, df = 1, p = 0.04123), verbs of emotion ($\chi 2$ = 0.04878, df = 1, p = 0.8252), presence of negation ($\chi 2$ = 0.01961, df = 1, p = 0.88864), reciprocal situations ($\chi 2$ = 1.18519, df = 1, p = 0.2763), future tense ($\chi 2$ = 2.31507, df = 1, p = 0.12813)[18].

If we look at the features associated with either of the markers (provided that they exhibit statistically significant differences in distribution) in terms of initial predictions, we can see they converge in many cases. In line with initial predictions, *siebie* co-occurs with animate subjects, volitional actions, emphatic contexts, and verbs of communication. In the case of *się*, we have non-volitional actions. Surprisingly, the correspondence analysis on the full set of variables does not indicate that the presence of a reflexiva tantum verb is associated with *się*.

The plot contains an array of different variables, both morphosyntactic and semantic. It could be the case that only one of the groups governs the choice of the marker, e.g. the choice of the marker is sensitive only to semantic variables. Two more plots were generated to explore this issue: one with morphosyntactic and the other with semantic and pragmatic variables. The correspondence analysis of morphosyntactic variables explained 45.25% of inertia, while the one for semantic and pragmatic variables explained 31.6% of inertia. The explained inertia figures for both analyses are greater than the one for the analysis of all variables, so their interpretability is somewhat higher than the analysis of all variables. The plot of morphosyntactic variables is presented in Figure 3.2. below.

[18] Multiple chi-squared tests were run. To account for multiple comparisons, the p-values were adjusted using Bonferroni correction method. In Bonferroni correction, we divide the p-value (0.05) by the number of tests run (twelve in this case). The corrected p-value thus amounts to 0.00417.

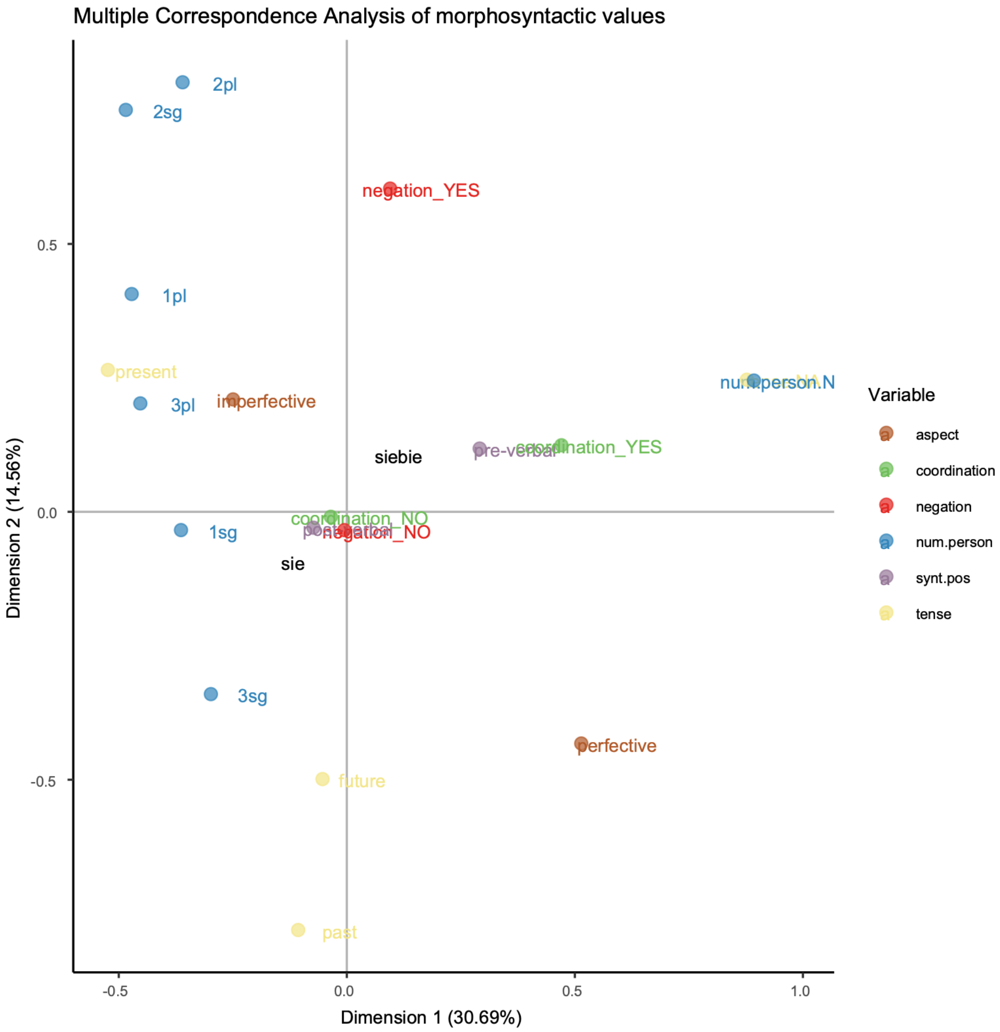


Figure 3.2. Correspondence analysis: morphosyntactic variables

As far as morphosyntactic variables are concerned, *się* and *siebie* are located near the origin. The only feature that is associated with *się* seems to be first person singular – the difference in the distribution of the variable between the markers was statistically significant ($\chi 2$ = 15.04167, df = 1, p = 0.00011). In the case of *siebie,* it is the preverbal position of the marker, but this variable missed statistical significance ($\chi 2$ = 0.04712, df = 1, p = 0.82815) [19]. Such an arrangement means that the correspondence analysis could not "find" many features that are distinctly associated with either of the markers.

As the last step in the correspondence analysis, pragmatic and semantic variables were examined (Figure 3.3.). No variables seem to clearly correlate with either of

[19] Again, the *p*-values were corrected for multiple comparisons with Bonferroni correction. The corrected *p*-value is 0.05/2 = 0.025.

the markers in this analysis. *Siebie* seems to be somewhat associated with communication verbs – the difference in the distribution of the variable is statistically significant ($\chi2$ = 24.73529, df = 1, p < 0.00001)[20]. The distances between *siebie* and the variables present in its quadrant of the plot are very large, and many variables lie near the axes, which suggests that the association might be small. We could say that *się* in this analysis correlates somewhat with perception verbs and contexts in which there is no *sam* and no emphasis (all statistically significant)[21]. Overall, the analysis of semantic variables also did not find many variables that would correlate with either of the markers.

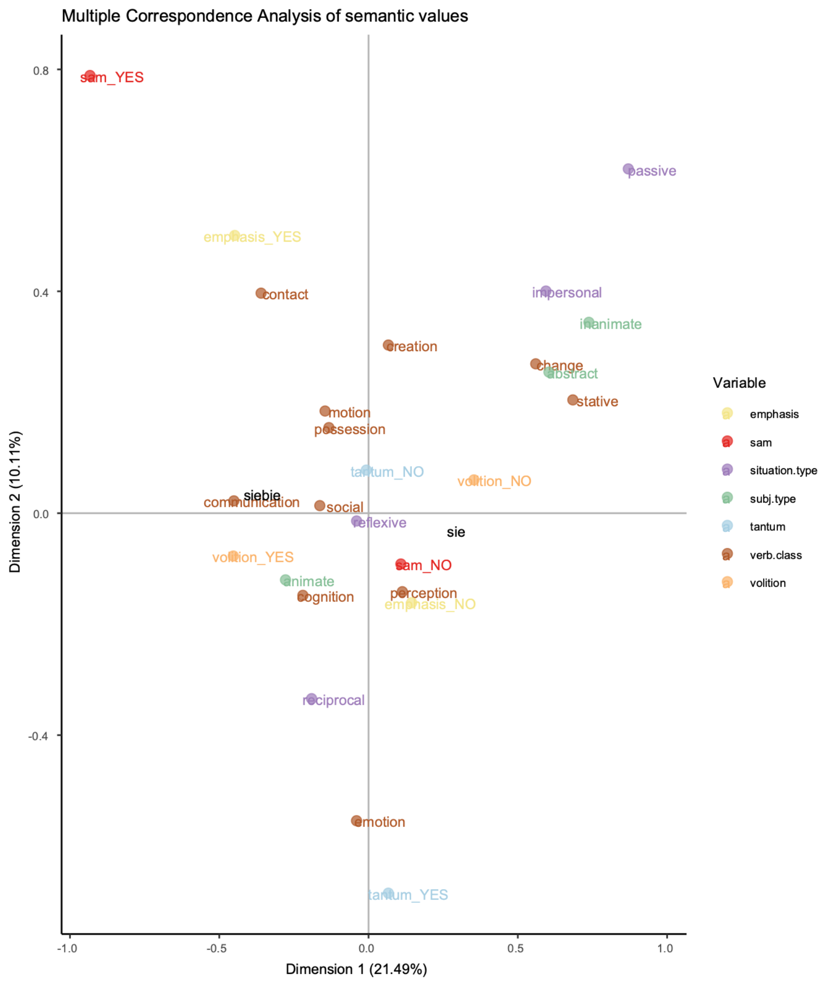


Figure 3.3. Correspondence analysis: semantic and pragmatic variables

---

[20] Corrected Bonferroni *p*-level 0.05/4 = 0.0125. Four analyses were run altogether for the variables co-occurring with *się* and *siebie*.

[21] Chi-squared tests: absence of emphasis ($\chi2$ = 29.60556, df = 1, p < 0.00001), absence of sam ($\chi2$ = 26.02696, df = 1, p < 0.00001), perception verbs ($\chi2$ = 15.62162, df = 1, p = 0.00008). Corrected Bonferroni *p*-level 0.05/4 = 0.0125.

The analyses of morphosyntactic variables and semantic variables suggest that neither of the groups can explain the choice of the marker effectively.

All conducted analyses are characterised by relatively low levels of explained inertia (~28-45%), which implies that their interpretability is moderate at best. Splitting the variables into separate groups did not improve the levels of explained inertia, which means that we cannot conclude that one of the groups (morphosyntactic vs semantic variables) explains the phenomenon better. Overall, *siebie* correlates with many more variables than *się* when take all the analyses into consideration. *Siebie* also seems to rely on more general properties: volition, emphasis, and animate subjects, which might be an early indication that *siebie* could form a general category in speakers' minds. When it comes to *się,* the analysis could not find many properties that would correlate with the marker, which might be an indication that it does not form one coherent general category – it might actually rely on a number of less general categories.

We must treat the results of the correspondence analysis with caution because it is only an exploratory technique. In order to make more robust conclusions about the data and extrapolate the results onto the phenomenon under investigation as a whole, we need to use a confirmatory statistical technique such as logistic regression. The results of the CA will be taken as an indication of the possible avenues to explore with the regression analysis. The following section will present the results of a logistic regression analysis of the corpus data for *się* and *siebie*.

3.2.5.2. Logistic regression

As mentioned in the previous section, an exploratory statistical technique can yield only an indication of the relationships within the data. To draw any robust conclusions, we must analyse the data with a confirmatory technique. The tagged contexts used in the correspondence analysis were used to build a binomial logistic regression model. A binomial logistic regression model seems particularly suited to the data on Polish reflexive-type markers, because it can model a choice between two options – *się* and *siebie* in this case. In a nutshell, a binomial logistic regression model can tell us whether a variable impacts the choice and, if it does have an impact, it can tell how large this impact is. To give an example, a binomial logistic regression model can tell whether *siebie* is indeed more likely to appear in contexts containing volitional actions and, if so, how much more likely. Before we proceed to the report of the analysis and its results, let us first get acquainted with the method.

3.2.5.2.1. Method

Binomial logistic regression allows one to model the relationship between a binary response variable, such as the choice of a reflexive-type marker, and one or more explanatory variables, e.g. pragmatic and semantic features. Logistic regression tests the statistical significance of the impact each predictor has on the predicted category. Owing to this, we can verify whether the results we have obtained for our sample are likely to occur again if we take another sample of the same "population" – this makes logistic regression an inferential statistical technique.

Binomial logistic regression is a type of a linear regression model. Essentially, all types of linear regression try to fit a linear function (plotted as a line) to data points, such that the *residuals* are as small as possible. Residuals can be visualised as vertical distances between the regression line and the data points. Figure 3.4. shows a regression line (blue) fitted to a "cloud" of data points (black dots); the figure also shows residuals as red vertical lines.

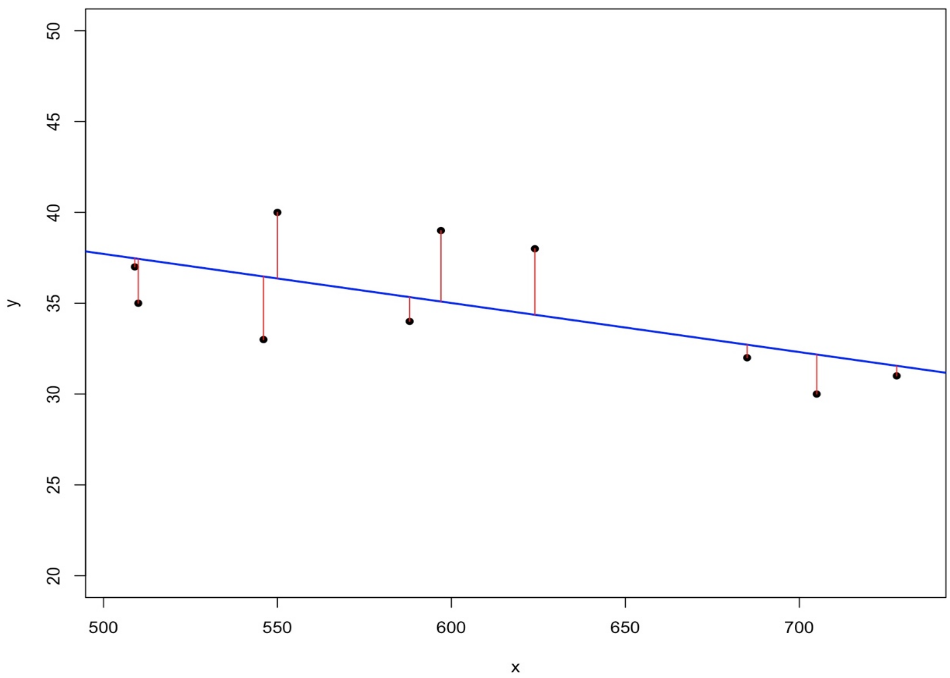


Figure 3.4. Regression line

Linear regression was originally developed to analyse how continuous variables, such as reaction times, 'behave' in response to changes in other variables. A 'standard' linear regression model outputs a set of coefficients that indicate how much the response variable will increase or decrease if we increase an explanatory variable by one unit. For instance, in the case of response times in a reading task, the coefficient for word length might indicate that increasing word length by one letter will result in longer response times.

Logistic regression can be seen as an extension of linear regression capable of analysing categorical data. Since logistic regression deals with categorical variables, we cannot inspect whether a change in one of the explanatory variables will result in an increase (or decrease) in the response variable – the response variable in logistic regression typically has two possible values (but it can have more). To overcome this problem, logistic regression analyses the **odds** with which the values occur, in response to changes in explanatory variables. Odds, just like probability, are a measure of the likelihood that an event will occur. We express probability as the ratio of the number of successes to the total number of events (both successful and not). Odds, on the other hand, are calculated as the ratio of the number of successes to the number of failures[22]. If we consider an example of a deck of cards, the probability of drawing an ace would be 1/13 (or 7.7%). A deck consists of 52 cards and contains 4 aces, so the probability amounts to 4/52 = 1/13. If we were to express the likelihood of this event as odds, the calculation would look as follows: 4:(52-4) = 4:48 = 1:12[23] because we need to subtract the number of successes from the base of the calculation. In terms of a binary choice (or rather a binary outcome), the odds ratio of 1 to 12 would indicate that it is 12 times more likely that a non-ace card will be drawn from the deck.

Similarly to "standard" linear regression, logistic regression also outputs coefficients as a result. The coefficients in logistic regression ultimately indicate how much more (or less) likely the occurrence of one level of the response variable is given the value of the explanatory variable. For instance, the coefficient for inanimate subjects could indicate that if the subject of a sentence is inanimate, it is four times more likely that *się* will occur in this sentence. Unfortunately, the coefficients in logistic regression are not as easy to interpret as those of "standard" linear regression. In logistic regression, odds undergo a logarithmic transformation into logits (log-odds), and the logit scale is not "a very 'natural' scale" (Speelman 2014: 498). The method used to estimate the parameters of logistic regression also causes problems with the interpretation of results; for instance, it is difficult to assess the exact amount of variance explained.

3.2.5.2.2. Procedure

Following the correspondence analysis, the data were analysed by means of binomial[24] logistic regression. The logistic regression analysis was conducted in R (R Core Team 2017), using the package *rms* (Harrell 2015). Before any regression models were built, the categorical variables that contained more than two levels were broken down into binary variables that indicated the presence or absence of a given property. For example, *semantic verb class* contained twelve different levels

[22] In this context, a success means that a desired even occurs, while a failure means that a desired event did not occur.
[23] Both measures of likelihood are written in a different way: probability is expressed as a fraction or a percentage, whereas odds are expressed by a ratio.
[24] The term 'binomial' means that the response variable (i.e. the variable explained by the regression) can assume **two** values.

and, subsequently, it was converted to twelve binary variables that said whether the verb in the example sentence was of a particular semantic class. With all variables prepared in this fashion, the model building process commenced; at first, a "full" model containing all usable variables was built. Subsequently, any variables that missed statistical significance were pruned stepwise until each variable that the model contained was statistically significant. The final model for the sample was parsimonious (i.e. it contained only statistically significant variables) and it had high values of R^2 and C indices as well as a high prediction accuracy.

The best model for the data includes animate subjects, *sam*, volition, emphasis, reflexiva tantum, impersonal situation types, and four semantic verb classes: change, motion, social, and stative verbs. For the sake of brevity, we will discuss only the most important aspects of the model; the full model output can be found in Appendix 7.

3.2.5.2.3. Results and discussion

Let us first examine the model's general properties and goodness-of-fit indicators. The model was tested against a constant only model – that is, a model that always predicts only one of the options (e.g. always predicts *się*). The difference between the constant only model and the model based on the ten variables was statistically significant, indicating that the predictors as a set distinguished between *się* and *siebie* ($\chi 2 = 619.49$, $p < 0.0001$ with 11 degrees of freedom) better than chance. The most important goodness-of-fit statistics to consider are the model's prediction accuracy and the index of concordance (*C*). The model's prediction was high: 82.40% (84.33% for *siebie* and 80.58% for *się*), which means that this model is good at predicting the markers, and it fares much better than chance. The index of concordance takes the value of 0.907, which suggests that the model is a very good fit and its predictions are robust. The model was also more likely to predict *siebie* (515 times) than *się* (485 times). Additionally, to see whether the problematic contexts that included verb forms other than finite and infinitive (discussed in Section 3.2.4.2.) would affect the analysis significantly, the final model was also run for a dataset where those contexts where removed. The results were virtually identical, which means that those problematic contexts do not bear a significant impact on the analysis. The output of the model ran for the dataset where those contexts were removed can be found in Appendix 7.

A logistic regression analysis yields a coefficient for each variable that indicates which option the response variable level is more likely to take (i.e. *się* vs *siebie* in the case of this study) if a given variable level is present. For example, the logistic regression model could predict that *się* will more likely occur in contexts where a reflexiva tantum verb is present. The values of all coefficients along with the marker they predict (i.e. whether they predict *się* or *siebie*) are presented in Table 3.4. The table also presents the *p*-values for each variable; the final logistic regression model contains only statistically significant variables, so all *p*-values fall below 0.05.

PREDICTION	VARIABLE	COEFFICIENT	*p*-value
SIEBIE	*sam*	3.6917	<0.0001
	subject type: animate	1.4870	<0.0001
	semantic class: perception	1.0995	0.0021
	volition	0.8508	<0.0001
	emphasis	0.4919	0.0456
SIĘ	*tantum*	-5.8411	<0.0001
	semantic class: motion	-2.7312	<0.0001
	semantic class: stative	-2.7164	<0.0001
	situation type: impersonal	-1.9289	0.0001
	semantic class: change	-1.5098	<0.0001
	semantic class: social	-0.9551	<0.0001

Table 3.4. Model coefficients

For the purpose of analysis, binomial logistic regression encodes the levels of the response variable numerically – one option is assigned 0, whereas the other option is assigned 1. *R* makes this choice automatically based on the alphabetical order of possible outcomes. Accordingly, *się*[25] was coded as 0, because it precedes *siebie* (coded as 1 in the model) in the alphabetical order. If a coefficient for a variable level obtains a positive value, it means that the model predicts *siebie* to be more likely to occur when this variable level is present, while it predicts *się* to be more

[25]For practical reasons, *się* was coded as *sie*, i.e. without Polish diacritics.

likely to occur if the coefficient for a variable takes a negative value. The magnitude of the coefficient also matters – the higher the absolute[26] value of a coefficient, the more strongly it predicts one or the other level of the response variable. Let us take animate subject types as an example: the coefficient for animate subjects is positive (1.4870), which means that the model predicts *siebie* to be more likely to occur in contexts with animate subjects. Overall, we have five variables that predict *siebie* (*sam*, animate subjects, perception verbs, volition, and emphasis), out of which *sam* was the strongest predictor – its coefficient value was 3.6917. In the case of *się*, we have six variables that predict this marker (reflexiva tantum, motion, stative, change and social verbs, and impersonal situation types), and the single strongest predictor was the presence of a reflexiva tantum verb – its coefficient value was -5.8411.

The regression analysis provided evidence in support of most of the predictions discussed in Section 3.2.2 (Table 3.1., repeated here for convenience).

properties predicted to co-occur with *się*	properties predicted to co-occur with *siebie*
reflexiva tantum	*sam*
reciprocal and impersonal situation types	emphasis
preverbal clause position	co-ordination
overt subjects	volitional actions
verbs of motion and bodily care	verbs of communication, perception, and cognition
	animate subjects

Table 3.1. Properties of contexts predicted to co-occur with *się* or *siebie.*

Siebie is more likely to occur with verbs of perception and in contexts containing emphasis, volitional actions, animate subjects, or the word *sam*. As expected, *się* correlates with the tantum variable, which is a very strong predictor (the coefficient is above 5) – this implies that whether a verb is a reflexiva tantum or not very strongly predicts *się* against *siebie.* As far as the verb classes predicting *się* mentioned in Section 3.2.2. are concerned, the sample did not contain enough data to properly investigate whether bodily care verbs correlate with *się,* because bodily

[26] The absolute value means the distance of a given value from zero. For instance, the absolute value of -2.56 would be 2.56.

care verbs appeared only three times in the dataset. Motion verbs, however, do predict the occurrence of *się*. Apart from providing evidence in favour of the predictions, the regression model indicated that three more verb classes predict *się*: stative, change, and social verbs. Finally, impersonal situation types also predict *się*.

Siebie mostly correlates with general variables, such as animate subjects, emphasis (also indicated by *sam*), or volition. The regression model suggests that *siebie* has a clear context it occurs in: volitional actions performed by animate subjects on themselves or actions in which the involvement of the subject is emphasised. This type of context is relatively general – it is not limited to a few particular actions (e.g. seeing or hearing). *Siebie*'s correlation with general variables, such as volition or emphasis, suggests that it does not rely on the low-level semantic or lexical characteristics of particular verbs. Since *siebie* does not rely on low-level verbal semantics and occurs in relatively general contexts, it is possible that speakers have one general category for *siebie* as a whole.

Się's status appears to be different from that of *siebie*. First and foremost, it is most strongly predicted by the *tantum* variable – what is more, *tantum* constitutes the single strongest predictor in the model. Such a result suggests that *się* can be in large part predicted by the lexical properties of some reflexive verbs (that is, the reflexive tantum verbs). It also means that some *się*+verb combinations rely on specific (lexical) constructions. Apart from being part of a reflexiva tantum verb, *się* can be predicted by the semantic class of the verbs it appears with – the model suggests that *się* correlates with motion, change, stative, and social verbs. The correlation of *się* and social verbs suggests that *się* might appear in contexts where more than one party is involved in the same action (where all the parties are both agents and patients of this action), for instance, meeting one another (*spotkać się*) or sharing things with one another (*dzielić się*) – we could say that these are actions of the RECIPROCAL type. Change and stative verbs imply situations without volition, because they usually occur "autonomously" or as a result of other processes – in other words, the events conveyed by the change and stative verb classes could be INCHOATIVE/RESULTATIVE or PASSIVE. Some researchers (e.g. Wilczewska 1966; Niedzielski 1976; Frajzyngier 2000) have postulated that *się* can carry RECIPROCAL, INCHOATIVE/RESULTATIVE or PASSIVE meanings – social, change, and stative verb classes correlating with *się* are an indication that native speakers could build categories for those senses postulated for *się* from the exposure to linguistic input. The only more general variable that predicts *się* is the impersonal situation type. The impersonal situation type variable indicates that the context contains a situation typical for the IMPERSONAL sense of *się* postulated by some researchers – an event or action instigated by a maximally diffused agent.

The results of the logistic regression model suggest that *się* might be better described as a "bundle" of different (less general) senses or functions – "TRUE" REFLEXIVE, RECIPROCAL, INCHOATIVE/RESULTATIVE, PASSIVE, and IMPERSONAL – than one general construction. To put it differently, the results of the logistic regression model do not provide evidence supporting the hypothesis that native speakers could build a general construction for *się* as a whole, however, it does provide some evidence suggesting that speakers might build less general constructions for different senses of *się*.

3.2.5.3. Conditional inference trees and random forests

Finally, the data were analysed with conditional inference trees and random forests. Conditional inference trees are also a confirmatory statistical technique, but they differ from logistic regression quite significantly in the way they analyse data. In regression, the model considers all the variables "at once" and tries to fit a regression equation. In contrast, conditional inference trees only consider one variable at a time – they split the dataset into many binary subsets until all observations are explained, and no more splits can be made. The final outcome is a decision tree that can predict, for instance, whether the example will contain *się* or *siebie* based on the values the variables take. The difference between regression and conditional inference trees is that the regression analysis tells us about the impact of each variable on the final outcome, while conditional inference trees can provide us with the most likely combinations of variables for which one or the other marker would occur. If the results of the logistic regression analysis and the conditional inference tree analysis converge, it would mean that the conclusions we can draw from the data are very robust and they remain identical (or similar) regardless of the method used to analyse the data. The next subsection (3.2.5.3.1) will introduce the method, while Section 3.2.5.3.3. will present the results of the conducted analysis.

3.2.5.3.1. Method

Conditional inference trees can predict which outcome of a response variable is more likely, e.g. *się* or *siebie,* given a set of predictors (Tagliamonte & Baayen 2012: 22). During the estimation of a conditional inference tree, the model goes through the data and seeks to establish which variables constitute useful predictors. The model considers only one variable at a given time and attempts to split the data into two binary subsets according to the values of this variable. To see whether such a split is statistically significant, the model runs a test of independence. If the predicted value (e.g. *się* vs *siebie*) is statistically independent from the variable under consideration, the model rejects the variable and proceeds to another variable. If, however, the predicted variable is statistically dependent on the variable under consideration, the model labels it as a useful predictor. The model repeats the procedure with the remaining predictors, and if there are more than two useful predictors, the model chooses the one whose association with the predicted value is the strongest. After it has chosen the predictor, the model splits the dataset into two subsets according to the values of the predictor[27]. The algorithm then works recursively through the data (i.e. the procedure is repeated) until no more justified splits can be made. The final outcome of the model is an inference tree, which visualises the choices made by the algorithm. An important feature of conditional inference trees is that once the algorithm makes a decision, it cannot go back and reconsider the choice.

A conditional inference tree constitutes a good visualisation of multivariate data. One tree, however, can overfit the data and thus describe some part of the random noise in the data as meaningful patterns (Divjak 2015: 61). To mitigate this problem, we can "grow" a random forest of conditional inference trees. A random

[27] If a given predictor can take more than two values, the algorithm puts the values into two bins, in order to enable the model to make a binary split.

forest creates multiple trees from the same dataset by permuting the data – the model selects random subsets of the data and grows trees for each of them. Once the algorithm has grown many trees, it averages over all decisions made in all trees. On the basis of a random forest, we can eventually compute the relative importance of all variables used in the estimation of the model.

3.2.5.3.2. Procedure

The conditional inference trees and forests analysis was conducted in R (R Core Team 2017), using the functions *ctree* (Hothorn, Hornik & Zeileis 2006) and *cforest* (Strobl et al. 2008) from the package *party*. The first step in the procedure was to compute a single conditional inference tree; subsequently, a random forest was generated, and the variable importance within the forest was computed. The first version of the model included all variables available in the tagged dataset. A large number of NA values in two variables: *tense* and *number and grammatical person* prevented the forest model from computing the relative importance of variables. The final model was run on all variables except *tense* and *number and grammatical person*. The following section (3.2.5.3.3) will present the results of the analysis.

3.2.5.3.3. Results and discussion

As mentioned before, conditional inference trees predict which outcome of the response variable is more probable, given a set of predictors. Conditional inference trees arrive at a prediction by making multiple binary decisions until no more decisions remain that would be justified by the data. The decisions the conditional inference tree algorithm makes are visualised as binary splits on the plot (i.e. two forking branches). The plot of the conditional inference tree generated from the corpus data is presented in Figure 3.5..

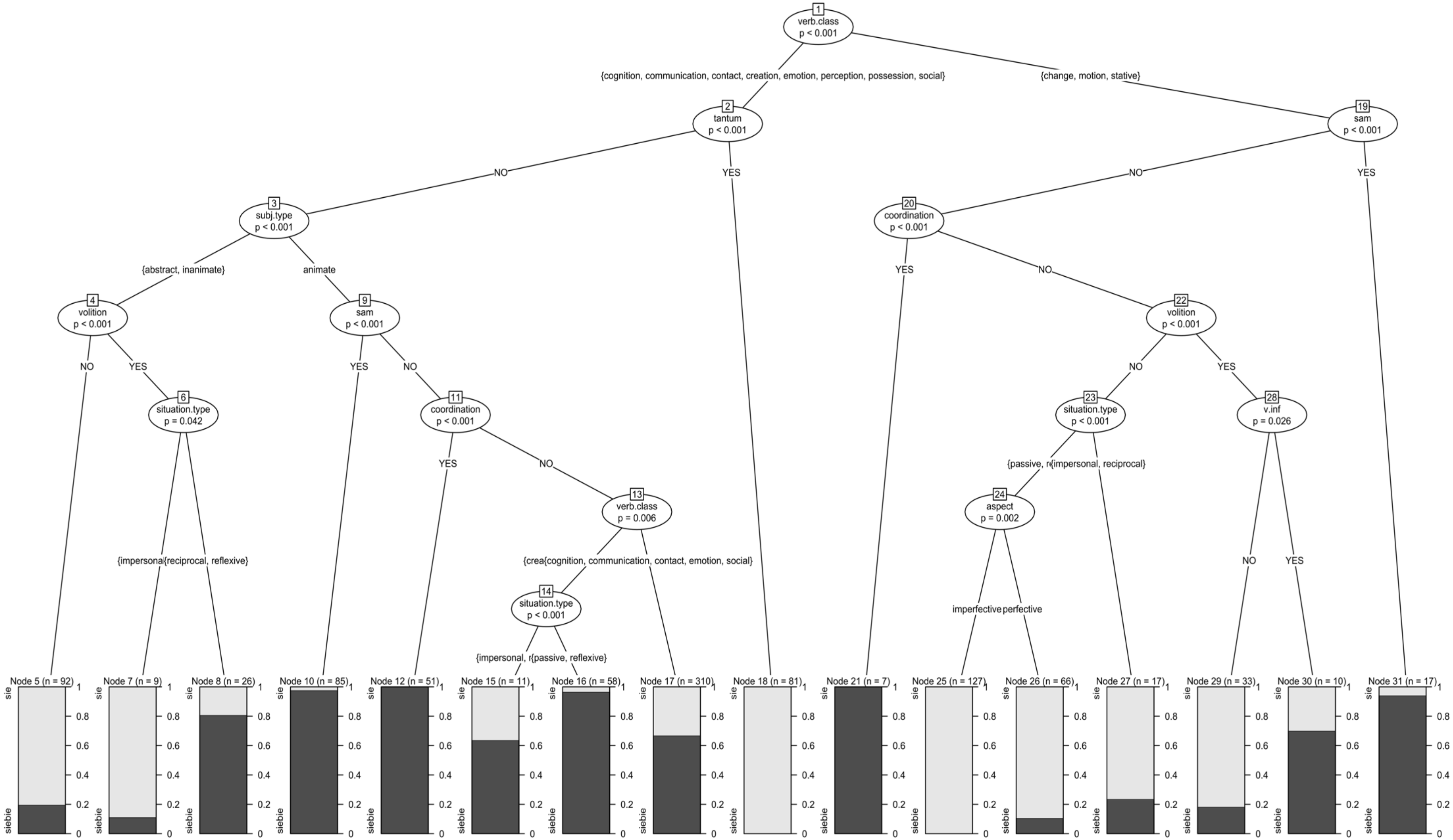


Figure 3.5. Conditional inference tree for the corpus data

Each node of the tree (presented as an ellipse), e.g. Node 1, stands for a variable according to which a split in the dataset was made; inside the ellipse, we can see the name of the variable as well as the *p*-value of this variable. The lines that proceed from one node to two other nodes denote the variable levels included in the binary split. For instance, Node 3 splits into abstract and inanimate subjects in one branch and animate subjects in the other. Finally, the terminal nodes at the bottom of the plot, e.g. Node 5, "show a bar plot of the output label distribution considering only the observations at each respective leaf, and denote with *n* the number of observations that were assigned to that leaf" (Sardá-Espinosa, Subbiah & Bartz-Beielstein 2017: 31). In the case of this analysis, the bars which indicate the ratio of *się* to *siebie* in the final split for a particular "path" from the top node to the leaf node – the black part of the bar indicates the relative frequency of *siebie*, whereas the grey part indicates the relative frequency of *się*. When summed up, all leaf nodes jointly account for the whole dataset.

The first decision the algorithm made for the data (Figure 3.5.) was to split the dataset according to the class of verb. We can see two branches in the initial split [node 1]: change, motion and stative verbs in the right branch and all the remaining classes (cognition, creation, communication, contact, emotion, perception, possession, and social) in the left branch. The change, stative and motion verbs were further split by *sam* [Node 19]. If *sam* was present, the model predicted almost exclusively *siebie* [Node 31]. If *sam* was absent, the branch split again, into contexts with and without grammatical co-ordination [Node 20]. For the contexts with coordination, only *siebie* was predicted [Node 21]. The contexts with co-ordination were further split according to whether the action was volitional or not [node 18]. The contexts with volitional actions were split into contexts containing a VERB+INFINITIVE construction (or not) [Node 28]; the contexts with the V+INF construction predicted mainly *siebie,* while contexts without the V+INF construction predicted mainly *się*. Non-volitional actions were further split into two groups, depending on the situation type [Node 23]: impersonal and reciprocal situations in one group and reflexive-type and passive in the other. The impersonal and reciprocal group did not split any more, and it mostly predicted *się* (about 90%). Finally, the passive and reflexive-type group was split into two more groups based on aspect. The imperfective aspect branch contained exclusively *się,* while the perfective aspect branch contained some *siebie*, but it also mainly consisted of *się* (about 90%).

The left branch – i.e. cognition, creation, communication, contact, emotion, perception, possession, and social verbs – was first split by reflexiva tantum [Node 2]. If the verb was a reflexiva tantum verb, the model predicted exclusively *się*. The non-reflexiva tantum branch first split according to the type of subjects [Node 3]. For inanimate and abstract subjects, one more split occurred – according to whether the action was volitional or not [Node 4][28]. In the case of non-volitional actions, the model predicted *się* in ~80% of observations. Volitional actions were further split by situation types: impersonal and passive in one branch (predicting mostly *się*; about 90%), reciprocal and reflexive in the other (predicting mostly *siebie*; about 80%). The branch with animate subjects contains a few more splits, but if we look at prediction bars in the bottom part of the tree, we can see that the model predicted *siebie* in an overwhelming majority of cases (see Leaf Nodes 10, 12, 15, 16, 17). As in the case of the rightmost branch [Node 19 and further down], contexts with co-ordination and *sam* almost exclusively predict *siebie.* The picture for the contexts with neither co-ordination nor *sam* is somewhat more complicated – such contexts are further split by the verb class [Node 13]. With cognition, communication, contact, emotion, and social verbs, the ratio of *siebie* to *się* is approximately 70 to 30 [Node 17]. Creation, perception, and possession verbs are further split according to situation type: impersonal and reciprocal situations predict *siebie* in more than 60% of cases, while passive and reflexive situations predict *siebie* in more than 95% of cases.

The fact that the splits according to verb class and reflexiva tantum were located high in the tree indicates that these variables are important for the model. The

[28] One might wonder how inanimate or abstract agents can perform volitional actions. Nouns denoting institutions were coded as abstract in this study, and institutions are capable of instigating actions that would be interpreted as volitional. The following sentence involves an abstract (i.e. institutional) agent and an action that ought to be seen as volitional: *…władze Rudy … nie* ***zgodziły*** *się na takie rozwiązanie* [‘Ruda’s town council did not accept this solution’]. Moreover, the agents for impersonal sentences were also coded as abstract. Technically, impersonal sentences in Polish are subjectless, and we can see the agent of the action/process encoded in an impersonal sentence as maximally diffuse and thus abstract.

tantum variable says directly that a reflexive verb is a specific (lexical) construction. Reflexiva tantum verbs do not have a ’bare’ counterpart, for instance, the verb *bać się* (’be afraid’) does not have a non-*się* counterpart *bać*. In consequence, in *tantum* verbs, arriving at the verb + *się* pairing by means of combining a ’bare’ verb and *się* is impossible, because the ’bare’ verb does not exist. Such verbs must be stored and produced by speakers as specific (lexical) constructions. The (*semantic*) *verb class* variable reflects low-level semantic properties, which also suggests that (at least some) reflexive verbs might be specific low-level constructions, because they depend on low-level (lexical) semantics. Some verbs in the main left branch (the one split by reflexiva tantum) might be heavily lexicalised with *się* – a number of emotion and cognition verbs are reflexiva tantum in Polish, for instance, *bać się* ’be afraid’, *wstydzić się* ’be ashamed of’ or *wahać się* ’dither’.

As a means of summary, let us now inspect the branches (all the way down to the leaf nodes) in which more than 80% of predicted cases were either *się* or *siebie*. We could think of these branches as contexts that contain particular properties – by inspecting the branches that contain a large proportion of one or the other marker, we will be able to establish the “typical” contexts for them. Table 3.5. presents only the branches for which one marker constituted more than 80% of predicted cases.

MARKER	CONTEXT PROPERTIES	CASES
SIĘ	cognition/communication/contact/creation/emotion/perception/possession/social verbs + not *tantum* verb + abstract/inanimate subjects + no volition	92
	cognition/communication/contact/creation/emotion/perception/possession/social verbs + not *tantum* verb + abstract/inanimate subjects + volition + impersonal/passive situation type	9
	cognition/communication/contact/creation/emotion/perception/possession/social verbs + *tantum* verb	81
	change/motion/stative verbs + no *sam* + no coordination + no volition + passive/reflexive situation type + imperfective aspect	127
	change/motion/stative verbs + no *sam* + no coordination + no volition + passive/reflexive situation type + perfective aspect	66
	change/motion/stative verbs + no *sam* + no coordination + volition + no *verb+INF* construction	33
SIEBIE	cognition/communication/contact/creation/emotion/perception/possession/social verbs + not *tantum* verb + animate subjects + *sam*	85
	cognition/communication/contact/creation/emotion/perception/possession/social verbs + not *tantum* verb + animate subjects + no *sam* + coordination	51
	cognition/communication/contact/creation/emotion/perception/possession/social verbs + not *tantum* verb + animate subjects + no *sam* + no coordination + creation/perception/possession verbs + passive/reflexive situation type	58
	cognition/communication/contact/creation/emotion/perception/possession/social verbs + not *tantum* verb + abstract/inanimate subjects + volition + reciprocal/reflexive situation type	92
	change/motion/stative verbs + no *sam* + coordination	7
	change/motion/stative verbs + *sam*	17

Table 3.5. Branches in the conditional inference tree with more than 80% of predicted cases predicting one marker

Overall, the picture seems quite clear: if an action was performed volitionally by an animate subject, *siebie* is much more likely to be found as the marker, unless it is “overridden” by the verb being a reflexiva tantum. *Siebie* also appears in contexts where *sam* is present, which indicates that emphasis might play a role as well. *Się,* on the other hand, seems to be associated with change, motion and stative verbs, and reflexiva tantum verbs. Change and stative verbs might indicate low volition – one can argue that change is a process that often occurs independently of people’s actions, or at least it is conceptualised as such. States, on the other hand, are not even processes – they are properties. Being in a state indicates no volition on the part of the entity that remains in this state. Let's take the verb *nazywać się* ('have a name') as an example. Having a name is a property rather than a volitional action. We could volitionally **apply for a new name**, but having the new name will again be a non-volitional state. In contrast to change and stative verbs, motion verbs cannot be said to carry little volition as class – verbs such as *wspinać się* 'climb' or *położyć się* 'lie down' usually denote volitional actions. Consequently, the lack of volition cannot be a defining characteristic of *się* as a whole. The correlation of *się* with particular classes of verbs (motion, change, and stative) is rather an indication that *się* relies on properties of lower generality, such as the type of verb. Additionally, *się* appears whenever it is a lexical prerequisite (reflexiva tantum verbs), which suggests that *się* could be considered as a lexical phenomenon, at least to some extent. This result – *się* associated with certain verb classes and tantum verbs – provides more support for the hypothesis that *się* relies on constructions of lower generality.

According to Divjak (2015: 61) "[a] single tree is likely to overfit the data", i.e. a model based on a single tree might interpret some random noise as a significant relationship in the data. To alleviate this problem, the data were further analysed with random forests – a procedure which generates multiple conditional inference trees by permuting the data from a given sample. Random forests also allow for calculating the variable importance, i.e. we can see which predictors were the most and the least important in the estimation of the model. A variable importance plot is presented in Figure 3.6..

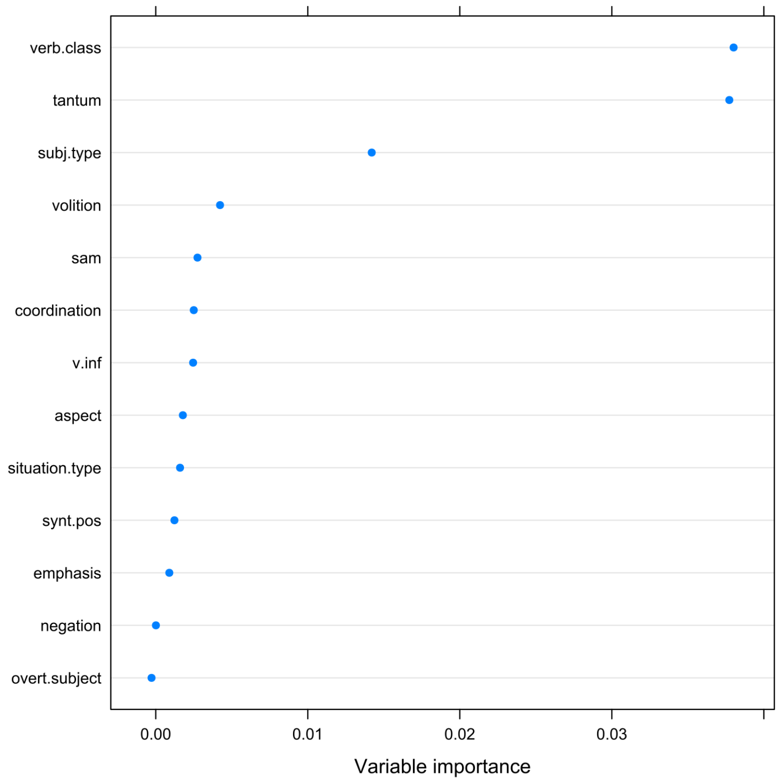


Figure 3.6. Variable importance plot for the corpus sample

The variable importance plot suggests that verb class (variable importance = 0.038), reflexiva tantum (variable importance = 0.0377), and subject type (variable importance = 0.0142) were the three most important predictors in the dataset. Overall, the model with all the variables presented in the variable importance plot predicted the correct marker in 86.3% of cases, which is a very high accuracy (0.863). Hence, the model is quite robust and predicts much better than chance[29].

The most important variable was verb class, which, in combination with reflexiva tantum being the second most important variable, suggests that the lexical semantics of the verb can explain a large part of the variation. This result corroborates the findings from the tree in Figure 3.5, where the initial split was made in accordance with the verb class, and reflexiva tantum formed another important split, because it further divided the verb class variable. The subject type itself came second, and it is the only high-level (or general) variable of relative importance in the trees and forests model. This outcome seems to corroborate the

[29] The chance level is 0.5 in this case, because the choice the model needs to make is binary (*się* vs *siebie*), and the dataset is balanced, that is, *się* and *siebie* have an equal number of examples.

results of the logistic regression model in which *siebie* was strongly associated with animate subjects.

The remaining variables (volition, *sam*, co-ordination, V+INF constructions, aspect, situation type, syntactic position, emphasis, negation, and overt subject) did not have a substantial impact on the prediction accuracy, with volition being the most important out of the less influential variables. This means that these variables are not crucial for the estimation of the model and perform more of a "fine-tuning" (Divjak 2015: 62) function – they improve the model's prediction only by a small fraction. A model including only three variables (verb class, tantum, and subject type) was run in order to compare prediction accuracies. The three-variable model predicted the marker correctly in 80.7% cases (0.807 accuracy), so leaving out ten out of thirteen variables in the model decreased the prediction accuracy only by 0.056 in comparison to the model with thirteen variables, whose accuracy was 0.863.

If we consider the conditional inference tree output (Figure 3.5.) and the variable importance calculations from the forest analysis, the results appear to corroborate the findings of the correspondence analysis and the logistic regression model. *Siebie* correlates with relatively general properties: animate subjects, *sam* (which could be an indicator of emphasis), and volition. Consequently, it seems likely that speakers of Polish build a coherent usage-based category for this marker as a whole. *Się*, in contrast, is predominantly predicted by low-level properties (being a lexical prerequisite in reflexiva tantum verbs), which might suggest that some *się* verbs might rely purely on low-level (lexical) semantics. The tree also indicated that *się* might be strongly associated with change, stative, and motion verbs. The first two classes (change and stative) indicate that *się* could carry PASSIVE and INCHOATIVE/RESULTATIVE meanings. The third class, motion, might be an emanation of the TRUE REFLEXIVE sense of *się*. Motion is typically instigated by the same entity that moves – it is usually an action performed by the subject on the subject, which is the definition of a true reflexive event (see Section 2.2.2.1.1.). Overall, since *się* correlates with low-level lexical properties or verb semantics, it is quite unlikely that speakers could build one general construction for *się* as a whole. It is more likely that speakers build less general constructions for the different senses of *się* (e.g. PASSIVE or INCHOATIVE/RESULTATIVE), and, additionally, some reflexive verbs (reflexiva tantum verbs) rely exclusively on low-level lexical semantics.

3.3. Interim conclusions

The primary purpose of the corpus study presented in this chapter was to investigate whether the language input speakers are exposed to could enable them to build general categories for reflexive-type markers *się* and *siebie.* The results of exploratory correspondence analysis suggest that there is much variation in the data that cannot be easily accounted for by only two general categories. The analysis of the full set of variables revealed some properties associated with *siebie:* animate subjects, contexts with volitional actions, contexts with emphasis, and verbs of communication. These properties were predicted to co-occur with *siebie*, based on the literature on Polish reflexive-type markers (see Section 3.2.1.). In the case of *się*, the correspondence analysis algorithm could not find any variable levels that would correlate with the marker. The situation did not change when the variables were split into form- and meaning-related sets. Such a result – some variables correlating with *siebie* and hardly any variables correlating with *się* – might be an indication that speakers could build a general category for *siebie*, while the same would not be very likely for *się*.

Correspondence analysis, being an exploratory statistical technique, cannot provide robust evidence in favour or against a hypothesis, and thus two confirmatory (or explanatory) analyses were conducted subsequently: a logistic regression analysis and conditional inference trees. The results of both conditional inference trees and regression converge and suggest that *siebie* is likely to appear in contexts where the word *sam* and animate subjects are present. In Polish (and some other Slavonic languages), speakers use the word *sam* to indicate and **emphasise** that the action was performed independently by the subject of the sentence and that there was no need for external help. Consequently, *siebie* seems to appear in contexts involving volitional actions and some degree of emphasis. The properties that predict *siebie*: animate subjects, volition, and emphasis are quite general and independent from the semantics of individual verbs (or classes of verbs); it is, therefore, likely that native speakers of Polish could build one general construction for *siebie* as a whole.

When it comes to *się*, the results of logistic regression and conditional inference trees and forests also converged. The results of both techniques suggest that native speakers of Polish might not be able to build one general category for *się* as a whole. *Tantum* was the single most important predictor for *się* in the logistic regression and one of the two strongest predictors in the trees and forests analysis. This implies that *się* might rely on the semantics of a given verb, and at least some verbs might be lexicalised with *się*, i.e. some verb + *się* combinations would form lower-level (lexical) categories in the minds of native speakers of Polish. The meaning of some verb + *się* combinations, namely reflexiva tantum verbs, could not be arrived at compositionally by means of taking a bare verb and *się*. The other important variable that explained the behaviour of *się* was the semantic class of the verb – the statistical models indicated that motion, change, stative, and reciprocal verbs correlate strongly with *się*. Change and stative verbs denote events that occur "spontaneously" or as a result of other processes. In other words, these verbs classes denote events of INCHOATIVE/RESULTATIVE or PASSIVE nature. The social verb class correlating with *się* can indicate that it might also have a propensity to appear in RECIPROCAL contexts, because social situations usually involve more than one party doing the same action as the other parties (e.g. meeting one another or sharing with one another). Lastly, motion verbs could be a representation of the TRUE REFLEXIVE sense. These kinds of events (PASSIVE, INCHOATIVE/RESULTATIVE, RECIPROCAL, and TRUE REFLEXIVE) correspond to some of the functions that researchers postulated for *się* (see Section 2.2.1.). This result might indicate that *się* could be better described as a bundle of different less general categories rather than one maximally general category.

Summing up, the corpus study has yielded evidence supporting the hypothesis that speakers of Polish could build one general category for *siebie* as a whole. On the other hand, the results obtained from the corpus study did not support the hypothesis that speakers could build one general category for *się*. Instead, speakers might build a few less general categories for *się,* each with a different meaning. Additionally, some verb + *się* combinations might be fully reliant on low-level (lexical) verbal semantics. Some researchers (see e.g. Wilczewska 1966; Niedzielski 1976) have already postulated that *się* could have a number of different senses, but this hypothesis has never been tested empirically. To bridge this gap, an experimental study was conducted to investigate whether native speakers of Polish could build constructions for the different senses of *się* – the results of the study will be presented in Chapter 4.

Chapter 4: Reflexives experimental study

4.1. Introduction

The results of the corpus study presented in the previous chapter suggested that speakers might not build a single general construction for *się* based on the input they receive. The study showed that the behaviour of *się* can be predicted in large part by low-level lexical properties – that is, whether a verb is already lexicalised with the marker. Apart from that, there was also some indication that native speakers could formulate (less) general constructions for the different senses/functions of *się*, such as the IMPERSONAL sense or the INCHOATIVE/RESULTATIVE sense.

If *się* can indeed express multiple senses, we would expect native speakers to build a construction for each sense of the marker. Following that, if a postulated construction exists in the minds of speakers, they ought to be able to classify different verbs carrying that construction as members of the same category. In other words, speakers should see those verbs as similar. For instance, they should categorise the two TRUE REFLEXIVE verbs *myć się* ('wash') and *ubierać się* ('get dressed') as belonging to the same category. If that is the case, and speakers do perceive similarities between different verbs with the same sense of *się*, we could surmise that they might be able to build a general category for this sense. Conversely, if speakers do not perceive similarities between different verbs with the same sense of *się*, we should take it as an indication that they might not have a category for this sense.

Some authors, including e.g. Wilczewska (1966) or Kubiński (1982), suggested that *się* can take multiple different senses, but to my best knowledge, the existence of these senses in speakers has not yet been investigated empirically – this study seeks to bridge this gap. To investigate experimentally whether speakers build and use these senses, we could present speakers with tokens of a construction and ask them whether they perceive those tokens as similar. An experimental paradigm exists that allows performing such a study: sentence-sorting. This chapter will discuss the results of a sentence-sorting experiment investigating whether native speakers of Polish could build constructions for the different senses of *się*.

4.2. Are verbs with the same sense of *się* similar? A sentence-sorting experiment

In a nutshell, in a sentence-sorting experiment, subjects are presented with a number of sentences (containing constructions under investigation) and asked to group sentences into a number of bins. The results of the experiment are subsequently analysed to see how participants grouped the sentences. If sentences with the same construction were put into the same bin, we could surmise that participants perceived the sentences as similar and, consequently, that they might have a category in their minds for the proposed construction. For this reason, sentence-sorting was used as experimental paradigm to investigate whether speakers of Polish really use the general constructions for the different senses of *się.* The primary hypothesis of the experiment is that if participants reliably group sentences according to the senses of *się*, they might have constructions for those senses, because they perceive sentences with verbs with the same sense of *się* as similar. In the opposite case, if participants group the sentences according to other criteria, they might not have built constructions for the senses of *się*.

In this experiment, participants were given sentences with five different senses of *się* postulated in the literature ("TRUE" REFLEXIVE, RECIPROCAL, INCHOATIVE/RESULTATIVE, PASSIVE, and IMPERSONAL); the participants were then asked to group these sentences into five bins. The primary hypothesis is that if participants have categories for the different senses of *się,* they will reliably group experimental sentences according to these senses. The following section will present a more comprehensive overview of the sentence-sorting method.

4.2.1. Sentence-sorting experiments on Polish *się*: theory and construction

Sorting is a linguistic experiment in which participants are presented with a number of stimuli which they have to sort according to how similar they find them. The central theoretical assumption of sorting experiments is that the way participants sort stimuli will reflect the mental categories they hold in their minds. In usage-based linguistics, sentence sorting has been used successfully in polysemy and synonymy research (Sandra & Rice 1995; Divjak & Gries 2008) and research on L1 syntax (Bencini & Goldberg 2000) and L2 syntax (Gries & Wulff 2009), while sorting experiments sensu largo have achieved great popularity in psychology, sociology, and anthropology (Coxon 1999: 1–2).

Sorting experiments are constructed in a simple way. First, participants are given sentences, words, or phrases. The participants are asked to divide the stimuli into groups containing stimuli that the participants find similar. The number of groups is usually predefined by the researcher, depending on the study's objectives; one can also either give similarity criteria to participants or leave the choice entirely to their discretion. A large number of available experimental software packages allows sorting experiments to be administered electronically and via the Internet.

According to (Sandra & Rice 1995) "distinctions made in a sorting experiment are assumed to reflect in a relatively straightforward way the distinctions that language users have learnt to make in the course of language acquisition". This would mean that sorting experiments directly tap into the categorisation systems of language users. One caveat is, however, that in off-line sentence-sorting experiments, users are given ample time for deliberation during the experiment. Consequently, they can analyse the stimuli in much detail and use their **metalinguistic** knowledge to complete the task. What follows is that sorting experiments differ substantially from online experiments, such as self-paced reading or eye-tracking experiments, in terms of the language faculties into which they can give insights. The latter can shed some light onto language as it unfolds in the moment, whereas the former can give us some idea about how speakers use language when faced with more cognitively demanding tasks and how general linguistic categories can be. That notwithstanding, when investigating whether speakers can build a general construction postulated by linguists, a sentence-sorting task can tell us if the speakers possess a category for this construction **at all**. In other words, a sentence-sorting task can tell us whether any possibility exists that speakers have a category for a given construction, even if using this category would require more time and deliberation than what is usually needed in online speech comprehension or production.

4.2.2. Procedure

In a sentence-sorting task, the respondents were asked to group 15 or 25 sentences (depending on the version) into five bins. The number of bins corresponded to the number of different senses of *się* that were considered in this experiment: *true* reflexives, *passive* reflexives, *inchoatives/resultatives, impersonals*, and *reciprocals.* Each sentence contained a verb and the marker *się* carrying one of the five senses. Participants were not provided with any criteria that would influence their decisions, so they could sort the sentences in any way they saw fit; they were only asked to "put sentences in groups according to their similarity". The bins had no suggestive labels – they were labelled *Grupa 1, Grupa 2* ['Group 1', 'Group 2'], and so on. No limit was imposed as to how many sentences a participant could put in one group, as long as they put at least one sentence in each of them. Participants could take as much time as they needed to complete the questionnaire, and they could change their choices an unlimited number of times – the questionnaire was considered complete only when a participant pressed the "submit" button. Full task instructions can be found in Appendix 1A.

Participants were also requested to provide answers to a number of questions related to demographics, education, and reading habits – not all of them were mandatory so as not to make participants wary of disclosing too much personal data and thus not completing the questionnaire.

The study obtained ethics approval by The University of Sheffield. The questionnaires were delivered electronically via the Qualtrics[30] platform and distributed through an anonymous link over the period from 23 August 2017 to 1 September 2017. The system randomly assigned one version of the experiment to each participant with a 50-per cent probability.

4.2.3. Stimuli

Two versions of the experiment were created: one with three sentences per meaning (15 in total), the other one with five sentences per meaning (25 in total). The aim behind that decision was to see whether the number of sentences in the experiment

[30] http://www.qualtrics.com

would affect participants' ability to perceive different meanings of *się*. All stimuli can be found in Appendix 2.

The meanings of *się* considered in this experiment are as follows: REFLEXIVE, RECIPROCAL, PASSIVE, IMPERSONAL, and INCHOATIVE/RESULTATIVE. All meanings were described in Section 2.2.1, but, for the sake of convenience, a short overview will be presented below. In the REFLEXIVE meaning, *się*-verbs (usually) have their non-reflexive base counterparts, and adding *się* to such base verbs does not alter their semantics significantly and only serves as a device that indicates that the subject also happens to be the object of a transitive verb. An English example of the REFLEXIVE sense would be the sentence *She wrapped herself in a blanket* – you usually wrapped something or somebody else, but in this situation, the subject happened to wrap herself (in a blanket). In verbs with the RECIPROCAL meaning of *się,* each of the two (or more) entities in a sentence performs the role of both the subject and the object of an action. A good example from English would be hating or loving, e.g. *John and Mary hated/loved* ***each other***. We could call the two above meanings – the REFLEXIVE and RECIPROCAL meanings of *się* – agentive, because actions expressed by sentences with these meanings always construe an agent that actively performs the action. The agentive meanings of *się* stand in contrast to non-agentive meanings, where the agent is demoted in some manner. In the PASSIVE meaning of *się*, the action implies an agent, but, similarly to the prototypical passive constructions, the agent remains outside the scope of the conceptualisation. In the IMPERSONAL meaning, the agent is maximally diffuse – we know that the action described in the sentence must have an agent, but the sentence speaks about general habits, customs, or trends, so the role of the agent can be fulfilled by anyone (or anything). The rough English equivalent of the IMPERSONAL meaning of *się* could be the pronoun *one* as in *One does not talk about that.* Finally, the INCHOATIVE/RESULTATIVE meaning of *się* implies that the action described by the sentence happened spontaneously, that is, without any involvement of the third party or even the object of the action, e.g. in the sentence *The door opened*.

The verbs that were included in the experimental sentences were taken from Wilczewska's (1966) study, as this is the most comprehensive study of *się* in Polish linguistics to date. All sentences consisted of a single clause. Apart from the single-clause criterion, the sentences were not standardised in any other manner, and they could contain different types of verbs, subjects, objects, adverbials, et cetera; the verbs in the sentences were marked for different tenses, numbers, and genders. Not controlling for other grammatical parameters was a principled choice that aimed to give participants alternative criteria for sorting apart from the different senses of *się* and to make the contexts more naturalistic. Had the sentences been too uniformly structured with regards to their formal properties, the participants might have been compelled to sort the sentences according to the different senses of *się* due to the lack of other possible criteria for sorting. With many possible criteria to choose between, if participants sorted the sentences according to the senses of *się,* it would mean that the categories for those senses are strong enough in participants' minds to be chosen over the other criteria.

4.2.4. Participants

Participants were recruited on the official Facebook groups for the University of Warsaw, Warsaw School of Economics, and the University of Gdansk. Participants were also encouraged to share the link to the study with their peers.

Overall, tallied over all conditions, 273 respondents took part in the study (197 females; 72.16%). An overwhelming majority of participants were either in university education (39.93%), had graduated from a university (50.18%), or had taken a university course but had not graduated (5.49%) – the overall proportion of respondents with at least some tertiary education amounted to 95.60%. Participants' mean age was 25.57 years (median = 25, sd = 5.68, min = 18, max = 60).

4.3. Data and data analysis

The sample sizes of either experiment did not differ greatly from one another: 149 responses were submitted for the shorter condition (i.e. three sentences per meaning) and 124 responses for the longer condition (i.e. five sentences per meaning); as a result, the samples were taken as is and no data were removed. The data were analysed by means of hierarchical agglomerative cluster analysis, because this technique allows for investigating whether any regularities exist in the way respondents sorted the experimental sentences. As a result, the samples remained untrimmed, and the analysis was performed on a full dataset in each case. The following section (4.3.1) will present an overview of hierarchical agglomerative clustering as a statistical method.

4.3.1. Data analysis method: hierarchical agglomerative clustering

Hierarchical clustering can be divided into two main types: **agglomerative** clustering and **divisive** clustering, which are direct opposites of each other. Agglomerative clustering algorithms proceed from the bottom up "by a series of successive fusions of the *n* individuals into groups" (Everitt et al. 2011: 72). In other words, the algorithm starts with creating a small cluster (or clusters) of data points that have the smallest distance between them, which it subsequently merges into bigger clusters until it ends up with one large cluster encompassing all the data in the dataset. Graphically, the results are eventually plotted as a dendrogram, that is a tree diagram that visualises all clusters and how and where they have been merged with one another. By means of an illustration, let us inspect one of the dendrograms from Divjak & Fieller (2014):

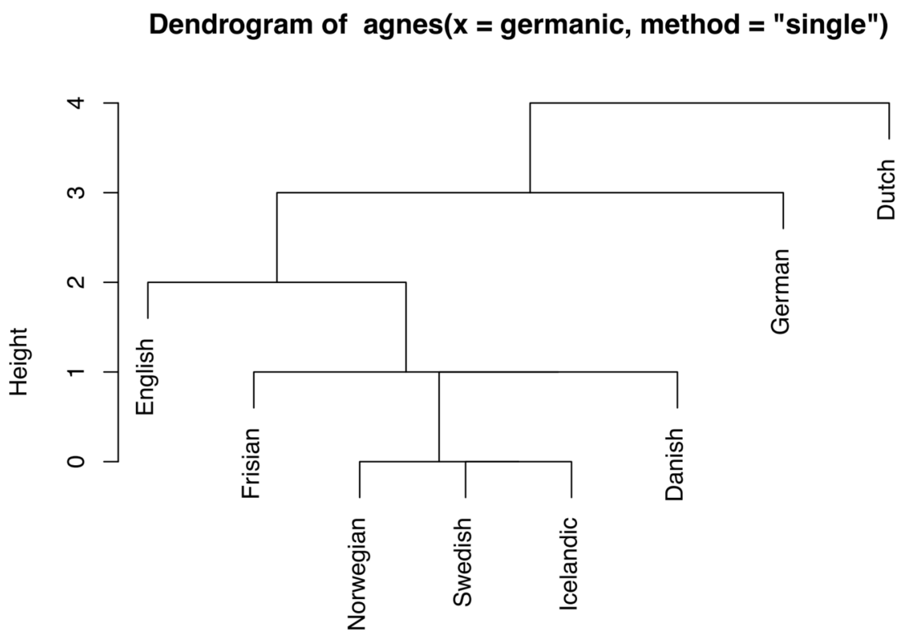


Figure 4.1. A dendrogram of agglomerative hierarchical clustering (Divjak & Fieller 2014: 114)

The above diagram presents a clustering of languages in the Germanic family based on their words for numbers. Languages grouped at the same height and on the same branch were clustered at the same stage and thus they are very closely related. We can see, that the algorithm clustered Norwegian, Swedish and Icelandic first, then it added Frisian and Danish, etc. Height differences reflect how similar given clusters are: the greater the height difference, the less similar clusters are. The difference in height between Dutch and the cluster of Norwegian, Swedish and Icelandic is the largest on the diagram, which makes those languages the least related (or similar) to each other from the entire Germanic family – at least according to this clustering solution.

Four main kinds of hierarchical agglomerative cluster analysis exist: single linkage (or nearest neighbour), complete linkage (or farthest neighbour), average linkage, and Ward's method. The methods differ in how they calculate the distance between clusters. Single linkage takes into consideration the minimum distance between a pair of objects from two different clusters (i.e. the nearest neighbours), complete linkage takes the maximum distance (i.e. the farthest neighbours), while average linkage calculates the average distance between all objects in two clusters (Everitt et al. 2011: 79). The workings of each method are illustrated in Figure 4.2:

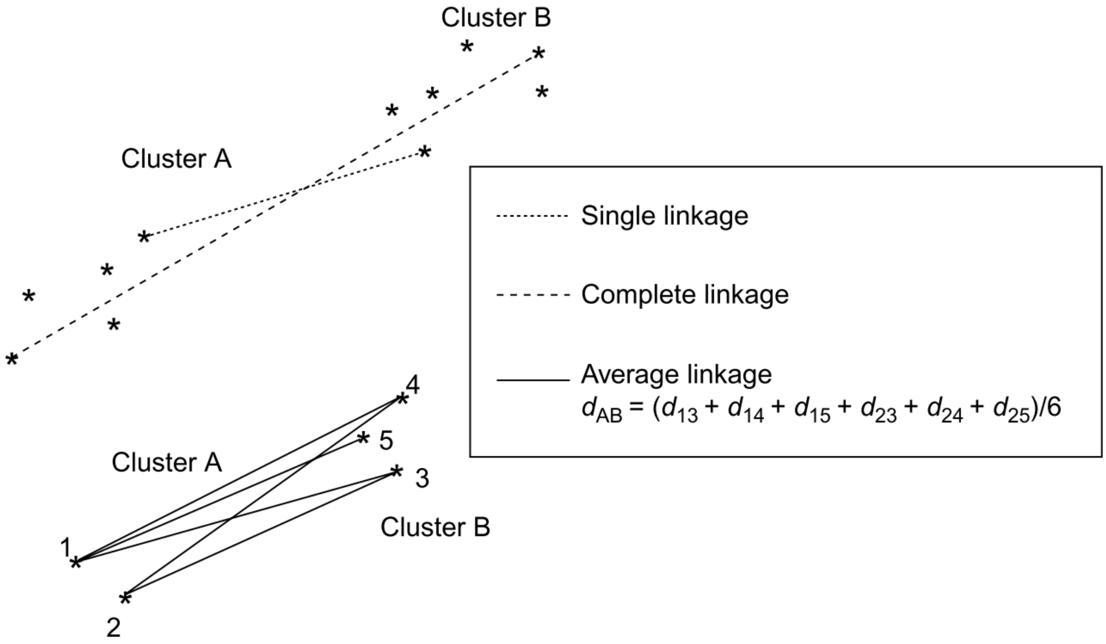


Figure 4.2. Inter-cluster distance measures: single linkage, complete linkage, and average linkage (Everitt et al. 2011: 77)

The single linkage method usually produces unbalanced clusters and is prone to a negative effect called chaining. The chaining effect manifests itself because the algorithm always takes the nearest neighbour and joins it with the existing clusters, which eventually results in elongated, "cigar-shaped" clusters (Kabacoff 2015: 374). This sort of an arrangement usually yields a suboptimal clustering solution, because objects that are dissimilar can be placed in the same cluster (Everitt et al. 2011: 92).

The complete linkage method does not exhibit chaining effects, which derives from its mathematical properties – complete linkage algorithms seek data points divided by largest distances (see Figure 4.2.). In practice, complete linkage algorithms give very compact, spherical clusters (Divjak & Fieller 2014: 116). Nevertheless, the complete linkage method is not without its own disadvantages: because it uses the distance between the farthest neighbours to determine which cluster a given data point should go into, outliers can distort the results produced by the method significantly. In the complete linkage method, an outlier can prevent two very close clusters from merging (Yim & Ramdeen 2015: 11) and thus yield a suboptimal solution, again. Being based on the average distance between points, average linkage tries to strike a balance between the two methods. It also tends to produce relatively compact and spherical clusters.

One more algorithm deserves a mention here, namely Ward's method. Ward's method clusters data in a less straightforward way than the three previous algorithms – it attempts to minimise the squared distance of data points from the cluster mean. Two clusters will merge if the sum of such distances is smaller than for any other possible merger (Divjak & Fieller 2014: 118). Ward's method usually yields spherical clusters of a roughly similar size (Everitt et al. 2011: 79).

Cluster analysis in sentence sorting experiments works on co-occurrence tables, that is, tables which contain the frequencies of how often sentences (or verbs) co-occur with each other in one group. Consequently, the data needed to be pre-processed before any analysis could take place. Co-occurrences are calculated pairwise – if one sentence occurs in the same group with another (regardless of which group it was), this pair scores one point. This procedure is repeated for every sentence in every response until a co-occurrence table such as the one presented in Table 4.1. is generated:

	perfumować.się	zepsuć.się	kłócić.się	budować.się	mówić	otulić.się	skończyć.się	licytować.się	wychować.się	pisać	zabić.się	utworzyć.się	przerzucać.się	leczyć.się	szacować
perfumować.się		27	9	6	8	22	43	7	13	31	35	7	2	20	1
zepsuć.się	27		35	41	28	12	33	5	22	24	63	17	10	22	7
kłócić.się	9	35		39	44	23	28	36	66	12	18	9	22	69	12
budować.się	6	41	39		41	36	36	26	39	28	10	54	46	17	22
mówić	8	28	44	41		15	23	35	20	31	19	41	30	29	56
otulić.się	22	12	23	36	15		40	46	38	10	40	52	51	35	6
skończyć.się	43	33	28	36	23	40		20	36	33	20	37	25	38	18
licytować.się	7	5	36	26	35	46	20		24	18	17	37	64	29	32
wychować.się	13	22	66	39	20	38	36	24		9	27	25	39	69	3
pisać	31	24	12	28	31	10	33	18	9		6	18	9	21	82
zabić.się	35	63	18	10	19	40	20	17	27	6		16	21	34	5
utworzyć.się	7	17	9	54	41	52	37	37	25	18	16		64	20	29
przerzucać.się	2	10	22	46	30	51	25	64	39	9	21	64		22	33
leczyć.się	20	22	69	17	29	35	38	29	69	21	34	20	22		16
szacować	1	7	12	22	56	6	18	32	3	82	5	29	33	16	

Table 4.1. Co-occurrence table for the *się* sorting experiment (3 sentences per meaning)

This table is the basis for creating distance matrices, which enable the clustering algorithm to determine how closely to each other the different points lie. Three distance measures (ways of calculating the distance matrices) will be considered in this study: Euclidean distance, Manhattan distance, and Canberra distance. Euclidean distance measures the straight-line distance between two points, that is the shortest possible distance. The Manhattan measure uses as grid to calculate distance – distance between two points is not a straight line but a series of small line segments that would move through a grid. We could compare it to a taxi driver travelling between two points in a perfect grid American city, hence the name: Manhattan distance. Lastly, Canberra distance is essentially a weighted version of Manhattan distance.

4.3.2. Data analysis and discussion

Three types of cluster analysis were run for each dataset (single-linkage, complete-linkage, and Ward's algorithm) on three different distance matrices (Euclidean, Manhattan, and Canberra), which yielded nine solutions for each dataset. There are no hard-and-fast rules or recommendations on which clustering method to use. The researcher must ultimately make the choice after careful consideration of the properties of a particular dataset and the solution which the researcher strives to obtain. Nevertheless, agglomerative clustering methods offer a useful heuristic that can help choose between different clustering solutions for a dataset: the *agglomerative coefficient*. The agglomerative coefficient is an index indicating the quality of a clustering solution or "the amount of clustering structure found" (Struyf, Hubert & Rousseeuw 1997: 28). In essence, the algorithm assesses how dissimilar the clusters are – the higher the value the more dissimilar the clusters. Optimally, we should always aim for clusters that are as dissimilar as possible, because we would like the clustering solution to reflect the differences in the data. Therefore, the higher the value of the coefficient the better the solution. One must be careful, however, when comparing clustering solutions for different datasets, especially for datasets of different sizes. The agglomerative coefficient is sensitive to sample size – its value grows for bigger datasets. This property of the agglomerative coefficient significantly limits its usefulness for comparing different datasets. Nonetheless, the agglomerative coefficient remains a powerful heuristic for comparing different clustering solutions for the same data.

The values of the agglomerative coefficient for all clustering solutions were compared, and the solution with the highest value was selected as optimal – a Ward's method solution on a Euclidean matrix for the 3-sentence experiment and a Ward's method solution on a Manhattan matrix for the 5-sentence experiment. The values of agglomerative coefficients for all solutions can be found in the Table 4.2.:

	WARD		
	Euclidean	*Manhattan*	*Canberra*
3 sentences	**0.6824**	0.6802	0.5871
5 sentences	0.8863	**0.89**	0.845
	COMPLETE LINKAGE		
	Euclidean	*Manhattan*	*Canberra*
3 sentences	0.5853	0.5835	0.5207
5 sentences	0.7768	0.7783	0.7088
	SINGLE LINKAGE		
	Euclidean	*Manhattan*	*Canberra*
3 sentences	0.2745	0.3175	0.3887
5 sentences	0.4627	0.4867	0.3628

Table 4.2. Agglomerative coefficient values, all clustering solutions, both versions of experiment.

The results produced by hierarchical clustering algorithms can be visualised as a dendrogram, in which small (usually pairwise) clusters merge into bigger clusters until no more mergers occur – the problematic part of interpreting such a dendrogram lies in choosing the optimal number of clusters. A clustering solution for the 5-sentence experiment (25 sentences in total) will produce 12 pairwise clusters and many more clusters of a higher order. The decision at which point to "cut" the tree and decide on the final number of clusters remains at the discretion of the researcher. Nevertheless, the decision must not be arbitrary and should be based on whether the produced clusters can be interpreted in the light of the research question and whether the clusters are statistically robust.

A method that helps assess the quality of a clustering solution are **silhouette** widths. Silhouette width is an index that "compares [an object's] separation from its cluster against the heterogeneity of the cluster" (Everitt et al. 2011: 128). In other words, silhouette width measures how similar the objects in a given cluster are and how a given object lies from the centre of its cluster. The index can take values from -1 to 1; values close to 1 mean that an object has been classified well, while values close to -1 mean that an object has been 'misclassified' (Everitt et al. 2011: 129). Silhouettes are calculated for each object in each cluster, but we can also use them to assess the quality of whole clusters or even the entire clustering solutions by calculating the average silhouette width for a cluster or the full solution. Silhouettes always take into consideration a predefined number of clusters – in order to compare different numbers of clusters, we need to recalculate silhouettes for each configuration.

The average silhouette widths for the 3- and 5-sentence-per-meaning version of the experiment are presented in Table 4.3.:

VERSION	2 clusters	3 clusters	4 clusters	5 clusters	6 clusters	optimal solution
3 sentences	0.2	0.24	**0.33**	0.27	0.22	**3 clusters**
5 sentences	**0.36**	**0.36**	**0.36**	0.35	0.35	**3 clusters**

Table 4.3. Average silhouette widths for both versions of the experiment

Analogously to choosing the optimal clustering with the agglomerative coefficient, the cluster number with the highest average silhouette width was chosen as optimal. For the three-sentence-per-sense version of the experiment, the four-cluster solution was selected, with 0.33 as the average silhouette width. In the case of the five-sentence version of the experiment, we have three solutions with the same silhouette width, thus the final optimal solution will be selected based on the *p*-values of the individual clusters. The process of selecting the optimal solution for the five-sentence version will be discussed in the subsection presenting the results of this version of the experiment.

4.3.2.1. Version A: three sentences per meaning

In the first experiment, participants were provided with three sentences per purported meaning of the reflexive marker, which makes 15 sentences in total. The dataset consisted of 149 responses and was analysed by means of hierarchical agglomerative cluster analysis. The finally chosen clustering solution was the one calculated using Ward's method on a Euclidean distance matrix, and optimal number of clusters was four, because the four-cluster solution had the highest average silhouette width. Let us first inspect the silhouette plot for the four-cluster solution to see the silhouette widths for particular clusters (Figure 4.3.):

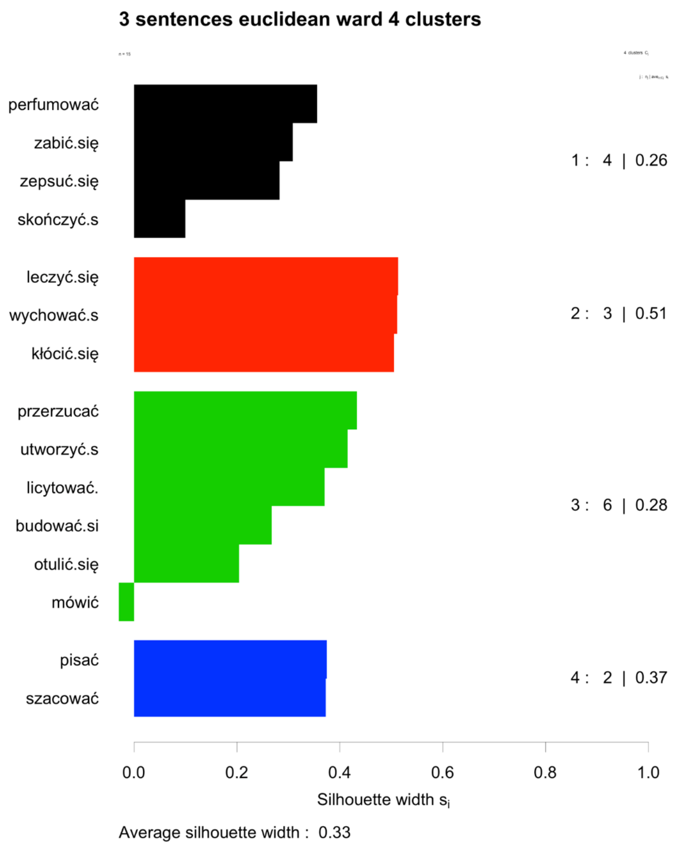


Figure 4.3. Silhouette plot for the solution with the highest average silhouette width (0.33); 3-sentence experiment.

As we can see, the item with the verb *mówić* 'speak, talk' is likely to have been assigned to a wrong cluster, because its silhouette width is negative. We will not consider the plots for all other solutions in detail here for the sake of clarity (they can all be found in Appendix 8), but three more solutions contained a misaligned item in one of the clusters. The only solution that did not contain an item that might have been assigned to a wrong cluster was the two cluster solution, with the lowest average silhouette width of 0.2. The average silhouette width of 0.2 is very low and means that "no substantial structure has been found" (Spector 2011: 172; see Section 2.3.4.). In other words, the clusters in the solution might have arisen due to pure chance, so we should not interpret this solution as a whole – even though it contains no misaligned items. To sum up, the solution with four clusters (average silhouette width of 0.33) will be considered as optimal and analysed further.

The four-cluster solution contains a misaligned item, which is a warning that the structure might be unstable, and we should investigate the stability of the clusters that the solution contains. A method that allows us to check whether clusters did not appear only due to chance is multiscale bootstrap resampling. The bootstrapping algorithm, developed by Suzuki & Shimodaira (2006), takes the analysed dataset and draws from it *n* random samples (e.g. 1000) of different sizes. Subsequently, for each of the random samples, the algorithm runs a cluster analysis, until *n* cluster analyses are produced. Then, the algorithm checks whether the clusters from the original solution are also present in each of the *n* clustering solutions generated for the random samples. Finally, the *p*-value for each original cluster is calculated based on how often it appeared in the *n* clustering solutions generated by the bootstrapping algorithm – the more often it appeared, the higher the *p*-value. A high *p*-value for a cluster will mean that the cluster does not exist only as a result of a sampling error, but can also be observed if the size and composition of the sample change (Shimodaira & Suzuki 2017).

Let us now analyse the p-values for individual clusters obtained by means of multiscale bootstrap resampling to see which clusters were statistically significant. The plot of the clustering result along with clusters' *p*-values can be found in Figure 4.4..

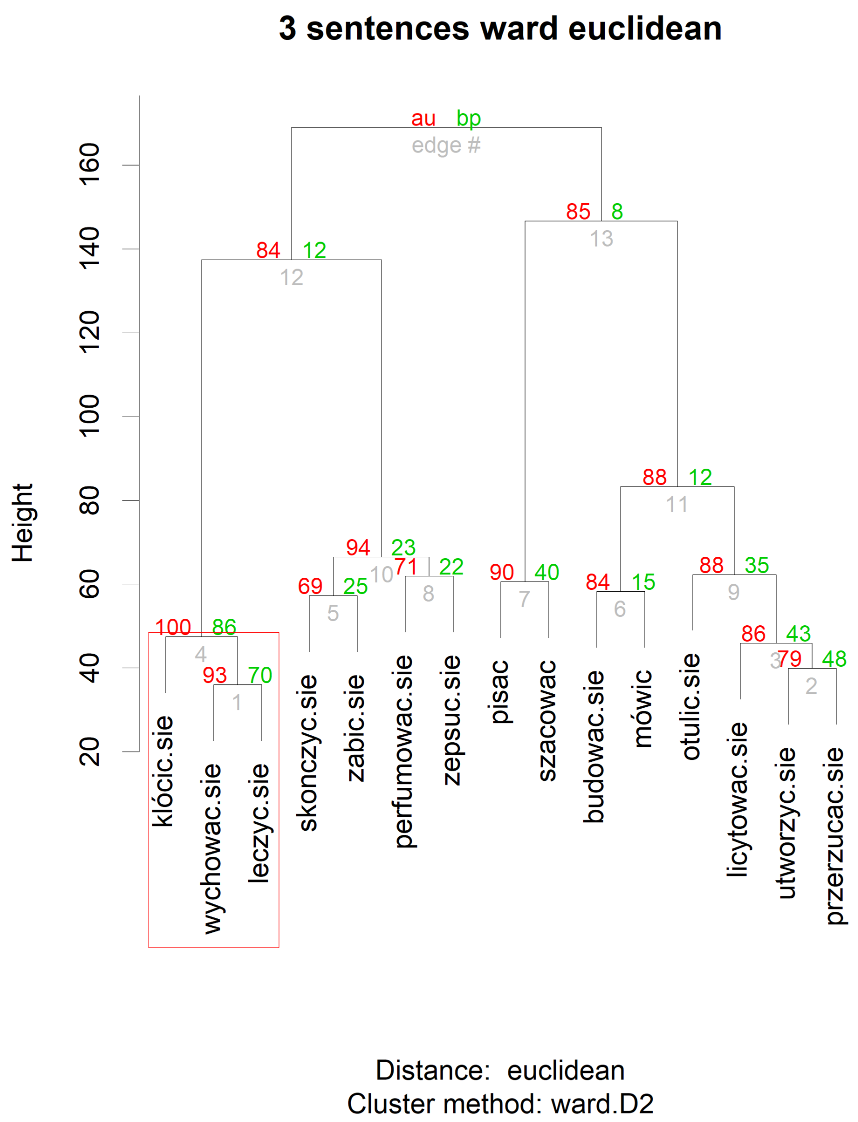


Figure 4.4. pvclust output for the three-sentence version of *się* sorting experiment

The plot contains two different estimations of *p*-values: AU (red numbers) and BP (green numbers). BP (Bootstrap Probability) calculates the probabilities by means of normal resampling, while AU (Approximately Unbiased) uses bootstrapped samples of different sizes in order to offset the possible bias of the same size for all *n* samples for bootstrapping (Shimodaira 2004: 2619). AU values provide a better approximation of the *p*-value (Divjak & Fieller 2014: 127) and should thus be used for the assessment of cluster quality in favour of BP probabilities. In order to decide whether a cluster is statistically significant (i.e. whether it has not only appeared by chance or as a result of an error), we generally use the "standard" probability threshold of 0.95[31], which means that we concede that there might be a 5-per cent chance that a cluster appeared "by accident". We will regard a cluster as statistically significant if its *p*-value is above 0.95; for the sake of convenience, the *pvclust* package marks the largest statistically significant clusters with red rectangles. We can see that only one cluster reached statistical significance (marked with a thin red rectangle), and all the other clusters were not statistically significant, including three out of four clusters from the optimal solution. This result suggests that the analysis did not find any substantial structure in the sentence-sorting data, even though the average silhouette width of the four-cluster solution was 0.33.

Overall, the clusters produced by the algorithm are not very strongly supported by the data for two reasons. Firstly, the average silhouette widths are quite low (0.2-0.33), and, what is more, the only solution without a "misassigned" verb exhibits the lowest average silhouette width, that is 0.2. Secondly, most clusters missed statistical significance – including three out of four clusters chosen as optimal in terms of silhouette analysis. Ultimately, these results mean that the algorithm could not find any stable and clear structure, and it produced the clustering solution only because it is designed to do so, not because the data cluster naturally. In other words, the structure produced by the algorithm might be artificial. We could compare this situation to someone looking at a handful of marbles thrown randomly onto a floor. If given a task to find smaller groups – for instance, three or four different "clusters" – this person will try to find groups among the marbles to complete the task. It does not follow, however, that the groups were "naturally" there, e.g. a group of blue marbles, a group of red marbles, a group of green marbles, and so on. The likelihood that those groups would actually be there is minuscule because the marbles have been thrown **randomly** on the floor. The clustering algorithm also worked on essentially random data with no natural structure and found very unstable – most likely artificial – structures as a result.

The participants in the experiment either could not find any clear similarities between the provided sentences but still grouped them into five groups because they had been asked to, or they all found different kinds of similarities, and the

[31] The *p*-values in the plot displayed in Figure 4.5. are multiplied by 100. Consequently, a value of 95 in the plot, for instance, would correspond to a probability level of 0.95.

clustering algorithm could not find any stable common structure. After a detailed inspection of the stimuli, it was impossible to find any clear patterns in which the participants grouped the sentences. In conclusion, we can say that the participants in the three-sentence-per-meaning version of the experiment **did not** group the sentences in accordance with the different meanings of the reflexive marker *się*.

4.3.2.2. Version B: five sentences per meaning

The experiment consisted of two versions that differed in the number of experimental sentences, because the hypothesis was that if the number of sentences was too low, participants would not have enough information to observe similarities between the stimuli and, thus, would not be able to group them consistently. The other version of the experiment, with five sentences per meaning, provided the participants with 25 sentences altogether, i.e. ten sentences more than in the first version. We will now discuss the results of the five-sentence-per-meaning experiment and see whether the increase in the number of experimental sentences changed the result in comparison to the three-sentence-per-meaning experiment.

The dataset consisted of 129 responses, and, analogously to the first experiment, it was analysed by means of hierarchical agglomerative cluster analysis. Only one clustering solution will be discussed in detail: Ward's method on a Manhattan distance matrix, because this was the solution with the highest agglomerative coefficient (see Table 4.2.). Analysing the average silhouette width did not help to choose the optimal number of clusters. Five different numbers of clusters (2-6 clusters) were assessed in terms of average silhouette width, but the results for all of them were virtually identical because the average silhouette widths ranged from 0.35 to 0.36. Due to the lack of differences between the different cluster numbers in terms of average silhouette width, the optimal number of clusters will be selected based on the *p*-values of individual clusters. A plot of the solution including the *p*-values of individual clusters is presented in Figure 4.5..

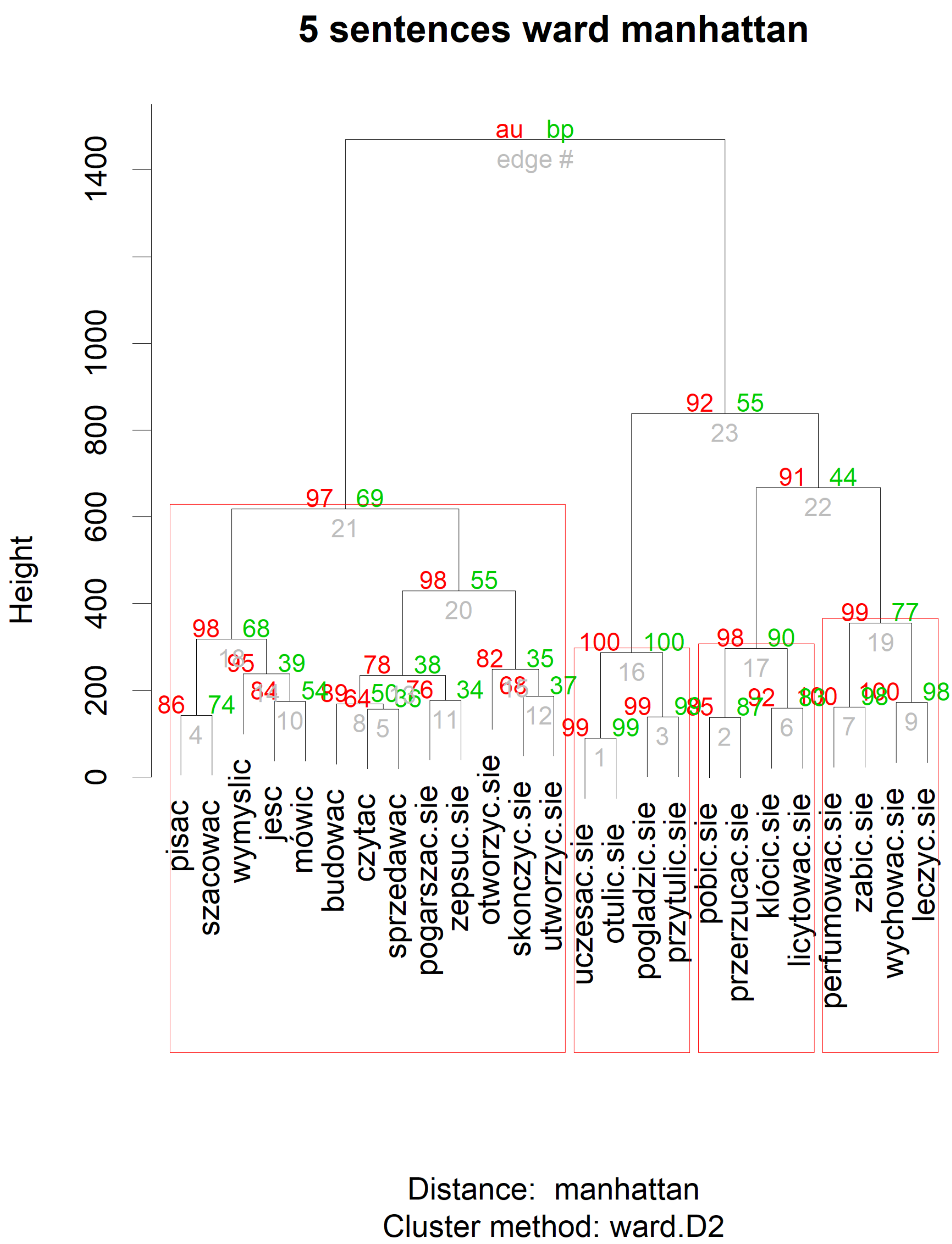


Figure 4.5. pvclust output for the five-sentence version of *się* sorting experiment

Red rectangles mark the highest-level clusters that achieved statistical significance. This does not imply, however, that lower-level clusters were not statistically significant – in fact, we can see many clusters on lower branches of the dendrogram with *p*-values higher than 0.95. The *pvclust* package in R, which was used to calculate and render the *p*-values for the solution, does not allow for the colour-coding of individual items, so a new plot with items colour-coded for meanings is presented in Figure 4.6..

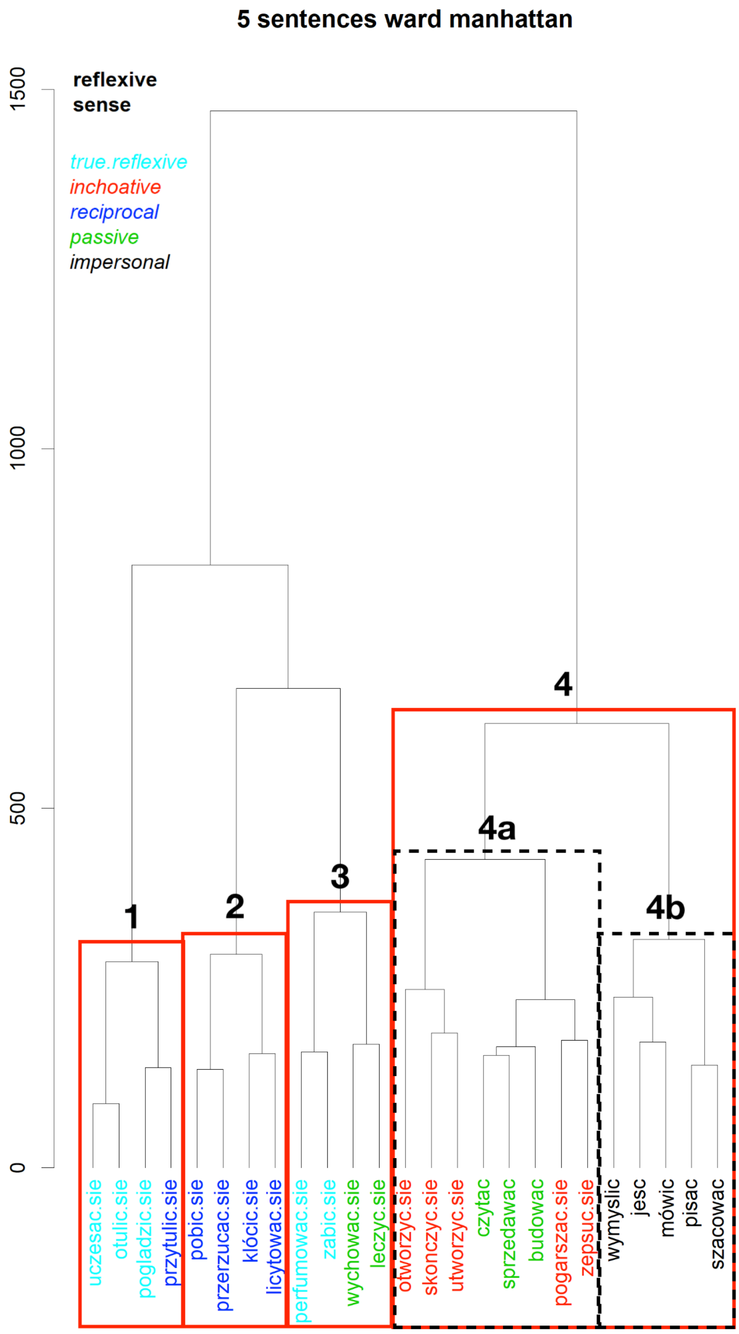


Figure 4.6. Dendrogram for the five-sentence version of *się* sorting experiment

The outputs in Figure 4.5. and Figure 4.6. differ in terms of the positioning of clusters (one is a mirror image of the other), but the solution itself remains identical as far as the composition of clusters and the relations between them are concerned – the R packages that rendered each solution use different drawing algorithms. The four largest statistically significant clusters were marked with red rectangles and were given numbers from one to four. Additionally, cluster four was divided into two (statistically significant) subclusters: (4a) and (4b). Clusters 1, 2, 3, 4a, and 4b make five clusters altogether, which is also the number of different bins the participants were asked to group sentences into and the number of different senses of *się* that were included in the experiment. Having five clusters in the solution, we will be able to see whether the clusters overlap with the different senses of *się*.

The items in the four clusters highlighted with red rectangles in Figure 4.6. form coherent groups in terms of the senses of *się*. In Cluster 1, there are three verbs with the REFLEXIVE meaning of *się* and one reciprocal; Cluster 2 consists exclusively of verbs with the reciprocal meaning of *się*; Cluster 3 contains two REFLEXIVES and two PASSIVES; Cluster 4a consists of five INCHOATIVE/RESULTATIVE verbs and three PASSIVES; lastly, Cluster 4b contains exclusively verbs with the IMPERSONAL meaning of *się*. It appears that the different senses of *się* could effectively explain the clustering solution presented in Figure 4.6. Nevertheless, we need to entertain other possible explanations of the clustering, that is, other possible criteria that the participants might have used to group sentences. Three formal criteria will be analysed: tense, verbal inflection (number, person, and gender), and the initial element of the sentence, because these are the most overt, and thus most transparent, cues the participants could have used for sentence grouping.

Let us first inspect a dendrogram of the same clustering solution as the one displayed in Figure 4.6. but colour-coded according to the inflection of the verb included in the experimental sentence. The dendrogram can be found in Figure 4.7.

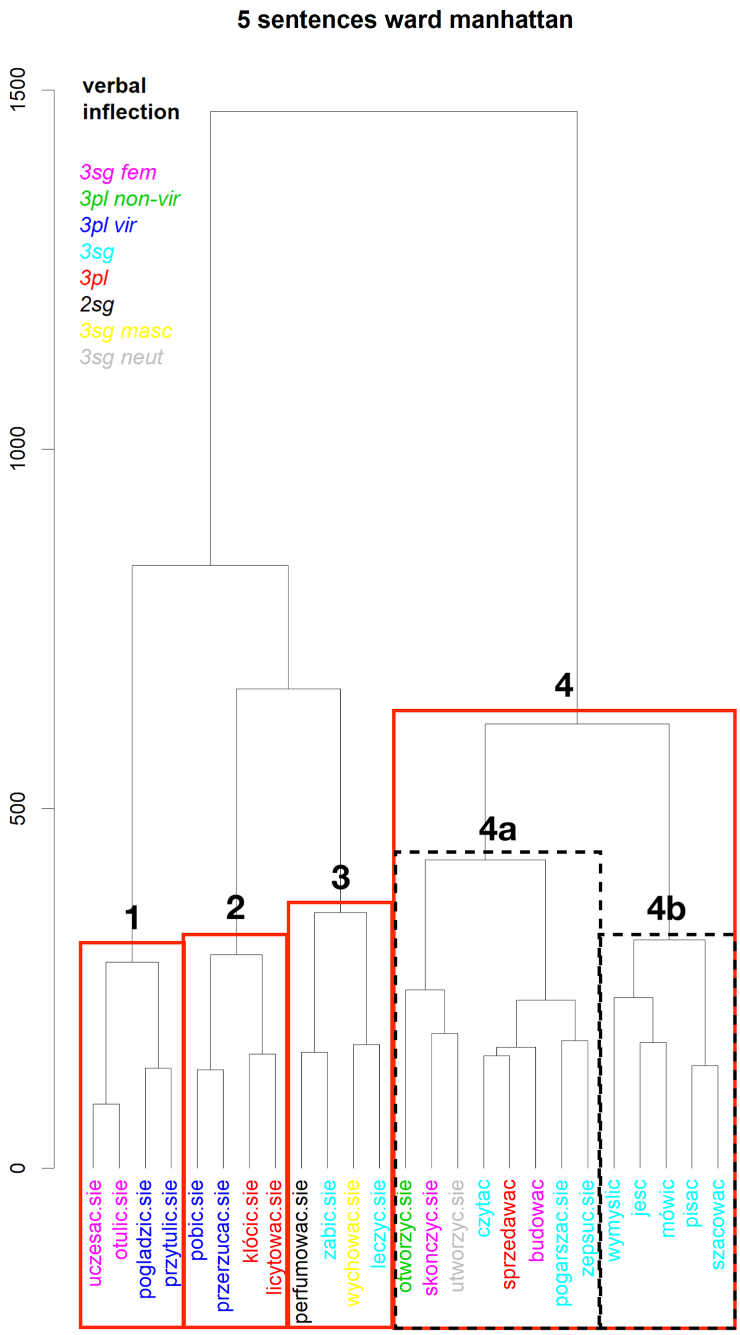


Figure 4.7. Dendrogram for the five-sentence version of *się* sorting experiment colour-coded for verbal inflection

The dendrogram suggests that the different senses of *się* explain Clusters 1, 2, 3, and 4a better than verbal inflection. When it comes to the sense of *się*, three out of four sentences in Cluster 1 are of the same sense (REFLEXIVE); when we look at the verbal inflection, the cluster contains two verbs in third person singular feminine and two in third person plural virile. Cluster 2 contains only the RECIPROCAL sense of *się,* but when it comes to verbal inflection, it contains two different kinds: third person plural and third person plural virile[32]. Cluster 3 contains three different types of inflections, while when it comes to senses, it only contains two different senses, and each small subcluster of Cluster 3 is made up of the same sense of *się.* Cluster 4a contains verbs with only two different senses of *się* (PASSIVE and INCHOATIVE/RESULTATIVE), but when we look at verbal inflection, we can see as many as five different types there. Only in Cluster 4b, verbal inflection can explain the clustering equally effectively as senses of *się* – this cluster contains verbs with only the IMPERSONAL sense of *się*, which all also happen to be inflected in third person singular[33]. In the case of the IMPERSONAL sense of *się*, it always correlates perfectly with the third person singular verbs, because sentences containing *się* in this sense always contain no subject **and** are inflected in third person singular. What follows, the fact that Cluster 4b contained verbs in the same inflection does not mean that inflection explains this clustering better – this inflection correlates with the IMPERSONAL sense of *się*. Overall, verbal inflection by number, person, and gender does not seem to explain the clustering solution better than the senses of *się*.

Another formal property that would be easy to perceive by the participants and could potentially explain the clustering solution (and participants' groupings) is the sentence structure – the initial element of the sentence in the particular. The sentences in the experiment could begin with an overt subject, a verb (i.e. implicit subject), or an adverbial. Examples of each type of sentence are presented in (4.1.):

(4.1.)	a.	*Spotka-l-i*	*się*	*w*	*barze.*
		meet-PST-3PL.VIR	*się*	in	bar
		'They met in a bar'			
		IMPLICIT SUBJECT			
	b.	*Wieczor-em*	*sie*	*odpoczywa.*	
		evening-DAT	*się*	rest.3SG	
		'One rests in the evening'			

[32] The gender inflection in Polish verbs applies only to the past and imperfective future tenses. For all other tenses (present and future perfective), the verb is only inflected for the number but not gender. Hence, the third person plural is not a superordinate category for third person plural virile.
[33] Here, again, third person singular means that the verbs were not inflected for gender, because they were either in the present or future perfective tense.

ADVERBIAL

c. *Szkoła* *się* *buduje.*

school$_{NOM}$ *się* build

'The school is being built.'

OVERT SUBJECT

Let us now inspect the dendrogram for the solution presented in Figures 4.6. and 4.7. but this time colour-coded for sentence structure:

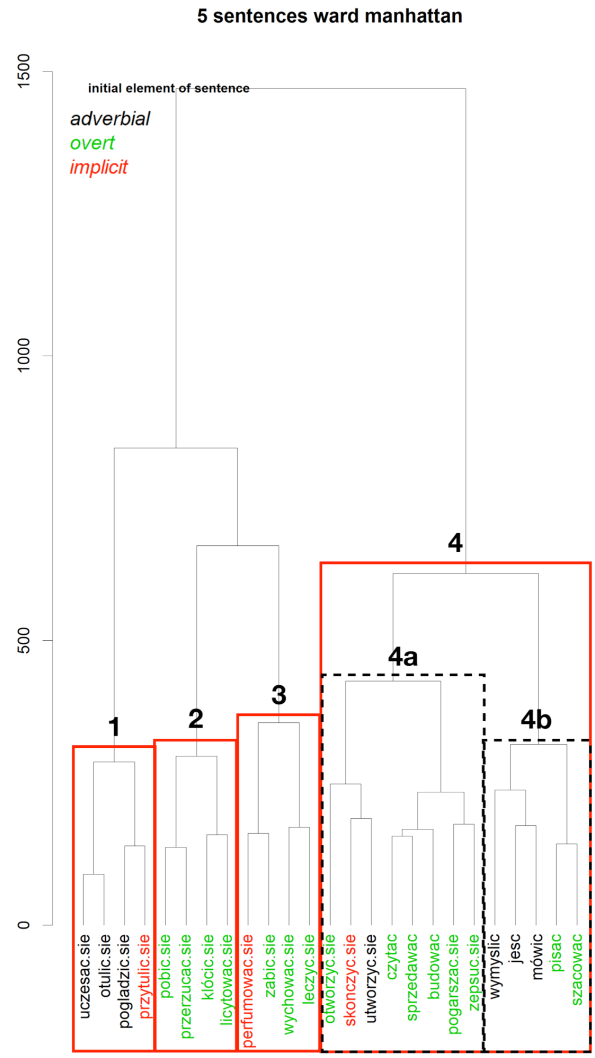


Figure 4.8. Dendrogram for the five-sentence version of się sorting experiment colour-coded for sentence structure

We can see that there is some correlation between the element that begins the sentence and cluster structure, for instance, Cluster 2 is fully made up of sentences that begin with an overt subject, and three out of four sentences in Cluster 1 begin with an adverbial. That notwithstanding, the variation in the initial element of the sentence is insufficient to explain the clustering correctly – there are only three types, and one type (overt subjects) appears in 15 out of 25 experimental sentences, which makes 60% of all cases. Only three types of options with one option dominant does not seem to be enough variety to "fill" five bins that the participants were asked to group sentences into. In conclusion, it appears unlikely that the participants used the initial element of the sentence as a criterion for grouping the experimental stimuli.

Cluster 1 contains almost exclusively reflexive-type meanings. Cluster 2 contains exclusively reciprocal meanings. We could say that clusters one and two contain the more "agentive" meanings of *się*. Cluster 3 is split in half: it contains two reflexive-type meanings and two passive meanings. This cluster most likely contains outliers, that is items with which people did not know what to do and probably binned them into one group after they had grouped all the remaining items. With 13 items, Cluster 4 is the biggest of all clusters established in this analysis. We could say that this cluster contains almost all "non-agentive" meanings of *się*: PASSIVE, IMPERSONAL, and INCHOATIVE/RESULTATIVE. The last cluster can be further subdivided into two smaller clusters: (4a) a cluster containing INCHOATIVE/RESULTATIVE and PASSIVE meanings of *się* and (4b) a cluster containing exclusively the IMPERSONAL meaning of *się*. The INCHOATIVE/RESULTATIVE and PASSIVE meanings of *się* display a range of similarities – most importantly, they both demote the agent and construe an event or action as if it had happened independently from any external forces. The similarities between the two meanings might have caused them to form a coherent cluster. The other subcluster (4b) consists exclusively of the impersonal meaning of *się*, which, albeit non-agentive, differs substantially from the two previous meanings. In contrast to all the previous meanings, verbs with the IMPERSONAL meaning of *się* **always** occur in third person singular neuter, and the sentences are always subjectless, as in example (4.1.):

(4.1.) *U*	*nas*	*jad-ł-o*	***się***	*obiad*	*o*	*czternastej.*
By	us	eat-PST-3SG.NEUT	***się***	lunch	at	fourteenth.

'Our family has lunch at 2 pm.'

In all other meanings, the verb to which *się* attaches must be transitive, and the marker *się* then renders it intransitive. The IMPERSONAL *się* does not impose this constraint and "accepts" both transitive and intransitive verbs. Besides its unique morphosyntactic behaviour, the IMPERSONAL *się* construction has a very clear and general semantic meaning – it indicates a habitual action/event and makes the agent maximally diffuse. The two factors combined – a special morphosyntactic behaviour and a very general and invariant meaning – might have contributed to the fact that the participants sorted all verbs with this meaning of *się* into one group.

Overall, we can see that in the five-sentence-per-meaning experiment, the participants were able to group sentences according to the meaning of the marker *się*. Such a result indicates that native speakers of Polish might have coherent usage-based constructions for the different meanings of *się*. If the participants were able to notice the similarities between the different meanings, which go beyond very clear morphologically marked features such as tense or number, the categories for the different senses of *się* may be active in their minds. The three non-agentive meanings also formed one large statistically significant non-agentive cluster, apart from small clusters corresponding to individual meanings. This might indicate that native speakers of Polish have some overarching general schema for the non-agentive meaning of *się*. In contrast, the higher-level cluster containing the two agentive meanings (REFLEXIVE and RECIPROCAL) missed statistical significance, meaning that they might rely only on mid-level constructions.

As far as the quality of the clusters is concerned, the judgment must inevitably be mixed. On the one hand, with only 0.36, the average silhouette width of the four-cluster solution does not attain a very high level (see Table 4.3.), which means that "[t]he structure is weak and could be artificial" (Spector 2011: 172). Since silhouettes measure the compactness of clusters and their distance from one another, a low average silhouette width indicates that the obtained clusters are either spread out (i.e. the items within a given cluster lie far away from one another), or the clusters lie close to one another. On the other hand, at 0.89, the value of the agglomerative coefficient is very high, and all clusters marked in Figure 4.6. (cluster analysis of the five-sentence-per-meaning experiment colour-coded for the senses of *się*) reach statistical significance. The agglomerative coefficient measures how dissimilar different clusters are or, in other words, it measures whether clusters are clearly delineated. The reason that the average silhouette width is low, while the agglomerative coefficient is so high might be that clusters differ substantially from one another, but the items **within clusters** are not very similar. The verbs within clusters might differ a lot in their precise semantics. Let's take *uczesać się* 'to comb one's hair' and *otulić się* 'wrap oneself' as an example – these verbs denote entirely different actions. Still, these verbs might also have something in common, and this property that they have in common might be the generalisation about the sense of the marker *się* – these generalisations might also be the property that differentiates the different clusters and the property that the participants used as a criterion for sorting the sentences.

4.4. Interim conclusions

The corpus study on Polish reflexive-type markers revealed that native speakers of Polish are unlikely to be able to build one general category for *się* as a whole on the basis of linguistic input. That study, however, indicated that speakers might build multiple, less general, categories for the different senses of *się*. To investigate whether speakers of Polish are capable of building categories for the different senses of *się*, an experimental study was conducted. In the study, experimental participants were asked to sort sentences containing verbs with five different senses of się: "TRUE" REFLEXIVE, RECIPROCAL, INCHOATIVE/RESULTATIVE, PASSIVE, and IMPERSONAL. The primary hypothesis of the experiment was that if the participants sort the sentences according to the senses of *się*, they are likely to have categories for those senses.

The study provides some evidence that native speakers of Polish could build general categories for the different senses of *się*. In the five-sentence experiment, the participants reliably grouped experimental sentences according to different senses of *się*. Moreover, when other possible explanations for the participant's groupings were considered (verbal inflection, sentence structure, and tense), they did not explain the groupings better. Consequently, we can say that the senses of *się* were stronger than any other criterion as a cue for the grouping of experimental sentences. The reliability of the clustering solution was assessed by average silhouette width and bootstrapping probabilities, which indicate that the solution is reliable enough to replicate if we took a different set of participants. Thus, we can assume that the results of the five-sentence experiment are indicative of a real language phenomenon.

The results of the three-sentence experiment diverge from those of the five-sentence experiment. The cluster analysis could not find any structure in the three-sentence experiment, which suggests that participants as a group did not use any coherent criterion for sorting the experimental sentences. Such an outcome indicates that speakers might need a "critical mass" of examples in order to recognise similarities and, consequently, assign examples to the same category. This suggests that linguists need to be cautious when creating stimuli for sorting experiments and make sure experimental conditions contain enough input in each category in order for the speakers to notice similarities consistently.

The results of the experimental study on Polish *się*, jointly with the results of the corpus investigation discussed in the previous chapter, suggest that it might not always be optimal for linguists to postulate the most general construction they can develop. If we look at real language data, we might find that speakers would not be able to build categories for those maximally general constructions – this was the case with *się*. It does not follow, however, that speakers cannot build general constructions at all. The results of the sentence-sorting experiment suggest that speakers are capable of building categories for the different senses of *się*, that is, constructions of lesser generality than one construction for *się* as a whole but still more general than specific constructions for each reflexive verb. Speakers' grammars might not always be identical to linguists' grammars, and linguists should exercise utmost caution when postulating very general constructions because they might not exist in the minds of speakers.

Chapter 5: Polish prefixed verbs

5.1. Introduction

The status of prefixed verbs in Polish remains a contentious issue – linguists have not reached an agreement as to whether the verbs belong to lexicon (as individual prefixed verbs) or grammar (as prefixes that are used for word formation or aspect manipulation). If prefixed verbs belonged to grammar, we would not have any prefabricated prefixed verbs, because each time we were to use a prefixed verb, we would have to generate the prefixed verb out of a non-prefixed verb and a prefix. For instance, in order to obtain and use the verb *potrzymać* 'hold for a while', we would need to take the prefix *po-* and the verb *trzymać* 'hold' and combine them in order to produce the verb each time we would like to use it. In other words, we would only have prefixes and verbs – we would have no prefixed verbs. In the other extreme version, if prefixed verbs belonged in the lexicon, prefixes would not be necessary at all. Each prefixed verb would be stored as unit in the lexicon. If a speaker wanted to use a prefixed verb, they would "just" retrieve the verb, e.g. *potrzymać* from the lexicon, without the need for combining a prefix and a non-prefixed verb.

At first glance, the issue of whether a construction belongs in the lexicon or in the grammar seems irrelevant to usage-based linguistics, because usage-based linguistics postulates a continuum of linguistic structures from lexis to grammar and eschews solid boundaries. If we look at it from the point of view of constructions, the seemingly void conflict between lexis and grammar might in essence be a different issue – that of general vs specific constructions. Instead of asking whether prefixed verbs belong to the lexicon or to the grammar, we would ask whether **speakers can abstract from individual prefixed verbs and build general constructions for prefixes**. Prefixes in prefixed verbs can be regarded as general high-level constructions when we see them as word-formation devices. On the other hand, if we see prefixed verbs as part of the lexicon, they would constitute an array of individual specific constructions. Certainly, drawing boundaries in phenomena that are inherently continuous and fluid seems like a futile endeavour, and prefixed verbs could also be both – specific constructions, out of which more abstract generalisations (i.e. general categories) are built over time. Nonetheless, we need to empirically explore the issue and see whether native speakers of Polish do attain abstract generalisations about prefixes and hold those general categories in their minds.

Studying Polish prefixes/prefixed verbs can shed more light on the more general question whether speakers can build and use the maximally general constructions postulated by linguists. So far, the this question has not been investigated empirically, and this thesis seeks to bridge this gap. The present chapter will discuss a corpus-based study of Polish prefixed verbs.

5.1.1. The choice of prefixes

Three prefixes were chosen based on their frequencies in the National Corpus of Polish. At first, three broad groups of prefixes were established: a low-frequency, medium-frequency, and high-frequency group, and one prefix from each group was ultimately selected for the analysis: *roz-, przy-,* and *po-*, respectively. The reason for considering three frequency groups was that patterns of different frequencies might behave differently. For instance, one might argue that if the frequency of a pattern is not high enough, it would be more difficult for the pattern to become a conventionalised unit of language. For each of the prefixes, a number of senses have been postulated in the literature (see e.g. Swan 2002 or Śmiech 1986). The purported senses of the prefixes, based on Śmiech (1986), are presented in Table 5.1:

PREFIX	SENSE 1	SENSE 2	SENSE 3	SENSE 4
PO-	**DEPARTURE:** When coupled with motion verbs, *po-* introduces the sense of departing from a location. *pójść* ‘walk, go’, *pojechać* ‘drive, go’, *polecieć* ‘fly’	**DISTRIBUTIVE:** Some verbs, when paired with *po-*, receive a distributive sense – the action denoted by the verb is then performed on multiple objects. *pokąsać* ‘bite (many things/people)’, *pomyć* ‘wash (multiple things)’	**SOMEWHAT:** In some cases, *po-* reduces the intensity of the action conveyed by the verb. *pogadać* ‘have a natter’, *poćwiczyć* ‘have some exercise’	**COVER:** With some verbs, *po-* emphasises that the action covered the full area of the object. *posrebrzyć* ‘silverplate’, *pomalować* ‘paint (over)’
PRZY-	**APPROACH:** Most verbs when paired with *przy-* mean that the subject of the action is approaching something. *przyjechać* ‘come (by driving)’, *przyciągnąć* ‘drag (over)’	**FIT:** Some verbs receive the sense of fitting/attaching something to something else. *przykręcić* ‘attach with screws, drive a screw in’, *przyłączyć* ‘connect, attach’	**COVER:** Similarly to *po-*, some *przy-* verbs can convey the COVER sense. *przykryć* ‘cover’, *przysypać* ‘cover with sand/soil/etc.’, *przykurzyć* ‘cover with dust’	**SOMEWHAT/ INTENSITY:** *Przy-* verbs can convey a reduced or, in some cases, heightened, intensity of a given action. *przybrudzić* ‘make a bit dirty’ *przyglądać się* ‘observe attentively’
ROZ-	**DISPERSION:** Many *roz*-verbs convey actions that involve scattering, spreading or distributing something in many directions or around. *rozsiać* ‘sow, scatter seeds around’, *rozchlapać* ‘splash around’	**OPPOSITION:** Some *roz*-verbs convey actions semantically opposite to the base verb. *rozszyfrować* ‘decode’ (as opposed to *(za)szyfrować* ‘encode’)	**INTENSITY:** Yet another group of *roz*-verbs generally mean that the action conveyed by the base verb has reached a certain intensity. *rozśpiewać się* ‘start singing with a lot of enthusiasm’, *rozpadać się* ‘start raining heavily’	

Table 5.1. Prefix senses (*po-*, *przy-* and *roz-*)

We can see that two senses of *po-* and *przy-* overlap: COVER and SOMEWHAT. As far as the COVER sense is concerned, the verb *kryć* 'cover$_{\text{IMPF}}$' will be a good example – both ***przy****kryć* and ***po****kryć* roughly mean 'cover$_{\text{PFV}}$', but they collocate with different types of objects.

(5.1) a. *Farba nie* ***po****-kryła całej powierzchni*

Paint not **po-**covered entire surface$_{\text{GEN}}$

'The paint didn't cover the entire surface'

b. *Tata* ***przy****-krył dziecko kocem*

dad **przy**-covered child blanket$_{\text{INSTR}}$

'Dad covered the child with a blanket'

One usually uses *pokryć* to talk about some substance (a mass noun) that covers an object with a very thin layer, such as water or paint (see example 5.1a.). On the other hand, *przykryć* is more likely to be used with count nouns denoting items that can cover another item, but they can be removed easily such as blankets or lids, (see example 5.1b.). Such examples suggest that at least some part of the meaning of prefixed verbs might rely on low-level semantics – at least some verb + prefix combinations seem to convey more meaning than the mere "sum of all parts" (i.e. verb+prefix). The slight differences between how the object is portrayed in the verbs above might arise from what other words the verbs collocate with, e.g. what kinds of objects the verbs take.

The senses conveyed by one prefixed verb can also heavily depend on the context they are used in, for instance:

(5.2) a. ***Po****malował wszystkie ściany w domu.*

'He painted every wall in the house' DISTRIBUTIVE

b. ***Po****malowała przez chwilę i zaraz skończyła.*

'She painted for a while and then stopped' SOMEWHAT

c. ***Po****malował całą ścianę.*

'He painted the entire wall' COVER

The same *po*-verb, *pomalować* 'paint', was used with three different senses of the *po*- prefix: DISTRIBUTIVE, SOMEWHAT, and COVER. The factors that would allow the language user to disambiguate between the three senses are the types of objects or adverbials and their grammatical properties. As far as objects are concerned, (5.2a) contains a plural object, while in (5.2c.), the object is singular – the plural object conveys that the action was performed on more than one item, thus the sense was DISTRIBUTIVE. One can argue, however, that the action conveyed by (5.2a) also involves covering the wall with paint, and thus the sense ought to be classified as COVER. When it comes to (5.2b), the adverbial of frequency, *przez chwilę* 'for a while', is probably the element of the sentence that would enable the hearer to deduce the SOMEWHAT sense.

Additionally, in many prefixed verbs, the function of the prefix has often been described as 'purely' aspectual, e.g. *myśleć* 'think$_{IMPF}$' – *pomyśleć* 'think$_{PFV}$':

(5.3)	a.	*Myśla-ł-em*	*o*	*przyszłości.*
		think$_{IMPF}$-PST-1SG.MASC	about	future
		'I was thinking about the future'		
	b.	***Po-****myśl-ał-em*	*o*	*przyszłości.*
		po-think$_{PFV}$-PST-1SG.MASC	about	future
		'I thought about the future'		

The only difference between (5.3a) and (5.3b) is that the former conveys an action that lasted some time (the verb is imperfective), while in the latter, the action is portrayed as a punctual event and the verb is perfective. In verbs such as *pomyśleć* in (5.3b), the prefix (*po*- in this case) is purported not to contribute any meaning besides changing the aspect of the verb from imperfective to perfective. The phenomenon of 'purely aspectual' prefix-verb combination has been the topic of a heated debate in Slavonic linguistics (see Janda et al. 2013 for a comprehensive overview). What is important for this study is that verbs where the prefix carries only an aspectual function can make it more difficult for native speakers to develop constructions for the other senses of the prefixes (e.g. DISTRIBUTIVE, SOMEWHAT, and COVER for *po*-). Prefixes, even in their 'more specific' senses (e.g.

DISTRIBUTIVE), are perfectivising. The verbs with the ASPECTUAL *po-* (or any other prefix) would contribute a lot of statistical noise to the system, as speakers would need to disambiguate between the purely aspectual sense and the other, more concrete senses.

Apart from the purely aspectual function of the prefix, in many verbs, it would be difficult to trace any meaning contributed by the prefix at all – let us label such verbs as IDIOSYNCRATIC. Those verbs would usually have been derived through prefixation at some stage in the language's diachronic development, but **synchronically,** they are no longer transparent. In some verbs such as ***po****wiedzieć* 'say' we could still decompose the prefixed verb into the prefix (*po-*) and the base verb (*wiedzieć* 'know')[34], but the native speaker would probably find it difficult to establish any meaningful relationship between the base verb and its prefixed 'derivate'. In other verbs such, as ***przy****pomnieć* 'remind, remember', the base verb would have gone or almost gone out of use and would appear only in historical texts or very literary official contexts, as is the case with *pomnieć* 'remember, bear in mind'[35]. If a verb has already gone out of use, its prefixed 'derivate' is unlikely to appear morphologically transparent to native speakers, and, thus, we should expect the whole 'derivate' to be perceived as an unanalysable unit. The entries in etymological dictionaries (e.g. Boryś 2008) can provide evidence that a diachronic relationship exists between the meaning of a derivate and the verb it was based on. Even though trained linguists postulate that the relationship still exists now, it seems highly unlikely that contemporary native speakers of Polish could establish a link between the senses of *wiedzieć* and *powiedzieć*, for instance.

If we consider all the properties of prefixes described above – senses overlapping between prefixes (example 5.1), the contextual modulation of senses (example 5.2), the purely aspectual function of prefixes (example 5.3), and idiosyncratic verbs – we can see that arriving at coherent general constructions for prefixes might

[34] According to Boryś (2008: 427–473), the word *powiedzieć* 'express something with words (say)' has existed in the Polish language since the 14th century, when it was derived by prefixing the verb *wiedzieć* 'know' with *po-*. Most likely, the rationale behind the derivation was that saying something implies 'announcing something that is known', hence the connection between *powiedzieć* and *wiedzieć.*

[35] See https://sjp.pwn.pl/slowniki/pomnie%C4%87.html [accessed 8 Aug 2017]. Interestingly, *pomnieć* is itself a derivate that came into use around the 14th century (Boryś 2008: 461).

be a difficult endeavour for a native speaker of Polish. All the four phenomena introduce large amounts of variation into prefixed verbs – variation that would need to be abstracted from in order for native speakers to arrive at general constructions for the prefixes. If the noise levels are so high, one might ask the question of whether general constructions for prefixes are plausible at all – this study will seek to answer to this question.

5.1.2. Analysing behavioural profiles

Similarly to the study on Polish reflexive verbs, in the corpus study on Polish prefixes, behavioural profiles of *po-, przy-,* and *roz-* will be constructed and explored statistically. The study is aimed at investigating whether stable behavioural profiles of prefixes could be built – if building such profiles would be possible, we could surmise that those three Polish verbal prefixes constitute coherent general usage-based categories. The different "behaviours" (i.e. semantic or collocational properties) of the prefixes would manifest themselves, for instance, in the types of objects or subjects, or classes of verbs they "prefer". In other words, the different senses of prefixes should correlate with differences in distribution, that is differences in behavioural profiles. The study aims to investigate whether there are higher-level properties in corpus data that could distinctly characterise each prefix. If such high-level properties are found, we could surmise that native speakers could attain higher-level generalisations about the prefixes – if a simple statistical model could distinguish between the prefixes on the basis of some properties, the mind of a native speaker could also be able to do it.

To build behavioural profiles 1500 examples of sentences with verbs containing the prefixes *po-, przy-* and *roz-* (500 for each prefix) were annotated for multiple semantic and morphosyntactic properties and subsequently analysed by means of multiple correspondence analysis (see Section 3.2.5.1.1. for an overview of the technique). The next section will outline the source of data, the annotation scheme, and the properties of annotated data.

5.2. Data

The data for the study come from two corpora: plTenTen (a Web corpus) and NKJP (the official corpus of Polish). The structure of data sources is identical to that in the corpus study on Polish reflexives (Chapter 3) – an overview of the two corpora can be found in Section 3.2.3. The ultimate reason for extracting data from different corpora was for the data to reflect the type of language input speakers are subjected to more accurately. It is estimated that almost 50% of the written content people consume currently comes from Web sources. Consequently, combining an official corpus (NKJP) – which contains mainly data from books and newspapers – and a Web corpus (plTenTen) will allow us to more accurately recreate the structure of the sources of language input for speakers. Altogether, 1500 random examples were drawn from both corpora – 250 examples per prefix from each corpus; a sample of this size should also guarantee a good coverage of all possible senses for each prefix. Each example consisted of an occurrence of a prefixed verb and an amount of context large enough[36] to ensure that it was possible to annotate the examples for all variables, both semantic and grammatical.

The prefixed verbs in the data extracted from the two above corpora included only finite verbs and infinitives. In Polish, three more types of verb-like constructions exist: past and present participles, subject-less impersonal constructions, and deverbal nouns. These categories, however, are not marked for tense, mood or person, and, consequently, they would yield a large number of NA values[37] for variables related to verbal morphosyntax, which could potentially cause difficulties in statistical analysis. Because participles, subject-less impersonals, and deverbal nouns are not very frequent, removing them from the data will not skew the results. Consequently, in order to avoid distorted results in the statistical analyses, these three verb-like categories were omitted.

[36] The search engines for plTenTen and NKJP work differently, so the amount of context was also different for each corpus: 20 words of left and right context each for NKJP and 200 characters for plTenTen.

[37] An example is given an NA, when it cannot be tagged for a particular feature. For instance, deverbal nouns are not marked for tense, so the variable "tense" receives NA in each example that contains a deverbal noun.

5.2.1. Annotation scheme

The data extracted from the corpora were annotated (or tagged) for three sets of variables: verb-related, clause-related and subject-/object-related variables (presented in Table 5.2.). Each of the sets contains a mixture of morphological, syntactic, and semantic variables.

Type of ID tag	ID tag	ID tag levels	Example	Translation
verb-related	tense-mood	indicative-past	*kupił*	He bought [sth].
		indicative-present	*kupuje*	He/she is buying [sth].
		indicative-future	*kupi*	He/she will buy [sth].
		subjunctive	*kupiłby*	He would buy [sth].
		infinitive	*kupić*	to buy
	aspect	perfective	*kupił*	He bought [sth].
		imperfective	*kupował*	He/she was buying [sth].
	voice	active	*Janek kopnął Tomka*	Janek kicked Tomek.
		middle/reflexive	*Janek się ogolił*	Janek shaved.
		passive	*Tomek został kopnięty*	Tomek was kicked.
	transitivity	transitive	*Janek kopnął Tomka*	Janek kicked Tomek.
		intransitive	*Janek zasnął*	Janek fell asleep.
	action completion	complete	*Janek kopnął Tomka*	Janek kicked Tomek.
		incomplete	*Janek kopał Tomka*	Janek was kicking Tomek.
	prefix sense	distributive	*Pozdejmowali obrazy ze ścian.*	They took paintings off the walls.
		somewhat	***Po**gadaliśmy chwilę.*	We talked for a while.
		cover	***Przy**krył stół obrusem.*	He covered the table with a tablecloth.
		departure	***Po**leciała do Singapuru.*	She went to Singapore (by plane).
		approach	***Przy**jedź do mnie jutro.*	Come to me tomorrow (by car/bus).
		fit/attach	***Przy**kleiłem wieszaczek do drzwi.*	I stuck a peg onto the door.
		idiosyncratic	*Joasia **po**mogła go odbudować*	Joasia helped rebuild it
		dispersion	*Listonosz **roz**niósł listy do adresatów.*	The postman delivered letters to the addressees.
		opposition	*Pacjent **roz**ebrał się do naga.*	The patient undressed.
		intensity	***Roz**padało się.*	It started to rain heavily.

	semantic verb class	motion	***przy****jść*	come
		perception	***przy****jrzeć się*	have a closer look
		contact	***roz****bić*	break (apart)
		communication	***przy****toczyć*	quote
		competition	***po****konać*	defeat
		change	***roz****wijać się*	develop
		cognition	***roz****umieć*	understand
		consumption	***po****żreć*	devour
		creation	***przy****gotować*	prepare
		emotion	***roz****bawić*	make somebody laugh
		social	***po****zwolić*	allow
		bodily	***roz****ebrać się*	undress
clause-related	negation	affirmative	*Wykonał zadanie.*	He completed the task.
		clause negated	*Nie można powiedzieć, że wykonał zadanie.*	I can't say he completed the task.
		ni-word	*Nikt nic nie powiedział.*	Nobody said anything.
		preceding verb negated	*Nie chcę przegrać.*	I don't want to lose.
		verb negated	*Nie wykonał zadania.*	He did not complete the task.
	sentence type	declarative	*Poszedłem do domu.*	I went home.
		imperative	*Pójdź do domu.*	Go home.
		interrogative	*Pójdziesz do domu?*	Will you go home?
	clausal dependency	independent	*Ubrał się i poszedł do domu.*	He got dressed and went home.
		dependent	*Powiedział, że pójdzie do domu.*	He said he would go home.
	type of dependent clause	adverbial: manner	*Pokaż mi, **jak to zrobiłeś.***	Show me **how you did this.**
		adverbial: spatial	*Spotkajmy się tam, **gdzie wcześniej się spotkaliśmy.***	Let's meet **where we met before**.
		adverbial: temporal	*Pójdę tam, **gdy skończę pracować.***	I'll go there **once I finish working**.
		jeśli clause	*Zrób to, **jeśli chcesz.***	Do it **if you want**.

		relative	*To jest ta nauczycielka,* ***o której ci mówiłam.***	This is the teacher **that I told you about**.
		że-clause	*Wierzę,* ***że masz rację.***	I believe **that you're right**.
		żeby-clause	*Zrobiła to,* ***żeby dostać podwyżkę.***	She did this **to get a payrise**.
	adverbial	duration	*przez dwa dni*	for two days
		intensity/degree	*mocno*	strongly
		futility	*nadaremno*	to no avail
		necessity	*koniecznie*	necessarily
		temporal	*zimą*	in winter
		location	*na rehabilitacji*	at physiotherapy sessions
		manner	*w zadowalający sposób*	in a satisfying manner
		causal	*po to, by*	in order to
		należy, trzeba, etc.	*warto*	it's worth…
		certainty	*na pewno*	for sure
	particle	intensification	*bardzo*	very
		restriction	*tylko*	only
		choć	*choć*	at least
		exhortation	*niech (Niech wyniesie śmieci)*	let… ('Let him take the rubbish out')
subject- and object-related	animacy	animate	*żyrafa*	giraffe
		inanimate	*demokracja*	democracy
	abstractness	abstract	*demokracja*	democracy
		concrete	*żyrafa*	giraffe
	countability	count	*kamień*	stone
		mass	*wiedza*	knowledge
	number	singular	*decyzja*	decision
		plural	*decyzje*	decisions
	pronominality	noun	***Tata*** *poszedł do domu.*	Dad went home.
		pronoun	***On*** *poszedł do domu.*	He went home.

		implicit	*Poszedł do domu.*	He went home.
	semantic class	human	*lekarz*	doctor
		animal	*kot*	cat
		plant	*trawa*	grass
		mental/psychological	*uwaga*	attention
		location	*hotel*	hotel
		manmade	*robot*	robot
		other natural	*dym*	smoke
		quality/property	*teraźniejszość*	present
		organisation/group	*firma*	company
	morphosyntactic type	dative	*Powiedziałam to* ***listonoszowi.***	I said it **to the postman.**
		genitive	*Przywitałem* ***listonosza.***	I said 'hello' to **the postman.**
		instrumental	*Poszedłem z* ***listonoszem.***	I went **with the postman.**
		infinitive	*Potrafi* ***tłumaczyć*** *trudne zagadnienia.*	She can explain difficult topics.
		other clause	*Rozumiem* ***dlaczego to zrobiła****.*	I understand **why she did this.**
		że-clause	*Rozumiem,* ***że to złe.***	I understand **that it's bad**.
		accusative	*Widziałem* ***listonosza.***	I saw **the postman.**

Table 5.2. Tagging scheme for corpus data on Polish prefixed verbs

The verb-related group contains a number of 'standard' morphosyntactic properties of Polish verbs: *tense, mood, aspect, and voice*. The morphological properties of a verb – especially its aspect – express rich semantic information pertaining to the portrayal of the situation in question. All verbs in Polish are obligatorily marked for aspect, be it finite verbs, or infinitives, or deverbal nouns. As far as the semantics of verbal aspect are concerned, it can signal completeness, habituality and generality (Dickey 2000:12). Aspect, as well as other morphological properties can help assess the properties of the situation construed in a given context. Divjak & Gries justify including multiple variables as follows: "given that we attribute a central status to distributional information of syntactic and semantic nature, we must take the syntactic structures seriously in which the verbs under consideration are used as well as the types of modifiers and the range of collocates these structures harbor" (2006: 30). Since prefixes have been postulated to express adjustments in how situations in base verbs are construed, these morphological properties can have a direct impact on the choice of a prefix.

Apart from the overt morphological variables, the verb-related set also includes a number of semantic variables: *transitivity, action completion, prefix sense,* and *verb class.* Semantic verb-related variables do not receive any overt morphological coding, but they represent the different ways of portraying actions or real-world events. The first variable, *transitivity*, is a binary variable that pertains to whether the verb in a given example is transitive or not, that is whether it can receive a direct object. *Action completion* is also a binary variable and says whether the action conveyed by the sentence has been completed or not. One might say that this largely overlaps with aspect, however, with the future tense or in infinitives, the perfective aspect does not signal completion. *Prefix sense* categorises the prefixed verbs into classes corresponding to the different senses postulated for the prefixes investigated in this study, e.g. for *roz-*, the possible senses were: DISPERSION, OPPOSITION, and INTENSITY; *aspectual* and *idiosyncratic* classes were also available for each prefix. Prefix sense is a variable that was tagged for the purpose of using it in subsequent experimental studies. It will not be used in the statistical analysis that is part of this study, because most senses are associated exclusively with one prefix. If a level of an independent variable always appears with the same level of the dependent variable (e.g. the DISPERSION sense can only appear the prefix *roz-*), it would give a trivial prediction in a statistical analysis, as it would always predict the same outcome. The final variable, *verb class*, stands for the semantic class of the analysed verb. The semantic classes used in this study were taken directly from the Polish equivalent of WordNet, Słowosieć[38]. WordNet's classification consists of 15 semantic classes of verbs, such as communication verbs (e.g. *discuss*), verbs denoting cognitive actions (e.g. *ponder*), or verbs of motion (e.g. walk). For the purpose of annotation, every verb was looked up in Słowosieć and, subsequently, coded based on the information obtained from the system. In a small number of cases, the semantic class in Słowosieć did not correspond to the context or no semantic class was given at all – the semantic class of each such verb was adjusted based on the annotator's judgment (analogously to the treatment of reflexives, presented in section 3.2.4.1). The adjusted verbs can be found in Appendix 6.

As far as the clause-related group is concerned, two variables – *adverbial* and *particle* – were predicted to be very important for disambiguating between prefixed verbs with different prefixes. The different senses postulated for prefixes are essentially adjustments to the portrayal of the situation the base verb conveys, e.g. a different destination (*przyjść* 'come' vs *pójść* 'go') or different rates of completion (*upić* 'take a sip' vs *wypić* 'drink sth up'). Adverbials (e.g. *całkowicie* 'entirely') and particles (e.g. *trochę* 'a little bit') also adjust the construal of a given situation and thus might correlate with prefixed verbs expressing similar senses. Besides *adverbials* and *particles,* the clause-based group also includes a number of relatively 'standard' clausal categories such as *negation*, *sentence type*, as well as *clausal dependency* and *type of dependent clause*. The first variable, *negation*, tells whether negation was present in the clause containing the verb analysed in a particular example. This variable has more options than 'yes' or 'no'; if a clause contains negation, it can occur in different forms, for instance, the verb can be negated or a modal verb before the analysed verb can carry the negation instead. The *sentence type* variable tells whether the sentence was affirmative, interrogative, or imperative. *Clausal dependency* is a binary variable which tells whether the clause containing the verb was dependent or independent. If the clause containing the verb was *dependent*, the *type of dependent clause* variable contained the information about the exact type of the clause (e.g. whether it was a *that*-clause). The reason for including these four variables is that at

[38] http://plwordnet.pwr.wroc.pl [Accessed 14 Oct 2018].

this point, we do not know which elements convey meaning and which elements might correlate with verbal prefixes. Consequently, we should strive to include as much information as possible, as we might uncover correlations that have not been postulated before.

Prefixed verbs might also be compatible with different types of subjects and different types of objects. For instance, *po*-verbs in which *po-* has the DISTRIBUTIVE sense need to take a plural object, because the action needs to be 'distributed' over more than one object. The last group of variables – *subject- and object-related* pertains to the properties of subjects and objects in the analysed examples. Properties of subjects, direct objects, and indirect objects were analysed separately, but the variables for each group were identical, since they all usually concern either nouns or pronouns. As far as semantic variables are concerned, the subject- and object-related groups include four of them: *animacy*, *abstractness*, *countability*, and *semantic class*. The variable *semantic class* pertains to the semantic class of nouns the subject or object belongs to – that is, whether the noun denotes a human, an animal, or a psychological property, for instance. Similarly to the *semantic class* of verb, this variable was annotated in accordance with the data from the Polish WordNet. *Animacy, abstractness,* and *countability* (each of them binary) constitute more general semantic variables, which tell whether a subject/object is animate, abstract, or countable. The reason for including those was to investigate whether such general properties of verbs' subjects and objects could possibly explain the behaviour of the verbal prefixes. The corpus contexts were also coded for four formal (morphological and syntactic) variables: *number, gender, pronominality,* and *morphosyntactic type*. This set of variables can be inferred from the structural properties of the context, and it was included in the coding to investigate whether (and how) overt structural properties can explain the behaviour of the three verbal prefixes studies in this thesis. *Number* is a binary variable that indicates whether a noun was plural or singular. *Gender* indicates which grammatical gender the noun carries[39]. *Pronominality* stands for whether the subject/object is a noun or a pronoun, or whether it is implicitly expressed – the latter

[39] Polish nouns – in contrast to, for instance, English nouns – carry grammatical gender. Grammars of Polish usually distinguish five genders: feminine, masculine, and neuter in the singular number, and virile and non-virile in the plural number. For the purposes of this study, there only three types of gender were considered: feminine, masculine, and neuter – the benchmark for annotating was the gender of the singular form of a noun.

category exists in the coding because in Polish (and other Slavonic languages), the subject is often encoded on the verb and then the sentence does not contain an overt subject (see example 5.4.).

(5.4.)	***Roz***-*wiąza-ł*	*buty*
	roz-untie-PST.3SG.MASC	shoes$_{ACC}$
	'He untied (his) shoes'	

The *morphosyntactic type* variable applies only to objects and stands for the different forms a given object may take, e.g. a noun in the accusative case (most direct objects will take this morphosyntactic marking) or a *że*-clause.

5.2.2. Properties of the data

The inspection of the frequencies with which different variables occur in the dataset revealed that some levels of some variables are too infrequent and thus could distort the results of the correspondence analysis. Those low-frequency variables are all related to the semantic-class properties for verbs and nominals derived from WordNet. We will proceed with the inspection of nominal variables first. Table 5.3. presents the frequencies of all nominal semantic classes for three sentence roles: subjects, direct objects, and indirect objects. The numbers in the table represent in how many annotated examples, the subject, direct object, and indirect object belonged to which semantic class. If we look at subjects, for instance, we will see that in 793 annotated examples, the subject was a person, in 30 examples it was an act, and so on.

SEMANTIC CLASS (LEVEL)	VARIABLE		
	subject	direct object	indirect object
act	30	41	18
animal	**10**	**4**	**3**
artefact	74	81	17
attribute	**8**	37	**8**
body	**5**	11	**4**
cognition	42	66	**10**
communication	45	59	11
emotion	**4**	14	**5**
event	65	85	25
food	**3**	13	**3**
group	54	18	**5**
location	30	**9**	**7**
natural.object	11	**3**	**0**

natural.phenomenon	13	**8**	**3**
person	793	84	147
plant	**7**	**5**	**1**
possession	19	24	**5**
process	**5**	**5**	**0**
quantity	**5**	11	**1**
relation	75	21	13
state	16	20	**9**
substance	**5**	**5**	**1**
time	27	12	**4**

Table 5.3. Frequencies of nominal variable levels derived from WordNet

Many levels of the variables displayed in Table 5.3. above had frequencies that did not exceed 10 (those variables have been highlighted with the colour red and a bold font). When the frequency of a variable's level falls below 10, it may cause problems for the correspondence analysis. Observations with low-frequency levels are likely to end up as outliers on the correspondence analysis plot and obscure the relationships between other variables. In more technical terms, such observations can contribute disproportionately high amounts of inertia to the analysis and, as a result, substantially alter the map produced by the algorithm. Sometimes, they can "dominate a map so much that the more interesting contrasts between the more frequently occurring categories are completely masked" (Greenacre 2007: 92). The reason for such result is that if only a few observations have a given feature (i.e. variable level) they are treated by the algorithm as very rare and thus very distinct from other observations. We could draw an analogy here: low-frequency levels are similar to reaction times in self-paced reading or word association experiments longer than, for instance, 2000ms (or shorter than 50ms). We are bound to notice unusually short or long reaction times because they stand out from "normal" reaction times, but we should not base any conclusions on them, because such reaction times usually appear as a result of processes **external** to the phenomenon under investigation, e.g. participants' lapse of attention. The same happens with low-frequency variable levels: observations that have them stand out (i.e. are distinct) from other observations, but they might not be the result of any significant trend in the data.

Coming back to the analysis, food-related subjects occurred only three times out of 1500, and the correspondence analysis will treat them as distinct in the map, because, mathematically, they will seem very distinct. In reality, however, the extremely low frequencies of food-related subjects would more likely mean that they hardly every appear as subjects of sentences in general rather than that they are distinctive for any prefix.

From the three nominal variables presented above (the semantic class of the subject, the semantic class of the direct object, and the semantic class of the indirect object), only two are of a quality sufficient to be considered for further analysis: the semantic class of subjects and direct objects. As far as indirect objects are concerned, they appear in 300 sentences, which translates into merely 20% of all examples. More importantly, however, 20 out of 24 variable levels attained a frequency lower than 10, with one level, *person*, accounting for almost 50% of the variable (it appears in 147 sentences). In effect, we have a very high number of low-frequency levels and one level that dominates the data for this variable, which means that there is not enough variation for the correspondence analysis to detect any patterns. So many low-frequency variable levels and – even more importantly – the fact that we have data only for 20% of the examples limit the usefulness of the semantic class of the indirect object as a variable, and it will be thus excluded from further analysis.

The other two variables, subject and direct object, appear in 1346 (89.7%) and 636 (42.4%) examples respectively, and they have fewer variable levels of frequency lower than 10 (7 variables and 8 variables, respectively). This makes them better candidates for further processing, once we alleviate the problem of low-frequency levels. There are two strategies for alleviating the effect of low frequencies: (1) remove all observations with low-frequency variable levels or (2) merge the low-frequency levels into larger ones to increase their frequencies (Greenacre 2007: 207). The first strategy would bear a very negative impact on the analysis – it would entail the loss of a very large amount of data, because the number of low-frequency variable levels is quite large (nine for subject semantic class and seven for direct object semantic class). For the above reason, this study will adopt the second strategy, and the low-frequency levels will be merged. Figure 5.1. presents the scheme according to which the low-frequency variable levels were merged (for the sake of readability, the diagram does not include the variable levels whose frequencies were sufficiently high):

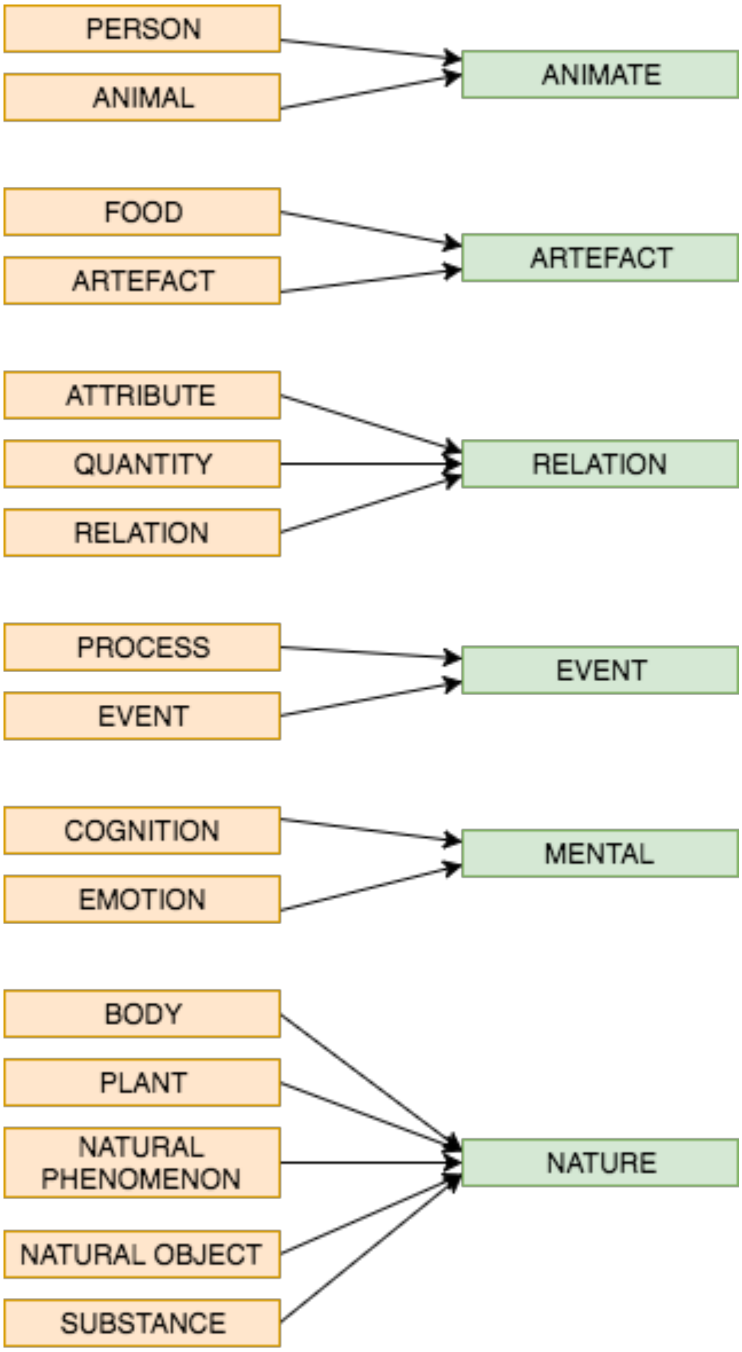


Figure 5.1. Merging of nominal variable levels – scheme

Firstly, the variable levels *person* and *animal* were merged into a single *animate* level because both human beings and animals are endowed with some sort of agency[40]. Additionally, in the Great Chain of Being, animals come only one level below than human beings (Lakoff & Turner 1989: 160). The levels *food* and *artefact* were pooled together into a joint *artefact* level. Most foods need to be

[40] Of course, the characteristics and the extent of agency in human beings and in animals differs to some extent – this is why they were separate categories to start with.

processed to some extent before humans can eat them – boiled, fried, cut, etc. – which essentially makes them products of human activity, that is artefacts. The next combined variable level is of a somewhat more abstract nature – *attribute*, *quantity* and *relation* were merged to create a larger level, *relation*. Both attributes (e.g. *red*) and quantities (e.g. *dozen*) can be seen as relational predicates predicating a relation between a trajector (an item described by the attribute or the quantity) and a landmark, where the latter is either a point on a scale or a region in some attribute space (e.g. the space of 'redness') (Langacker 1987a, p. 219). The final merged variable level, *nature*, encompasses five smaller levels: *natural object, natural phenomenon, body, plant* and *substance.* Natural objects and natural phenomena, as their names suggest, both belong to the realm of nature. The difference between the two rests in the fact that natural objects assume physical presence and thus are tangible (e.g. rocks or soil), while natural phenomena in this classification are intangible (e.g. light). *Substances* were added to the *nature* level on the basis of a close inspection of all instances of *substance* – all examples annotated as substances were substances of natural origin, e.g. water or carbon dioxide. *Plants* were included in *nature* because, even though they are essentially living organisms, humans perceive them as inanimate objects and they are part of the natural world. One might see *body* as a semantic class that could also be part of the *animate* level, because body parts belong to human beings or animals. Nevertheless, body parts do not possess any agency and their motions are instigated by the organisms they are part of. Consequently, they are perceived as objects rather than animate beings and thus should not be part of the *animate* level. Neither are body parts products of human activity, hence, the only variable level they could belong to was *nature.*

SEMANTIC CLASS	VARIABLE	
	subject	direct object
act	30	41
artefact	77	94
communication	45	59
event	70	90
group	54	18
location	30	9
person	793	84
possession	19	24
relation	88	69
state	16	20
time	27	12
living	22	20
mental	46	80
natural	29	16
NAs	154	864

Table 5.4. Frequencies of nominal variable levels after level merging

Table 5.4. above presents the frequencies of the nominal classes after the levels have been merged. As we can see, only one level remains with frequency lower than 10, namely *location* as direct object. It would be difficult to merge *location* with any other variable level, and it contains only nine examples, so this level will be removed from the dataset if it appears to be an outlier.

When it comes to semantic verb classes derived from WordNet, they exhibit a significantly better distribution, with the frequencies of only two levels dropping below 10: *bodily* verbs and *consumption* verbs (see Table 5.5.). Since the two classes combined occurred only 11 times in the entire dataset (0.73% of the data), we can safely remove them without a substantial loss of data.

SEMANTIC CLASS	FREQUENCY
bodily	4
change	243
cognition	225
communication	254
competition	26
consumption	7
contact	40
creation	57
emotion	34
motion	171
perception	72
possession	64
social	192
stative	111

Table 5.5. Frequencies of verb class before any changes to dataset

5.3. Analysis of the data

Following the merging of variable levels, the data were analysed by means of correspondence analysis. Correspondence analysis was chosen for this purpose, because the dataset for this study contains a multitude of variables, and correspondence analysis helps find patterns in multivariate data. A comprehensive introduction to the technique can be found in Section 3.2.5.1.1., but what correspondence analysis essentially does is "flatten" the data and reduce the number of dimensions[41] in the analysis to eventually plot the data on a two-dimensional or three-dimensional graph, which can then be interpreted by the researcher.

The correspondence analysis was conducted in *R* statistical software, version 3.4.0 (R Core Team 2017), using the *mjca* function from the *ca* package (Nenadić & Greenacre 2007). The particular mode of the correspondence analysis used is multiple joint correspondence analysis (Greenacre 2007: 145–152). Using multiple joint correspondence analysis can help reduce the impact of including many multi-

[41] Reducing the number of dimensions is crucial, because the more variables one includes in the analysis, the more dimensions one will need to visualise the data. Unfortunately, humans (including researchers) can only conceive of up to three dimensions.

level variables and thus increase the explained inertia levels. Multiple joint correspondence analysis requires a Burt matrix as input data (Greenacre 2007: 145–146), hence, this kind of matrix was chosen as input in the present study.

The data were annotated for 36 variables altogether, with some variables having a very high number of levels, and due to that, conducting a correspondence analysis on all the variables would yield an uninterpretable cloud of points of immense complexity. In order to avoid creating a correspondence analysis map with too many points, one large analysis was foregone in favour of a number of smaller-scale analyses exploring different aspects of the data. The following section will present separate analyses of: nominal (subject and object), verbal, and clause-related variables in various combinations. Two more analyses will be presented: (1) an analysis of only morphologically or syntactically (i.e. overtly) marked categories and (2) an analysis of only low-level semantic categories (i.e. WordNet-derived classes and adverbials). Contrasting the latter two types of variables (morphosyntax vs semantics) will enable us to see which of them has greater predictive power – overtly marked general categories or low-level concrete semantic categories.

Each of the conducted correspondence analyses treated the prefixes as a supplementary variable. Supplementary variables can be regarded as passive variables that do not influence the outcome of the analysis (Greenacre & Blasius 2006: 31). Such variables have no mass and thus do not change the position of other points on a CA plot. Treating prefixes as a supplementary variable makes it possible to see how the response variables (i.e. predictors) cluster together without the influence of the marker variable itself (see Section 3.2.5.1.2. for more details).

Before we proceed to the results of the correspondence analyses, a brief introduction to how CA plots are interpreted is in order. On a two-dimensional correspondence analysis plot, we will see many points scattered across the graph, and each point will correspond to one **level** of a variable. For instance, we will have one point for transitive verbs and another point for intransitive verbs, and so on. The lesser the distance between points, the stronger the association between them is. Consequently, to discover the categories that are closely associated with each of the prefixes, we should be looking for points that lie the closest to a given prefix.

5.3.1. Verb-related variables

We shall begin the analyses with investigating the verb-related variables. The ultimate aim of analysing verb-related variables is to see whether the properties of the verb itself can help differentiate between prefixes. Three different analyses were run and will be discussed in this section: all verb-related variables, low-level verbal semantics based on WordNet, and all verb-related variables **except** the semantic verb class. The investigation of verb-related variables was split into three analyses in order to see whether the levels of explained inertia increase when the variable with the highest number of levels (semantic verb class) is removed; the reason for including an analysis of semantic verb class only was to see whether the low-level semantic properties could explain the behaviour of the prefixes.

Let us begin with the analysis of all verb-related variables (tense and mood, transitivity, aspect, action completeness, voice, and semantic class). The plot of the correspondence analysis for this set of variables is presented in Figure 5.2.

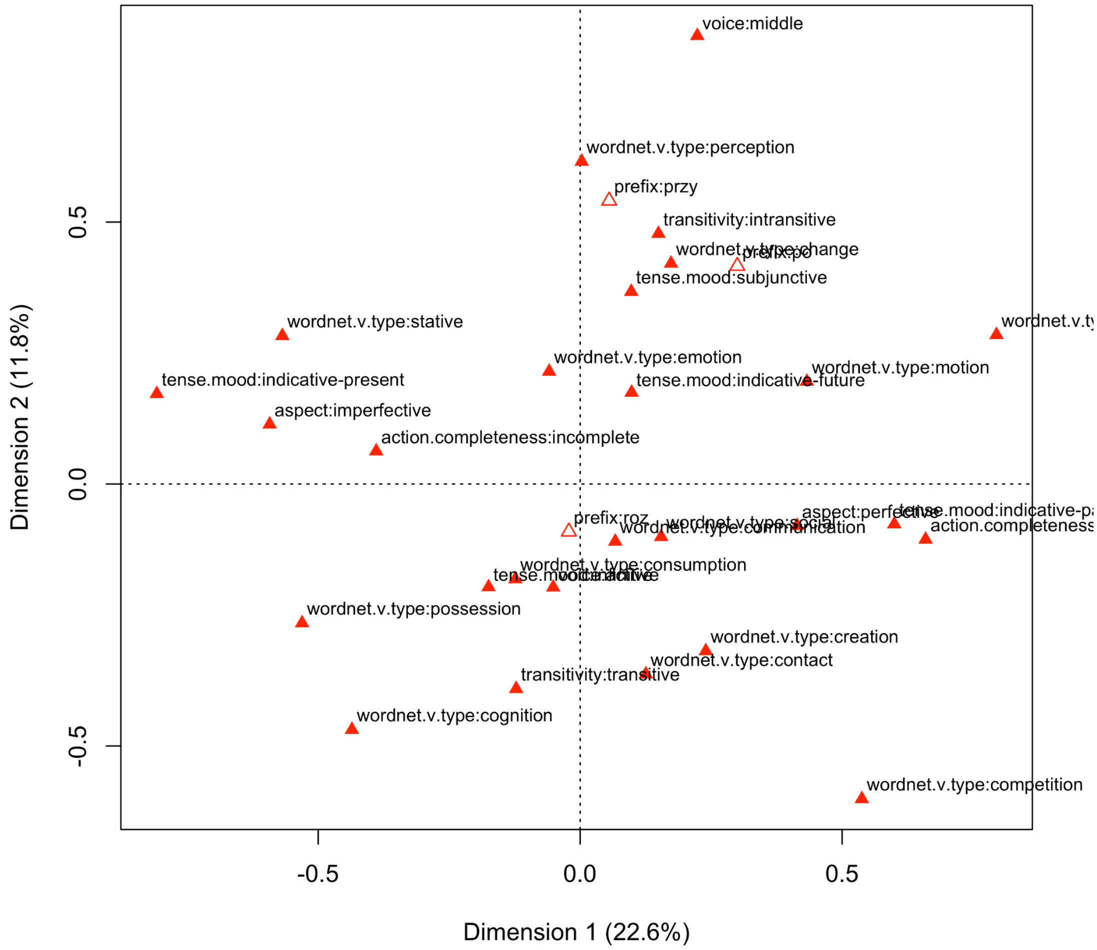


Figure 5.2. Correspondence analysis plot: all verb-related variables

The inertia levels are presented on the bottom and on the left-hand side of the plot. Altogether, the inertia explained by this analysis amounts to 34.4%[42], which constitutes quite a low result, as 65.6% of inertia remains **unexplained**. Such a low level of explained inertia indicates that the analysis is only moderately reliable and any results should be treated with caution. We can also see that many variable levels lie in quadrants other than where the prefixes are located, which means that they are not associated with any of the prefixes. Furthermore, the prefix *roz-* lies almost at the origin of the plot, and, consequently, is not associated with any of the variable levels. The three aspects discussed above – low explained inertia, many

[42] This pertains to the first two dimensions, i.e. the dimensions actually plotted in Figure 5.2.

properties in quadrants other than the ones with prefixes, and *roz-* almost at the origin – render the analysis moderately interpretable at best.

If we consider the axes of the plot, we will notice that the *x* axis represents an "aspectual" scale. Variable levels related to imperfectivity (imperfective aspect, present tense, and incomplete actions) appear on the left-hand side and variable levels related to perfectivity (perfective aspect, past tense, and complete actions) on the far right. The *y* axis represents a transitivity scale, with transitive events and the active voice in the lower part of the axes, and intransitive events and the middle voice[43] in the upper part.

The prefixes lie quite close to each other on the *x* axis, with *roz-* being the least "perfective" and *po-* being the most "perfective" (the most associated with perfective verbs). Linguists have postulated **perfectivisation** as the most general overarching function for all Polish verbal prefixes, which can explain why all three prefixes lie so close together on the *x* axis. *Po-*'s achieving the highest perfectivity score goes in line with ASPECTUAL being the most frequent sense in the dataset (apart from IDIOSYNCRATIC). If the most frequent function of *po-* were to render a verb perfective, it should also be high on the perfectivity scale.

As far as the *y* axis is concerned, we can see a substantial difference between *po-* and *przy-* on the one hand, and *roz-* on the other. *Po-* and *przy-* lie quite close together and are high on the "transitivity" scale (*przy-* is located somewhat higher than *po-*), preferring intransitive events. *Roz-*, on the other hand, lies much lower, slightly preferring transitive events. The most frequent type of sense for *przy-* in the dataset was APPROACH, which is inherently intransitive. An approach needs movement, and movement is usually conceptualised as an intransitive event, e.g. *przyjść* 'come' or *przyjechać* 'arrive driving'.

When we zoom in on the positions of particular points on the map, we can see that the prefixes *przy-* and *po-* lie close to each other in Quadrant I (upper right), which means that the analysis treats them as similar. *Przy-* is distinctly associated with verbs of perception, probably due to *przy-*'s being present in verbs such as

[43] Since the middle voice implies a clause in which the verb does not take a direct object, we can also see it as an indicator of intransitivity, albeit indirect.

przyglądać się 'observe intently', *przysłuchiwać się* 'listen closely'. Change verbs and intransitive situations go together on this plot and lie close to *po*- and *przy*-. This means that *przy*- and *po*- are roughly equally associated with intransitive verbs and change verbs. Change verbs usually convey intransitive situations in which change is construed as a spontaneous process, and where external forces are not profiled, e.g. *powiększyć się* 'grow larger' or *przytyć* 'put on weight'. As for the prefix *roz*-, it lies almost at the origin of the plot. Consequently, we cannot say that it is distinctly associated with any of the variable levels, because it is almost equally (dis)associated with all the levels.

The overall explained inertia for the analysis of all verb-related variables is 34.4%, which is a low result again. In an attempt to increase the amount of explained inertia, two more analyses were conducted, both with a reduced number of variables. The first analysis included only the WordNet-derived verb class, while the other analysis included all verb-related variables **except** the WordNet-derived verb class. WordNet verb class was removed because it was the single variable with the largest number of levels, so it might have been the one that introduced the most complexity and, consequently, the largest amount of unexplained inertia in the full analysis. Analysing the verb class alone did not increase the amount of explained inertia on the first two dimensions; in fact, the explained inertia dropped to 18.2%. Such a low score indicates a very poor reliability of the analysis, and this analysis will not be discussed in detail here (the plot of the analysis can be found in Appendix 9A). Figure 5.3. presents a plot for the analysis of all verbal variables except verb class.

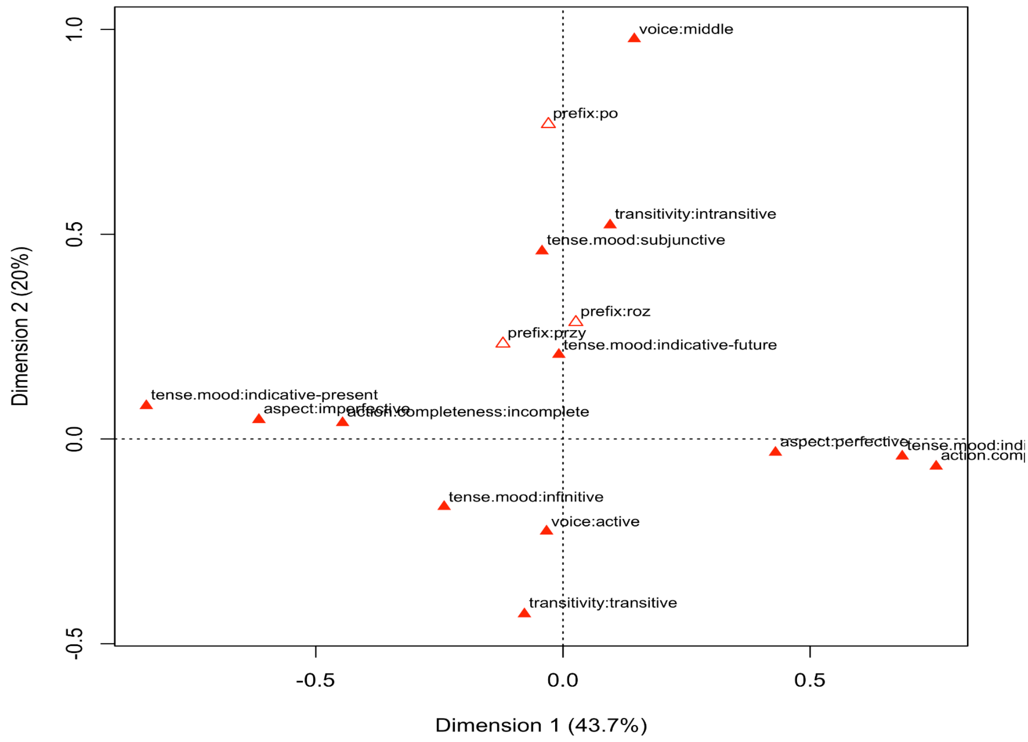


Figure 5.3. Correspondence analysis plot: all verb-related variables except verb class

In contrast to the verb-class-only analysis, the analysis of all verb-related variables **except verb class** exhibits much higher explained inertia: 63.7% on both dimensions, and can be considered as much more interpretable than the previous analyses. The results of this analysis seem very similar to the results of the analysis of all verb-related variables presented in Figure 5.2. The prefixes are not associated distinctly with many features: *roz-* and *przy-* seem to correlate with the indicative future tense and mood, and *po-* seems to correlate with the middle voice. The two axes in Figure 5.3. can be interpreted in the same way: the *x* axis represents perfectivity, while the *y* axis captures transitivity. All three prefixes, similarly to the first analysis, are situated very close to each other on the perfectivity scale; only their sequence differs – *przy-* is the least perfective, while *roz-* is the most perfective. In terms of transitivity, conversely to the first analysis, all tree prefixes lie on the "intransitive" part of the scale, which would indicate that all prefixes prefer intransitive situations.

If we sum up the three presented analyses in this section, the analysis of all verbal variables and the analysis of WordNet verb class only explained 33.4% and 18.2% of inertia respectively, which significantly limited their interpretability. With 63.7% of explained inertia (all verbal variables except verb class), the interpretability of the final analysis was significantly better, but the analysis did not yield many verb properties that we could reliably associate with any of the prefixes. If we compare the positions of prefixes against the *x* axis (perfectivity) and *y* axis (transitivity) between analyses, the picture is not coherent. For instance, in one analysis (all variables), *po-* and *przy-* prefer intransitive situations and *roz-* prefers transitive situations, while in another (all variables except WordNet verb class), all prefixes prefer intransitive situations. Interestingly, perfectivity did not turn out to be a reliable predictor for the prefixes in the correspondence analyses presented in this section. This result goes against the claim in most of the literature on Polish prefixes that transforming the verb from imperfective into perfective is one of the primary functions of all prefixes. The overall picture that emerges from the analyses of verbal variables presented above is that hardly any variable can be used to differentiate between the prefixes. Most variable levels are located in the quadrants other than the prefixes, which means that none of the prefixes are reliably associated with them.

5.3.2. Subject-related variables

After exploring verb-related variables, let us now investigate the nominal variables, that is the variables describing subjects and objects used with the prefixed verbs studied. First, we shall explore subject-related variables, and the analysis of object-related variables will be presented in the next section (5.3.3.).

The aim of analysing subject-related variables is to see whether the properties of the subject of a sentence can help differentiate between prefixes. Three different analyses were conducted, analogously to the analysis of verb-related variables: an analysis of all variables, an analysis of all variables except noun class, and an analysis of noun class only. The reasons for conducting three analyses were also similar to those in the case of verb-related variables: to see whether the removal of the variable introducing most variation (semantic noun class) would increase the amount of explained inertia and whether analysing noun class exclusively could explain the behaviour of prefixes. Only two analyses will be discussed and presented here: the analysis of all subject-related variables and the analysis of all subject-related variables except the WordNet-derived noun class. The analysis of only semantic noun class explained mere 15.4% of inertia on the first two dimensions. Such a low amount of explained inertia renders a correspondence analysis solution effectively random and, consequently, this analysis will not be discussed in detail (the plot of this analysis can be found in Appendix 9B). Figure 5.4. presents the plot of the first analysis, i.e. the analysis of all subject-related variables.

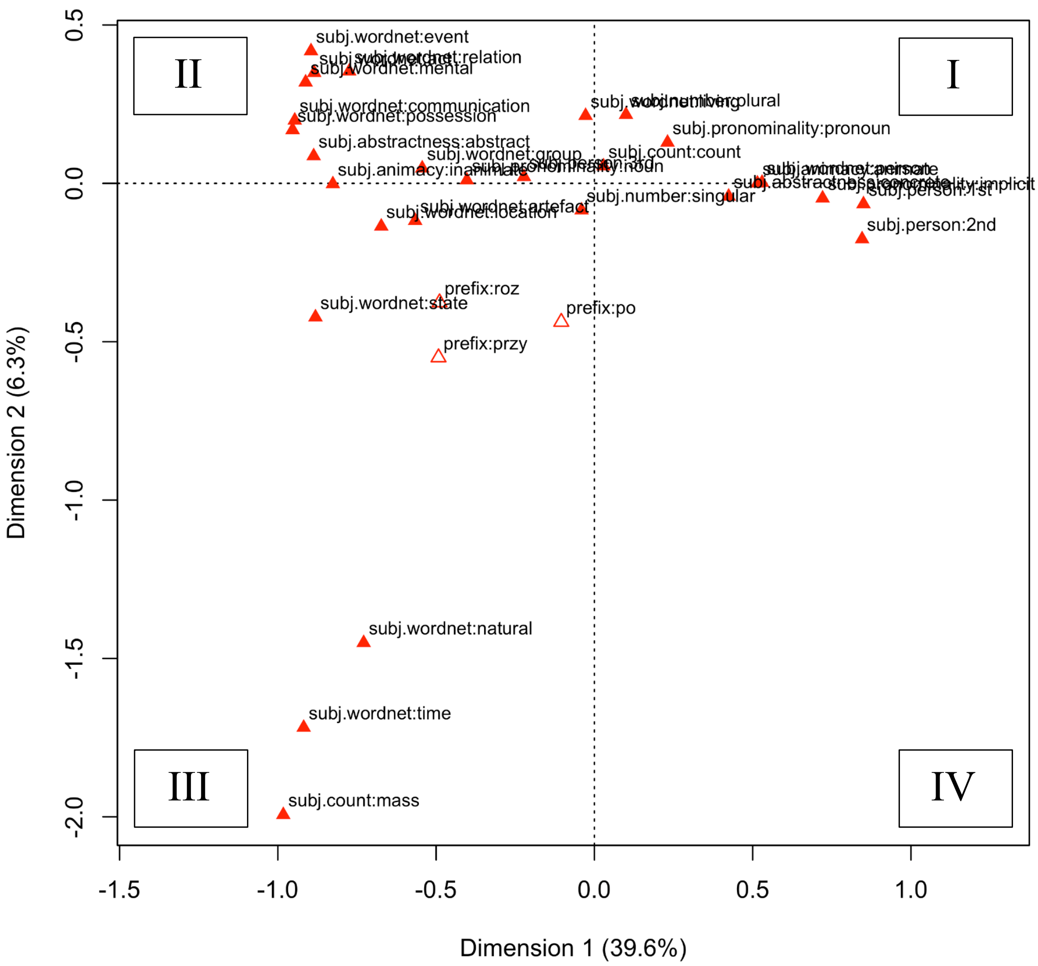


Figure 5.4. Correspondence analysis plot: all subject-related variables

The analysis in Figure 5.4. reached 45.9% in explained variation, which indicates moderately low reliability. Only the *x* axis can potentially be interpreted – it most likely represents an animacy/abstractness scale. On the left-hand side of the plot (in the negative area of the *x* axis), we can see the levels *abstract, inanimate,* as well as many inanimate or abstract WordNet properties, e.g. *communication, location, event* or *artefact.* The levels *concrete* and *WordNet.animate* are located in the positive region of the *x* axis on the right-hand side. If we take the *x* axis into consideration, it appears that all prefixes prefer inanimate or abstract subjects, because they are located in the negative region of the animacy/abstractness scale. *Przy-* and *roz-* exhibit a greater propensity to pair with inanimate or abstract subjects, because they appear farther to the left. All three prefixes are situated very close to each other in Quadrant III, while most of the variables lie either in other quadrants (mainly Quadrant I and Quadrant II) or near the origin of the plot. Such an arrangement of points on the plot means that the analysis could not find any properties that would differentiate the prefixes.

Analogously to the analysis of verbal variables, let us now inspect the analysis in which the WordNet-derived noun classes were removed and see whether it increases the amount of explained inertia and improves the readability of the plot. The plot for an analysis of all subject-related variables except noun class is presented in Figure 5.5.

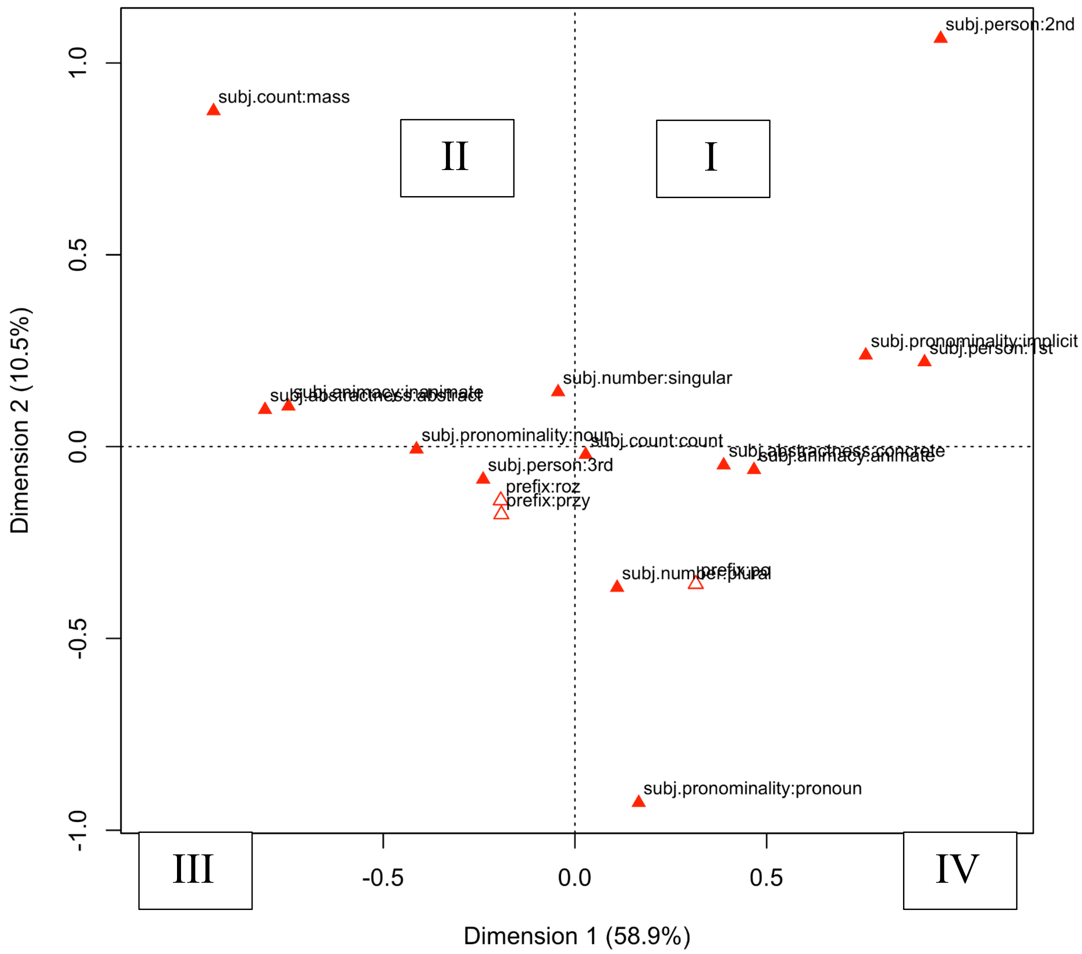


Figure 5.5. Correspondence analysis plot: all subject-related variables except noun class

The explained inertia statistics (69.4% in the first two dimensions) indicate that the analysis plotted in Figure 5.5. is more interpretable than the previous analysis. The noun class variable, having many levels, might have introduced too much complexity, reducing the amount of explained inertia of the analysis.

Similarly to the previous analysis, we can interpret the *x* axis as an animacy/abstractness scale, with the left-hand (negative) part of the axis indicating more abstract and less animate subjects and the right-hand part standing for more animate and more concrete subjects. *Po-* no longer prefers inanimate and more abstract subjects (contrary to the analysis plotted in Figure 5.4.), while *przy-* and *roz-* remain associated with lower animacy and higher abstractness.

When it comes to the position of individual points on the map, they are positioned differently than in the analysis of all subject-related variables presented in Figure

5.4. The prefixes lie in different quadrants: *przy-* and *roz-* are located in Quadrant III, while *po-* lies in Quadrant IV. That notwithstanding, the analysis sees *przy-* and *roz-* as nearly identical, because they were placed in virtually the same location. As with the previous analysis (all subject-related variables), it would be difficult to find properties that would reliably differentiate between prefixes – *po-* exhibits some association with plural subjects, while *przy-* and *roz-* prefer 3rd person subjects.

Overall, if we look at all three analyses of subject-related variables, it would be difficult to find properties that would help differentiate between the prefixes reliably. The only property that can be found in both interpretable analyses (the analysis of all variables and the analysis of all variables except noun class) is the animacy or abstractness of the subject. *Roz-* and *przy-* reliably correlate with inanimate and abstract subjects; the results for *po-* differ between analyses: in one, it correlates with inanimate and abstract subjects (see Figure 5.4.), while in the other, it correlates with animate and concrete subjects (see Figure 5.5.).

5.3.3. Object-related variables

The aim of analysing object-related variables as a group was to see whether the properties of the object of a sentence can help differentiate between prefixes. Different types of actions might require different types of objects: for instance, the DISTRIBUTIVE sense of *po-* might require a plural object, because the action is carried out on more than one object. Three correspondence analyses were conducted for object-related variables: (1) an analysis including all variables, (2) an analysis including all variables except noun class, and (3) an analysis of noun class only. The reasons for running three analyses were analogous to those in verb-related variables and subject-related variables. Analysis (3), just like the analogous analysis of subject noun class only from Section 5.3.2. had very low values of explained inertia (15.4% on the first two dimensions), and will not be discussed in more detail; the plot of this analysis can be found in Appendix 9C.

The plot in Figure 5.6. presents the analysis of all object-related variables. Two variables needed to be discarded: grammatical person, which did not contain sufficient variation (97.44% of all occurrences of this variable were 3rd person), and the syntactic type of object, which yielded many outliers and rendered the plot entirely unreadable (see Appendix 9D). Overall, the first analysis contained the following variables: animacy, abstractness, countability, number, pronominality, and noun class.

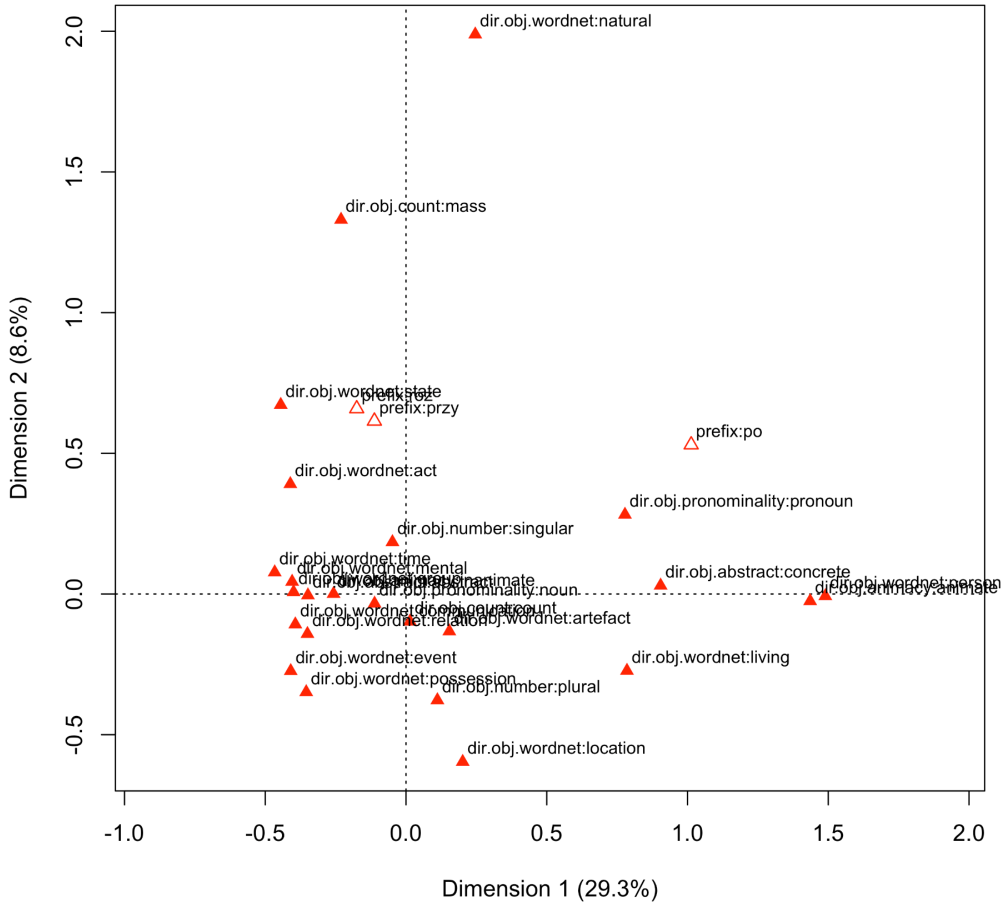


Figure 5.6. Correspondence analysis plot: all object-related variables

The analyses of all variables for verbs (see Figure 5.2.) and for subjects (see Figure 5.4.) did not reach a very high amount of explained inertia. This analysis is no different – it explains only 37.9% of inertia on the first two dimensions, which is quite a poor result. Because of this poor result, the analysis of all object-related variables can be regarded as only moderately interpretable.

It would be difficult to interpret the *y* axis in any meaningful way, but the *x* axis – similarly to the analyses of subject-related variables – corresponds to an animacy/abstractness scale. The left-hand side (the negative region) is more inanimate and abstract, while the right-hand side (the positive region) is more animate and concrete. *Przy-* and *roz-* prefer inanimate/abstract objects; *po-* is located much farther from the origin than *przy-* and *roz-* and exhibits a strong preference for animate/concrete objects.

In terms of levels distinctive for each of the prefixes, the analysis did not find many properties that would differentiate between the prefixes – most points are located either very close to the origin or in quadrants other than the quadrants the prefixes lie in. Two properties that both *roz-* and *przy-* seem to be associated with in this analysis are states and acts, whereas *po-* seems to be somewhat associated with pronominal objects.

Let us now remove the WordNet-derived noun class variable to see how this will change the amount of explained inertia. The plot of this analysis is presented in Figure 5.7.:

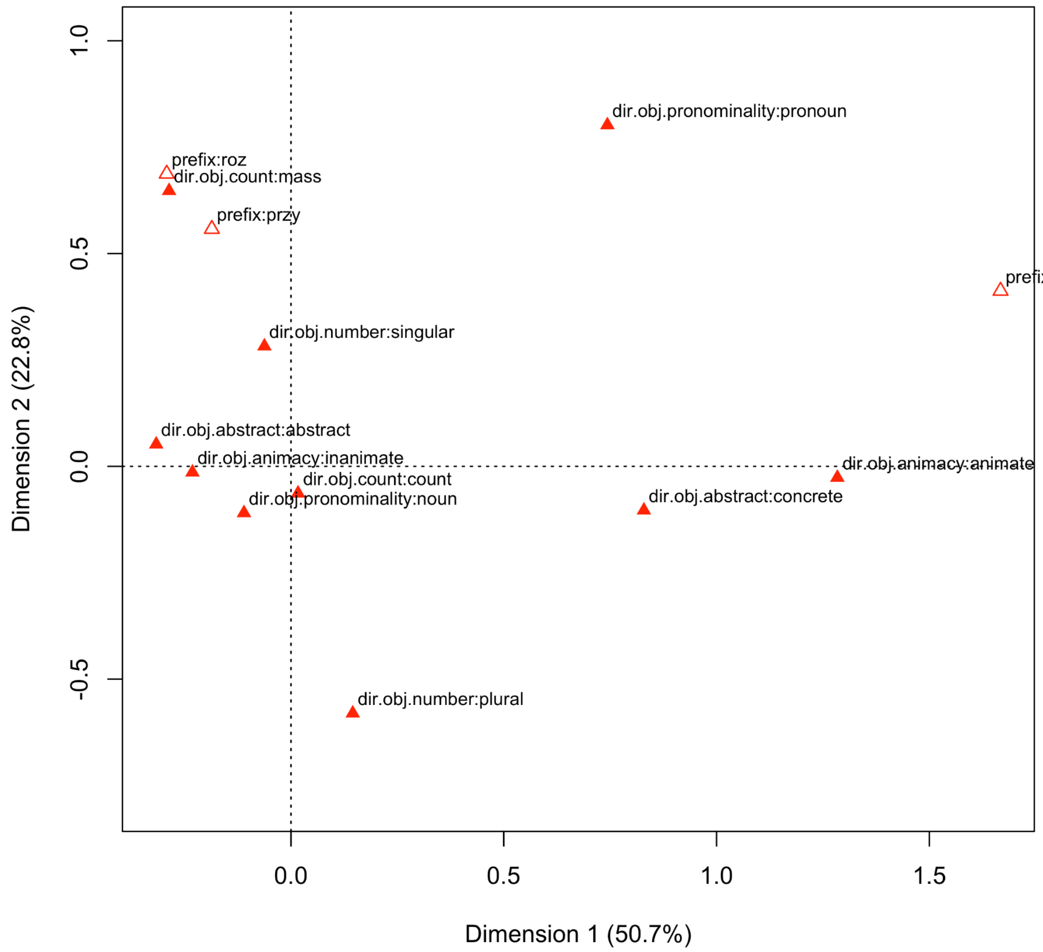


Figure 5.7. Correspondence analysis plot: all object-related variables except noun class

After the noun class has been removed, the amount of explained inertia on the first two dimensions improved substantially – it now amounts to 73.5%, and the analysis can thus be regarded as interpretable. The *x* axis can be interpreted identically as in the previous analysis: it stands for an animacy/abstractness scale. The left-hand side is more inanimate/abstract, whereas the right-hand side is more animate/concrete. Again, *po-* exhibits a strong preference for animate and concrete objects, while *przy-* and *roz-* slightly prefer inanimate and abstract objects. Contrary to the *x* axis, the *y* axis cannot be interpreted in any meaningful way. This analysis also sees *przy-* and *roz-* as similar, because the two prefixes lie next to each other on the plot. *Po-*, on the other hand, seems to be distinct from the other two prefixes, being located far away and in another quadrant. *Mass nouns* is the only individual variable level associated with any of the prefixes, namely *przy-* and *roz-*; the algorithm did not find any such properties for *po-*.

The three analyses of object-related variables are very similar to those of subject-related variables. It would be difficult to find properties that would help differentiate between the prefixes reliably. Again, the only property that can be found in both interpretable analyses (the analysis of all variables and the analysis of all variables except noun class) is the animacy or abstractness of the object. *Roz-* and *przy-* coherently correlate with inanimate and abstract objects, and *po-* correlates with animate and concrete objects. Other, more detailed, variables such as the WordNet noun class do not correlate reliably with any of the prefixes. This result, in combination with the similar result for subject-related variables, suggests that only one high-level feature could help speakers distinguish between prefixes. One feature of very high abstractness does not seem to be sufficient for the speakers to effectively differentiate between prefixes.

5.3.4. Clause-related variables

Lastly, clause-related variables were investigated to see whether properties such as negation or clause type can contribute to explaining the behaviour of the three prefixes. A procedure similar to the procedures employed in the previous analyses was applied also to clause-related variables. Three different analyses were produced: (1) an analysis of all variables, (2) an analysis of all variables except adverbials, and (3) a separate analysis of adverbials only. The purpose of splitting the analyses into three sub-analyses was to investigate whether high-level general properties (e.g. clause type) could be used to predict general categories (i.e. prefixes) or whether more specific variables do it better (e.g. adverbials). Adverbials received such a treatment because, just like WordNet-derived semantic classes, they are a granular semantic variable that can take a large number of levels and can thus introduce more complexity that it could explain.

The plot of the first analysis (the analysis of all clause-related variables) can be found in Figure 5.8. We can immediately see that the amount of explained inertia on the first two dimensions falls below 20% (to exactly 19.3%) – this is a very poor result, which renders the analysis uninterpretable. The axes also cannot be interpreted in any coherent way. Moreover, most points (i.e. variable levels) form one big ‘cloud’ around the origin of the plot, and all three prefixes are located in the same quadrant, which means that the analysis cannot explain the differences in the behaviour of the three prefixes.

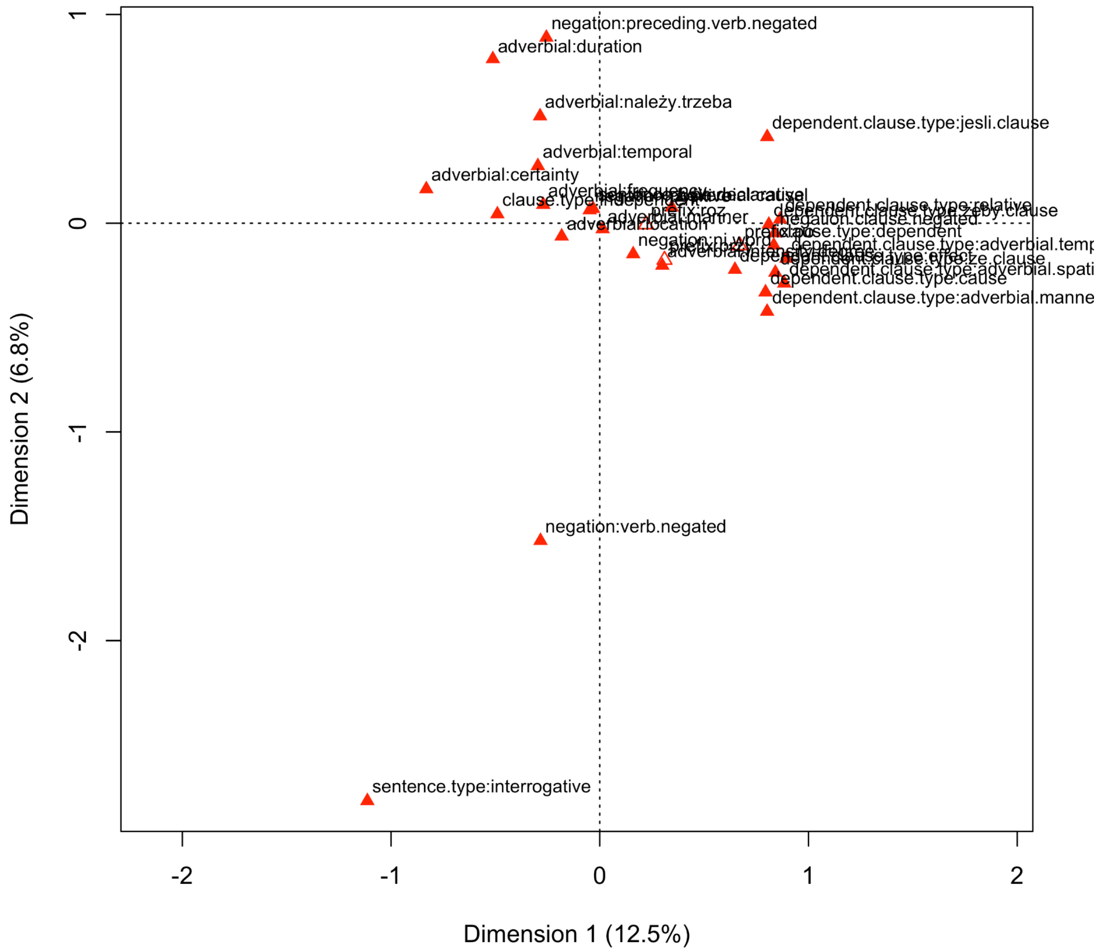


Figure 5.8. Correspondence analysis plot: all clause-related variables

When the *adverbial* variable was removed from the analysis (see Figure 5.9.), the amount of explained inertia improved slightly (27% on the first two dimensions). That notwithstanding, this analysis is still marred by the same problem as the analysis above: the axes cannot be interpreted, the prefixes are located close to each other in the same quadrant and near the origin of the plot, and most variable levels that could explain the behaviour of the prefixes clustered together to form a large cloud. Such an arrangement of prefixes and properties together with a low amount of explained inertia prevent this analysis from having any explanatory power as far as the behaviour of the prefixes is concerned.

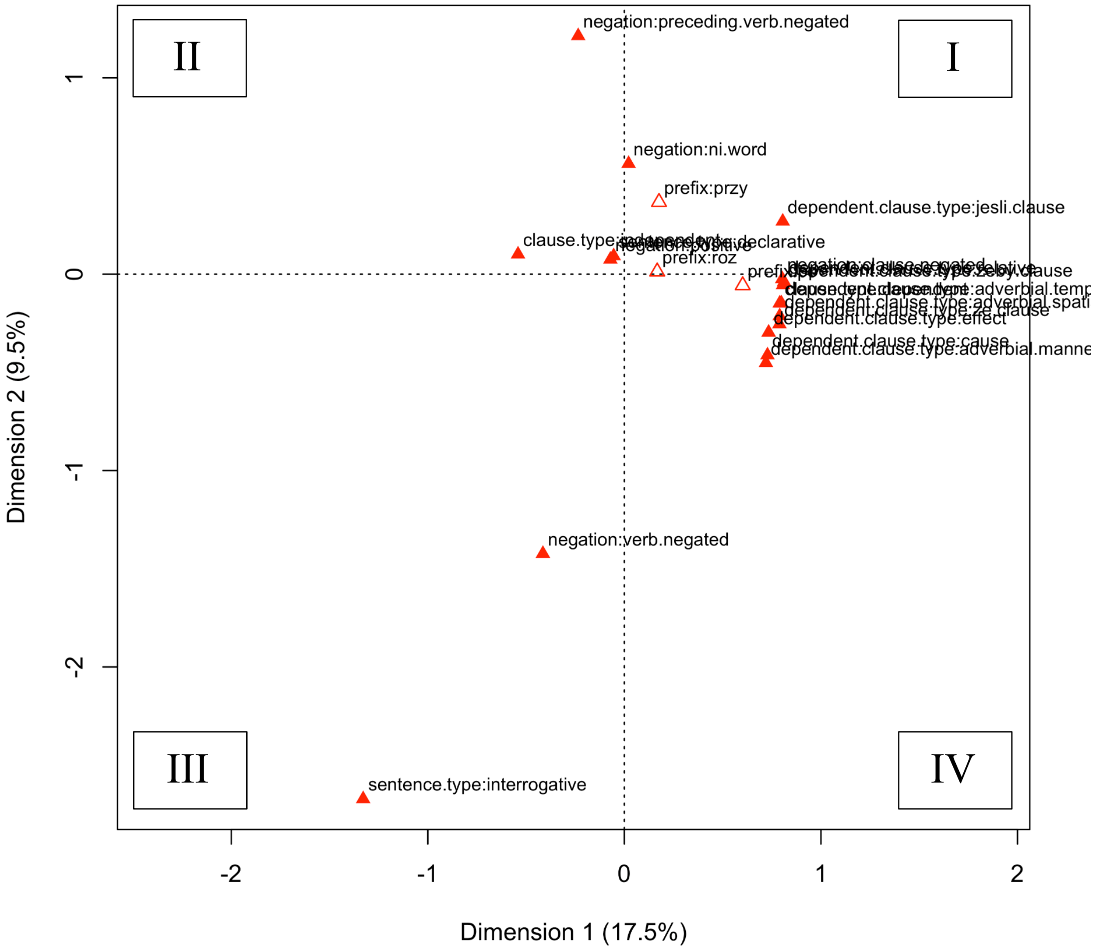


Figure 5.9. Correspondence analysis plot: all clause-related variables except adverbials

Finally, an analysis of adverbials only was conducted – it is presented in Figure 5.10. The explained inertia dropped by two percentage points (to 25%) in comparison to the analysis of all clause-related variables except adverbials presented in Figure 5.9. Similarly to the previous two analyses of clause-related variables (Figure 5.8. and Figure 5.9.), the axes defy interpretation. The properties do not form large clouds, unlike in the previous two analyses (Figure 5.8. and Figure 5.9.). The prefix *po-* lies almost exactly in the origin of the plot, which means that it is equally (dis)associated with all the properties in the analysis. In other words, none of the properties explain the behaviour of *po-*. *Przy-* and *roz-* are again located very close to each other in the same quadrant (Quadrant III), which indicates that the algorithm found the two prefixes similar. We can say that *przy-*

is distinctly associated with adverbials of certainty, and *roz-* with adverbials of location; both prefixes are equally associated with adverbials of duration.

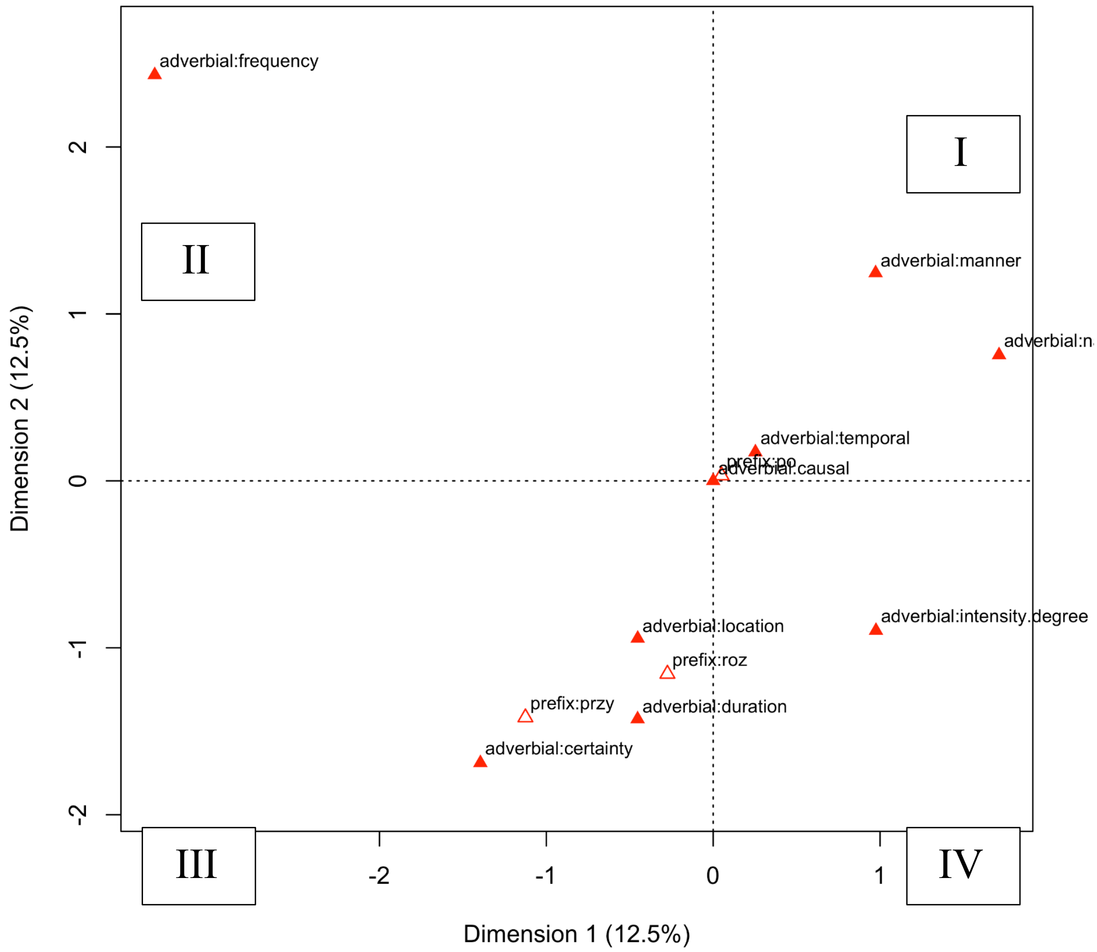


Figure 5.10. Correspondence analysis plot: only adverbials

Even though the plot in Figure 5.10. is much clearer than the plots for the two previous analyses, we must treat this result with utmost caution because the explained inertia amounts to only 25% on the first two dimensions, which indicates an analysis that is hardly interpretable.

In sum, the mappings in the three analyses of clause-related variables were not interpretable due to the low amounts of explained inertia. Such a result suggests that clause-related variables cannot be used by speakers to differentiate between prefixes.

5.3.5. Morphologically marked variables vs granular semantic properties

So far, the analyses were grouped according to the function of particular elements in the sentence (verbs, subjects, etc.) to see whether the properties of those elements can be used to differentiate between prefixes. This section will discuss two more analyses, which have a different purpose from the previous ones – their ultimate aim is to see whether the behaviour of prefixes can be explained better with general morphological variables or low-level granular semantic variables. The first analysis will consider only those properties that are overtly marked morphologically, while the second analysis will include only variables that pertain to low-level semantics (WordNet semantic classes and adverbial types).

As far as overtly marked variables are concerned, speakers might find it easier to notice overtly marked properties and use them to disambiguate between the prefixes. In consequence, overtly marked variables might be better predictors of the prefixes' behaviour – this is why they received a separate analysis in this study. The "overtly marked" group encompasses the following variables: tense/mood, aspect, voice, subject number (singular vs plural), and direct object number. The plot of the analysis can be found in Figure 5.11.:

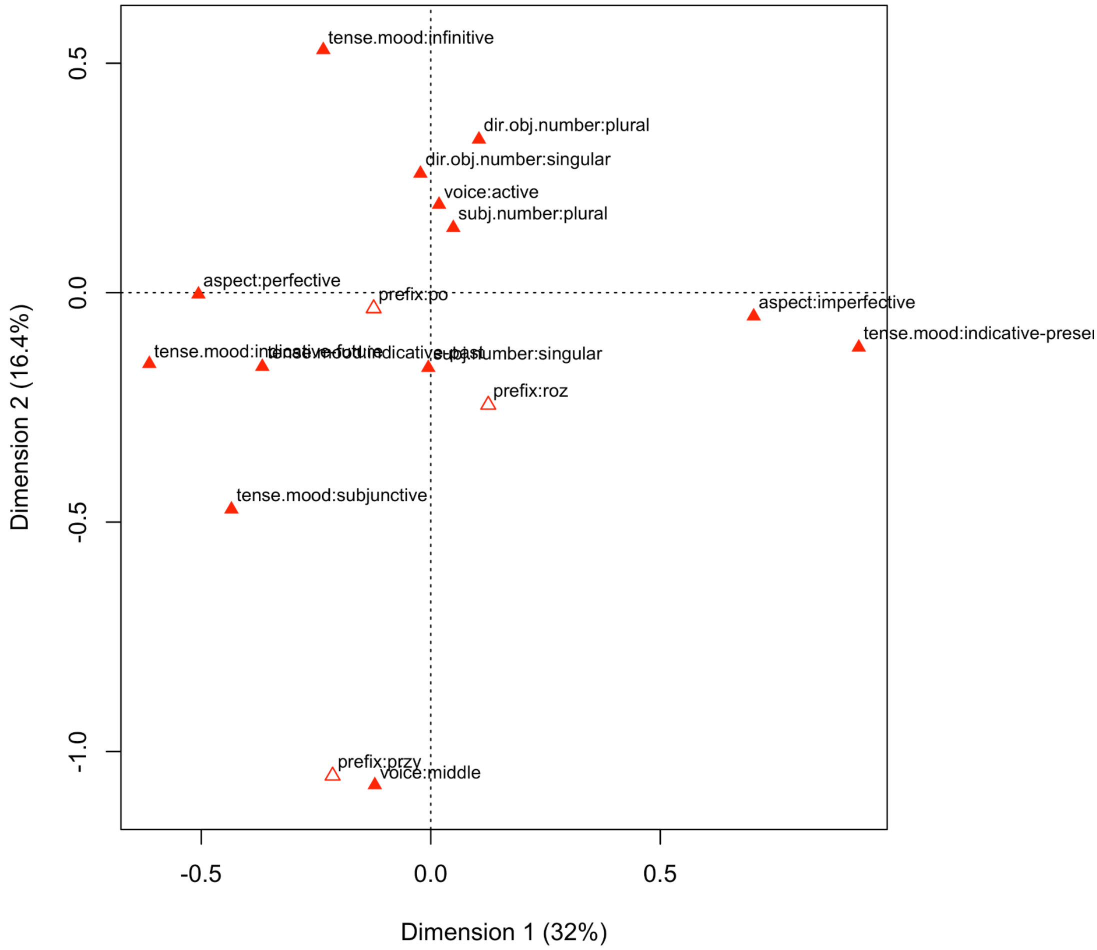


Figure 5.11. Correspondence analysis plot: all overtly marked variables

Overall, the analysis explains 48.4% of variance on the first two dimensions. Compared to some previous analyses (e.g. analyses of clause-related variables), it is a moderately good result, but more than a half of the variance remains unexplained.

Analogously to the analyses of verbal variables, we can interpret the horizontal axis (the *x* axis) as a perfectivity-imperfectivity dimension and the vertical axis (the *y* axis) as a voice dimension. In the case of the *x* axis, the left-hand side of the plot corresponds to increasing perfectivity, whereas the right-hand side indicates increasingly imperfective situations. As for the *y* axis, the top half contains the active voice, whereas the bottom half contains the middle voice. In this analysis, *po-* and *przy-* lie on the left-hand side of the plot but not far from the middle of the axis, so they slightly prefer perfective events. *Roz-* lies on the other side and thus (again slightly) prefers imperfective events. The differences between the prefixes in terms of perfectivity are not large – they are all located quite close to each other on the horizontal plane, i.e. the *x* axis. As far as voice is concerned, *przy-* and *roz-* prefer the middle/reflexive voice to some extent, but *przy-* lies farthest to the bottom of the plot and it is also located very close to middle/reflexive voice. *Po-* displays hardly any preference for either of the voices, because it lies almost in the middle of the *y* axis.

We can see that it would be difficult to associate any of the markers with any particular feature: *po-* lies almost at the origin of the plot. The only potential variable *roz-* could be associated with are singular subjects, but singular subjects are equidistant to *po-* and *roz-*, hence both prefixes are equally (dis)associated with singular subjects. Only *przy-* seems to correlate with a property, namely the middle/reflexive voice (i.e. the marker *się*).

We could say that granular semantic variables – such as verb or noun classes – are the opposite of the morphologically marked variables analysed above. Noun or verb classes pertain to quite concrete properties, while morphologically marked variables represent high-level general categories such as aspect. Let us now compare the analysis of morphologically marked variables with an analysis of **all** low-level semantic variables: verb class, subject nominal class, object nominal class, and adverbials. The plot of the analysis is presented in Figure 5.12.:

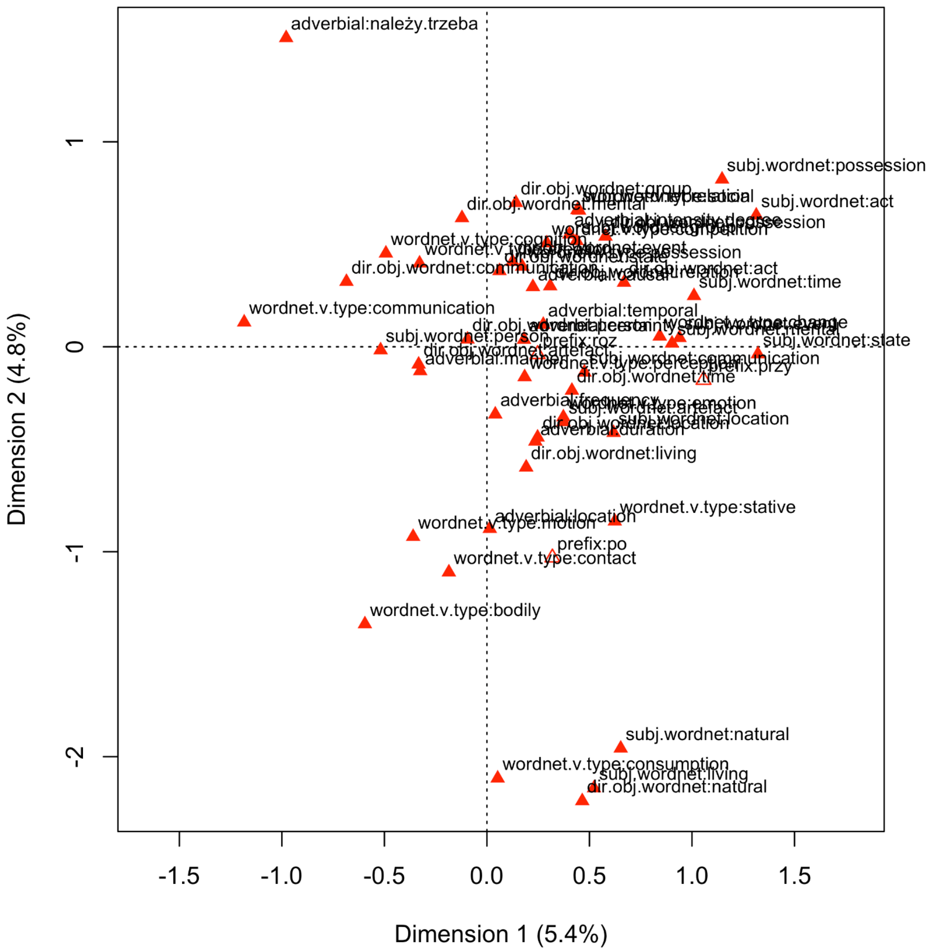


Figure 5.12. Correspondence analysis plot: all low-level semantic variables

The most striking feature of the above plot is that the explained inertia on the first two dimensions amounts to only 10.2% – this constitutes the lowest result of all conducted analyses. An amount of explained inertia this low means that so much variance remains unexplained that the arrangement of points on the plot is practically random. With such an arrangement of points on the map, we must not draw any conclusions on the basis of the analysis of all low-level semantic variables, because no clear patterns can be discerned. Taking both analyses (of all high-level morphologically-marked properties and all more concrete semantic properties) into consideration, we can say that high-level general properties better explain the behaviour of the prefixes than more concrete semantic properties.

5.4. Discussion

The most striking feature of the conducted analyses was that many of them explained minuscule amounts of variation – the levels of explained inertia stayed very low. The explained inertia dropped to extraordinarily low levels in each analysis including granular semantic variables (verbal and nominal semantic classes). Low levels of explained inertia constitute a sign that the algorithm could not find any clear patterns in the data. It seems that those variables introduced much more variance that they were able to explain – instead of telling more about the phenomenon, they rendered the analyses less interpretable. Any result with so little inertia explained should be treated with extreme caution and, most likely, would not extrapolate beyond the dataset the analysis was performed on. In other words, the analyses with the lowest levels of inertia do not allow us to draw conclusions about the language "in general". Even in the analyses with a decent level of explained inertia, most points (i.e. variable levels) appeared in quadrants other than the prefixes or far away from them.

The only properties that offered any explanation in the more reliable analyses (those with explained inertia greater than 60%) – but still did not correlate strongly with any particular prefix – were verbal transitivity, aspect, and the animacy of subject/object. Those three properties formed scales along which the axes in the correspondence analysis plots could be interpreted; they did not act as binary variables (e.g. transitive/intransitive) but rather as spectra of transitivity, aspect, or animacy. The differences between prefixes in terms of the scalar properties were not large, but it appears that *roz-* displays a slight preference for more perfective situations, *po-* does not display any preference, and *przy-*is slightly more likely to appear in imperfective situations. In terms of transitivity, all prefixes preferred intransitive contexts: *po-* displayed the greatest preference, while *przy-* and *roz-* remained on a par in this respect. In terms of subject-related variables, *przy-* and *roz-* were virtually identical, and they appear to have a slight preference for inanimate subjects, while *po-* displays approximately the same amount of preference but for animate subjects. The same structure of preference holds also for object-related variables: *przy-* and *roz-* prefer more inanimate objects, while *po-* prefers animate objects.

Overall, none of the conducted analyses showed any distinctive properties that would reliably distinguish one prefix from the others. This means that none of the prefixes is associated with any semantic classes of verbs, subjects or objects, neither does any of them correlate with any more general properties, e.g. nominal number or verbal aspect. If we compare the results of the present analysis with the results of the correspondence analysis of the corpus data on Polish reflexive verbs (Section 3.2.5.1.), we will see that the latter did indicate that some properties might be associated with either marker. The fact that the correspondence analysis of the data on reflexives indicated some properties justified running further, confirmatory, statistical analyses for these data (logistic regression and conditional inference trees). In contrast, finding no properties reliably associated with any of the prefixes does not warrant conducting confirmatory statistical analyses; hence, no further statistical analyses of the data will be pursued.

5.5. Interim conclusions

The primary aim of the corpus study was to investigate whether speakers of Polish could build maximally general constructions for the prefixes *po-, przy-,* and *roz-*. The correspondence analyses showed that hardly any of the properties for which the data were coded correlated with the prefixes. Consequently, it might be difficult for speakers to differentiate between the prefixes as a whole, because there are no prototypical contexts in which each of the prefixes would appear. In other words, no properties of the contexts that the prefixes appeared in would justify the existence of general constructions for the prefixes. It could also potentially mean that the behaviour of the prefixes could not be explained with **this set of variables**, but there might still exist some variables which have not been found yet that would be able to explain the distribution of the prefixes.

This corpus study explored only the question whether speakers could build maximally general constructions for prefixes based on the input they receive, and it did not find evidence in support of this hypothesis. Researchers, however, have also postulated a number of different senses for the prefixes *po-, przy-,* and *roz-*. Perhaps, if speakers might not be able to build maximally general constructions for prefixes, could they build (less general) constructions for the different senses of prefixes? This question will be investigated empirically in the next chapter.

Chapter 6: Polish prefixes: sentence-sorting experiment

6.1. Introduction

In the corpus study on three Polish prefixes presented in Chapter 5, we saw that the data indicate that native speakers might not be able to build coherent usage-based constructions for the prefixes *po-, przy-,* and *roz-*. The corpus study considered prefixes 'as a whole', that is, its aim was to investigate whether native speakers build **one** maximally general construction per prefix. The fact that the data did not support those maximally general constructions does not imply, however, that speakers do not build less general constructions for the different **senses** of the prefixes postulated in the literature. To give an example, speakers might not build a construction for *po-* as a whole, but they might build the constructions for the DEPARTURE, DISTRIBUTIVE, SOMEWHAT, and COVER sense of the prefix.

If native speakers of Polish develop general constructions for particular prefix senses, they should be able to perceive similarities between different verbs with the same sense of a prefix. For instance, speakers should see as similar the following verbs with the DISTRIBUTIVE sense of the prefix *po-*: ***po**kąsać* 'bite (many things/people)', ***po**myć* 'wash (multiple things)' and ***po**zamykać* 'shut/lock (many doors/windows/etc.)'. In other words, speakers should be able to classify the different verbs with the same sense of a prefix as belonging to the same category. If speakers, on the other hand, did not perceive prefixed verbs with a prefix with the same sense as similar, we could surmise that they do not build general constructions for the different senses of a given prefix. In that case, they would be likely to rely on more specific constructions for each prefixed verb.

In the literature on Polish prefixed verbs, researchers have postulated a range of different senses for the prefixes. Nevertheless, to the best of my knowledge, no empirical research has been conducted so far to investigate whether native speakers of Polish really build these general constructions. This chapter will present the results of a sentence-sorting experiment that seeks to bridge this gap and empirically investigate whether speakers of Polish can build general constructions for the different prefix senses.

6.2. Could speakers build the different senses of prefixes? A sentence-sorting experiment

Similarly to the corpus study reported in Chapter 5, the experiment will involve three Polish prefixes: *po-, przy-,* and *roz-*. The primary hypothesis of the experiment is that if participants reliably group sentences according to the prefix senses, they might have constructions for those prefix senses, because they perceive sentences with verbs with the same sense of the prefix as similar. In the opposite case, if participants group the sentences according to other criteria, they might not have built constructions for those prefix senses. For more details about the experimental method, please consult Section 4.2.1.

In a nutshell, in a sentence-sorting experiment, subjects are presented with a number of sentences (containing constructions under investigation) and asked to group sentences into a number of bins. The results of the experiment are subsequently analysed to see how participants grouped the sentences. If sentences with the same construction were put into the same bin, we could surmise that participants perceived the sentences as similar and, consequently, that they might have a salient category in their minds for the proposed construction.

6.2.1. Procedure

Participants were presented with sentences that included verbs with the prefixes *po-, przy-*, and *roz-* in each of the senses annotated for in the corpus study in Chapter 5. Similarly to the experiment discussed in Chapter 4, participants were asked to sort the sentences into a predefined number of "bins" according to whichever criteria they found relevant – no criteria were suggested to the participants. The participants could put any number of sentences in any chosen bin, and the bins were given unsuggestive labels: *Grupa 1* 'Group 1', *Grupa 2* 'Group 2' and so on. No time limits were imposed on participants. Moreover, participants were allowed to reconsider their choices as many times as they wished, until they chose to submit the questionnaire, by which time they could no longer alter their answers. To sum up, no constraints were imposed on the participants as to how they should perform the sorting. Full task description can be found in Appendix 1A.

The number of bins in each study corresponded to the number of senses each prefix could take – one bin for each sense. The senses for each prefix were identical to the ones annotated for in the corpus study presented in Chapter 5 (see Table 5.1. for more details): COVER, DEPARTURE, DISTRIBUTIVE, and SOMEWHAT for *po-*; APPROACH, COVER, FIT, and SOMEWHAT/INTENSITY for *przy-*; and DISPERSION, OPPOSITION, and INTENSITY for *roz-*. Consequently, the experiments for *po-* and *przy-* contained 4 bins, while the one for *roz-* contained 3 bins. Two versions of the experiment were created for each prefix: one version included 3 sentences per sense of a prefix, and the other version contained 5 sentences per sense. Two versions were created in order to investigate whether there was a "critical mass" of sentences to sort that would enable participants to perceive similarities (and differences) between the stimuli and group them into bins corresponding to prefix senses. In other words, two different experiments per prefix were conducted in order to see whether the number of sentences in an experiment would affect participants' choices. For instance, it might have been easier to notice the similarities between sentences containing verbs with the same sense of the prefix if the number of sentences was larger and, consequently, put the sentences into groups in accordance with the sense of the prefix.

Participants were also asked a number of questions related to demographics and reading habits. Not all questions were mandatory so as not to deter participants from completing the questionnaire. People are wary of disclosing large amounts of personal information (e.g. age or place of residence), and a too large number of questions might make potential participants quit before they even begin the experiment itself. The level of education attained, gender, and readership-related questions were chosen as mandatory[44] because they have been shown to have an impact on speakers' use of language (see e.g. Dąbrowska 2008a; Leaper 2014). In

[44] The experiment for *roz-* with 5/6 sentences per sense did not contain the question about education. This experiment included an early version of the questionnaire with socio-demographic questions, which was updated for the five remaining experiments.

the non-mandatory questions, participants were asked about their age, place of residence, foreign languages spoken, and their university major (if they have been in university-level education). The study was given ethics approval by the University of Sheffield.

6.2.2. Stimuli

As mentioned in Section 6.2.1., six different versions of the experiment were created – two for each prefix. The first version for each prefix contained 3 sentences for each sense of a prefix, which amounted to either 9 sentences (in the case of *roz-*) or 12 sentences in the case of *po-* and *przy-*. The second version contained 5 or 6 sentences for each sense of a prefix (17 for *roz-* and 20 for *po-* and *przy* each)[45]. *Roz-* contained 6 sentences for each sense (while *po-* and *przy-* contained 5) because the discrepancy between the overall number of sentences among the conditions would have been too great if each prefix had received 5 sentences for each sense. In such case, *roz-* would have contained 15 sentences, while *po-* and *przy-* would have contained 20, and the aim was for the experiments to be of a similar size. Two versions of the experiment for each prefix were created with the aim of seeing whether the number of sentences would have an impact on participants' ability to group the sentences together according to the prefix sense, similarly to the sentence-sorting experiment on Polish reflexive verbs reported in Chapter 4.

The verbs for sentences were derived from Śmiech (1986) – each sense of each prefix was retrieved, and verbs were chosen randomly from the example verbs listed by Śmiech. The stimuli for the experiments were always single-clause sentences. No other standardisation measures were imposed – the position of the verb in the sentence, type of subject or object, or tense were not uniformised. Thus, the sentences contained some variety and "random noise", which prevented them from sounding excessively artificial and introduced potential confounding criteria that participants could have used for sorting. Analogously to the sorting experiment on the different senses of *się* discussed in Chapter 4, not controlling for

[45] In the case of *roz-*, 18 (6 per sense) sentences were initially developed for the experiment and included in the software, but due to a software glitch participants only received 17 sentences.

grammatical features (other than clause number) was a principled design choice. The aim was to give participants many possible criteria for grouping so that if they chose to group sentences according to the senses of a prefix, it would mean that the categories for those senses are strong in participants' minds – strong enough to be chosen over other possible criteria. An example sentence is presented in (6.1):

(6.1) *Przez* *nieuwagę* *roz-gniót-ł* *stopą*

through inattention **roz-**squeeze-PST.MASC foot

ślimaka.

snail

'He inadvertently crushed a snail with his foot'

If participants decided to sort the sentences according to prefix senses, it would mean that the general prefix constructions are strong in the minds of the participants, because they were able to perceive similarities between the sentences despite many other (confounding) criteria they could have used for sorting.

Some verbs chosen from among the examples provided by Śmiech (1986) could also carry other senses of prefixes – not only those defined by Śmiech. For instance, the verb *podreptać* 'dawdle, walk very slowly', apart from the DEPARTURE sense, could also convey the SOMEWHAT sense; see the examples in (6.2.):

(6.2.) a. *Po porażce powoli* ***podreptali*** *zrezygnowani* ***do domu***.

'After the defeat, they **dawdled home** dispiritedly.'

b. ***Podreptała chwilę*** *i z powrotem usiadła.*

'She **moped around** for **a while** and then sat down again.'

The sentence 6.2a appeared originally as one of the stimuli in the experiment (see Appendix 3A and 3B) while sentence 6.2b is presented here for the sake of comparison. The two sentences contain different adverbials, which precisely disambiguate the potential senses the verb could carry. In the case of 6.2a, we can see an adverbial of place, *do domu* 'home', which enforces the DEPARTURE interpretation of the verb; the sentence 6.2b, on the other hand, contains an adverbial of duration, *chwilę* ‘for a while’, which indicates short duration and thus enforces the SOMEWHAT interpretation. In sum, the particular sense of a prefix can be disambiguated by the context the verb appears in.

The stimuli sentences contain more verbs – other than *podreptać* – that can convey more than one sense of a prefix. All the verbs used in the experiment that could carry more than one sense of the prefix are presented in Table 6.1.

prefix	verb	sense assigned by Śmiech (1986)	alternative sense	permits either interpretation?
po	*podreptać*	departure	somewhat	NO
po	*pokuśtykać*	departure	somewhat	NO
po	*pogalopować*	departure	somewhat	NO
po	*pochlapać*	cover	distributive	YES
po	*pomalować*	cover	distributive	YES
po	*posmarować*	cover	distributive	NO
po	*pomazać*	cover	distributive	NO
przy	*przydeptać*	cover	somewhat	NO
przy	*przykręcić*	attach	somewhat	NO

Table 6.1. Prefixed verbs used in the sentence-sorting experiment that permit more than one interpretation

Every attempt was taken in order to create stimuli that would permit only one interpretation of a verb. Nevertheless, the prefix *po-* in two sentences (with the verbs *pomalować* ‘paint’ and *pochlapać* ‘splash’) could be interpreted in terms of either the COVER[46] sense and the DISTRIBUTIVE sense, which is also indicated in Table 6.1. Let us now look in more detail at the experimental sentences that contained the two verbs (example 6.3.).

(6.3.) a. **Pomalowała** wszystkie **paznokcie** na śliwkowo.

*She **painted** all her **nails** purple.*

b. Skacząc do wody **pochlapała wszystkich** wokół basenu.

*Jumping into the pool, she **splashed** water on **everyone** around.*

Both sentences in example (6.2.) contain plural objects – *paznokcie* ‘nails’ and *wszystkich* ‘everyone’, respectively (marked in bold). A plural object paired with a

[46] COVER was the sense originally assigned by Śmiech (1986)

po-verb in those sentences permits a **DISTRIBUTIVE** reading of the verb, in addition to the COVER reading proposed by Śmiech (1986). If the object were singular, the **DISTRIBUTIVE** reading would not be permitted, because there would not be enough entities for the action to be distributed over. The fact that the two verbs permit alternative interpretations of the prefix ***po-*** will be taken into account when discussing the results of the experiment.

6.2.3. Participants

The questionnaires were delivered electronically via the Qualtrics[47] platform and distributed through an anonymous link posted on the official Facebook groups for the University of Warsaw, Warsaw School of Economics, and the University of Gdansk. The experiment was delivered in three batches over the period from 19 April 2017 to 1 July 2017. The system randomly assigned one version of the experiment to each participant[48].

Overall, tallied over all conditions, 1604 respondents took part in the study (1313 females). An overwhelming majority of participants were either in university education (41.62%), had graduated from a university (51.10%), or had taken a university course but had not graduated (3.19%) – the total number of respondents with at least some tertiary education amounted to 96.41%. The proportions of educational attainment were calculated based on the total number of 1002 participants who disclosed information about their education. The mean participant age was 28.57 years (median = 26, sd = 8.35, min = 15, max = 80).

6.3. Results

The sample sizes for the experimental conditions differ to a large extent – the largest sample contains 601 responses, while the smallest one contains only 164[49]. In order to make the analyses of all studies comparable, a random sample of 160 responses was drawn for each study, and those smaller samples will constitute the datasets for further analysis.

[47] http://www.qualtrics.com
[48] The odds of drawing each experiment were equal.
[49] The sample sizes were: 601, 203, 227, 175, 164, and 234 responses.

The data were analysed by means of hierarchical agglomerative cluster analysis (for more details on the technique, see Section 4.3.1.). Qualtrics does not generate data that could be immediately used for clustering, and the data needed to be transformed. Pairwise co-occurrences of sentences (and verbs) were calculated per participant – a sentence was classified as co-occurring with another sentence if they appeared in the same bin. The procedure for calculating co-occurrences was repeated for every participant, and, eventually, all co-occurrences were added together and stored in a form of a co-occurrence matrix (see Table 6.2.).

VERB	*rozszyfrować*	*rozpędzić się*	*rozgnieść*	*rozlecieć się*	*rozejść się*	*rozśpiewać się*	*rozpakować*	*rozplątać*	*rozchorować się*
rozszyfrować		63	9	56	39	82	44	26	5
rozpędzić się	63		17	113	11	54	7	17	22
rozgnieść	9	17		34	37	10	38	87	111
rozlecieć się	56	113	34		9	45	5	18	36
rozejść się	39	11	37	9		64	129	24	32
rozśpiewać się	82	54	10	45	64		54	16	16
rozpakować	44	7	38	5	129	54		29	31
rozplątać	26	17	87	18	24	16	29		92
rozchorować się	5	22	111	36	32	16	31	92	

Table 6.2. Co-occurrence matrix of *roz*-verbs in the sorting task (3-sentence experiment)

For example, we can infer from the table above that the participants put the sentence with the verb *rozpakować* in the same bin as the sentence with the verb *rozejść się* 129 times, while for *rozlecieć się* and *rozejść się*, the number of co-occurrences was only nine. The above matrix and the remaining co-occurrence matrices (i.e. the matrices for the other 5 experiments) were used as the basis for the calculation of the distance matrices needed for the cluster analyses, which will be presented in the following section.

6.3.1. Choosing the best clustering solution

For each dataset, three clustering algorithms were used: complete linkage, average linkage, and Ward's method. The single-linkage method was not considered as it tends to produce elongated cigar-shaped clusters that often contain dissimilar items, which would be an undesired effect. Ward's method and the complete-linkage method produce compact spherical clusters that are clearly separated; they are also not prone to the distortions of the single linkage method. Both methods (complete-linkage and Ward), however, are sensitive to outliers and might sometimes put similar items in different clusters, which would be a suboptimal outcome. The average linkage method is not so sensitive to outliers, but it does not always produce clusters of the same size. Optimally, we would like to have clearly delineated clusters of roughly the same size, because the experiment contained an identical number of sentences for each sense of a prefix. Each of the used methods implies a trade-off between some desired and undesired characteristics, and the choice of the final solution for each experiment will depend on the solutions' performance measured by the agglomerative coefficient. The clustering algorithms were run on three different dissimilarity matrices: Euclidean, Manhattan, and Canberra (for short characteristics of each type of dissimilarity matrix see Section 4.3.1.). In the end, nine different clustering solutions for each experiment were produced. The solutions were then compared using the agglomerative coefficient (see Section 4.2.2. for more details on the agglomerative coefficient) in order to choose the best solution for each dataset. The coefficient values for all experiments and all solutions can be found in Table 6.3.:

3 SENTENCES EACH			
PREFIX	WARD		
	Euclidean	Manhattan	Canberra
po-	0.74	**0.78**	0.69
przy-	**0.87**	**0.87**	0.74
roz-	0.9	**0.91**	0.83
	COMPLETE LINKAGE		
	Euclidean	Manhattan	Canberra
po-	0.62	0.69	0.6
przy-	0.79	0.8	0.63
roz-	0.85	0.86	0.75
	AVERAGE LINKAGE		
	Euclidean	Manhattan	Canberra
po-	0.54	0.59	0.48
przy-	0.77	0.76	0.54
roz-	0.83	0.84	0.7

5 SENTENCES EACH			
PREFIX	WARD		
	Euclidean	Manhattan	Canberra
po-	0.89	**0.9**	0.83
przy-	**0.95**	**0.95**	0.91
roz-	0.82	**0.84**	0.81
	COMPLETE LINKAGE		
	Euclidean	Manhattan	Canberra
po-	0.8	0.81	0.71
przy-	0.89	0.9	0.81
roz-	0.68	0.71	0.67
	AVERAGE LINKAGE		
	Euclidean	Manhattan	Canberra
po-	0.76	0.76	0.64
przy-	0.86	0.87	0.76
roz-	0.6	0.63	0.58

Table 6.3. Agglomerative coefficient values for all clustering solutions for all experiments

Table 6.3. shows that the Ward's algorithm on a Manhattan distance matrix was the best solution for every experiment. For both experiments involving *przy-*, Ward's clustering solution on a Manhattan matrix tied with Ward's clustering on a Euclidean matrix. Ultimately, a Ward's clustering on a Manhattan distance matrix was chosen for each experiment, so that the solutions are consistent.

The output of hierarchical agglomerative clustering is graphically represented as a dendrogram (see Figure 4.1. in Chapter 4). In a dendrogram, smaller clusters merge together to form larger ones, until no more mergers can be made and the data become one large cluster. The researcher must choose the optimal number of clusters, or, in other words, cut the tree at some height and thus ignore any further mergers. No hard-and-fast rules are available to automate the selection process. Nevertheless, a useful heuristic that can aid the researcher in choosing the number of clusters are silhouettes (Rousseeuw 1987). Silhouettes provide a visual representation of every cluster "based on the comparison of its tightness and separation … [, and] the average silhouette width might be used to select the 'appropriate' number of clusters" (Rousseeuw 1987: 53).

In Figures 6.1a. and 6.1b., two silhouette plots for the five-sentence experiment are compared. Figure 6.1a. presents the silhouette widths for a configuration with two large clusters, whereas Figure 6.1b. displays a silhouette plot for a configuration

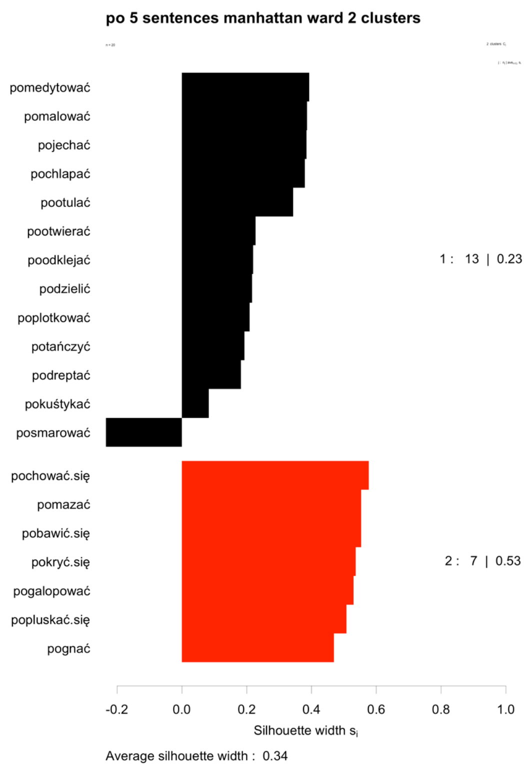


Figure 6.1a. Silhouette plot for a configuration with 2 clusters

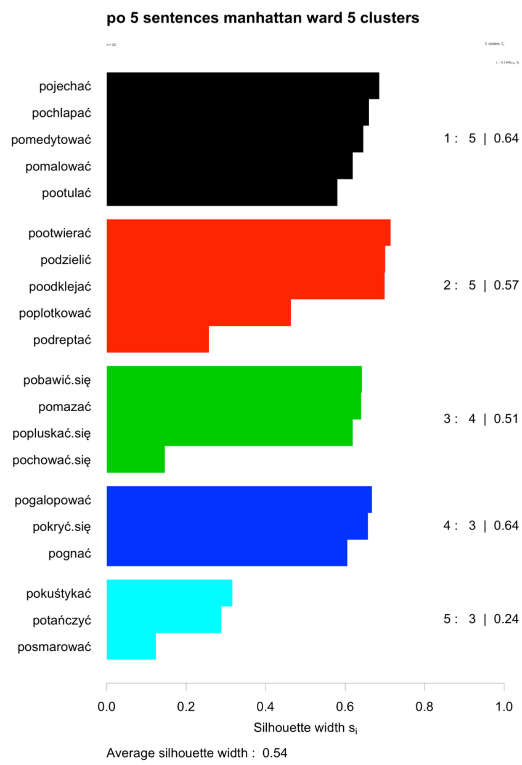


Figure 6.1b. Silhouette plot for a configuration with 5 clusters

with five smaller clusters. The average silhouette width of the solution in Figure 6.1a. is 0.34, whereas in Figure 6.1b., it is 0.54.

In conclusion, we should choose the five-cluster solution because the higher average silhouette width indicates that the clusters in this configuration are tighter and better separated. For each study, silhouettes for five different cluster configurations were plotted (the number of clusters considered in the silhouettes ranged from two to six). The average silhouette widths for every configuration are presented in Table 6.4.:

3 SENTENCES EACH						
PREFIX	2 clusters	3 clusters	4 clusters	5 clusters	6 clusters	**optimal solution**
po-	0.36	**0.4**	0.38	0.37	0.34	**3 clusters**
przy-	0.42	**0.61**	0.58	0.51	0.5	**3 clusters**
roz-	0.53	0.65	**0.76**	0.58	0.43	**4 clusters**

5 SENTENCES EACH						
PREFIX	2 clusters	3 clusters	4 clusters	5 clusters	6 clusters	**optimal solution**
po-	0.34	0.48	0.48	**0.54**	0.52	**5 clusters**
przy-	0.55	**0.63**	0.6	**0.63**	0.54	**3/5 clusters**
roz-	**0.4**	0.25	0.25	0.29	0.29	**2 clusters**

Table 6.4. Average silhouette widths

For each experiment, the cluster configuration with the highest average silhouette width was selected – this choice was then used for the annotation and interpretation of the dendrograms. As we can see, the optimal configurations range from two to five clusters. Each optimal cluster configuration was also checked for any items that might have been misaligned (i.e. having negative silhouette width) – no items were misaligned. For the sake of brevity and clarity, no silhouette plots will be presented in the following subsections reporting on particular experiments; the silhouette plots for the optimal solutions for each experiment can be found in Appendix 10.

The results of cluster analyses will be discussed in the alphabetical order of the prefixes: the results for *po-* will come first, then *przy-* and, finally, *roz-*. For each prefix, the analysis of the smaller study (3 sentences per sense) will be presented first and then compared with the analysis of the larger study (5/6 sentences per sense). All analyses were based on random samples of 160 responses because the lowest number of responses for the experiment was 164, and 160 is the nearest round number. A different random sample was drawn for each analysis.

6.3.2. Results for the prefix *po-*

The first experiment for *po-* consisted of twelve sentences containing verbs with the following senses of the prefix: COVER, SOMEWHAT, DISTRIBUTIVE, and DEPARTURE (three sentences per prefix sense). The sentences and their translations into English can be found in Appendix 3A. A random sample of 160 observations from the data provided by the participants was analysed with Ward's hierarchical agglomerative clustering method on a Manhattan dissimilarity matrix. The dendrogram for the final clustering solution is presented in Figure 6.2a.:

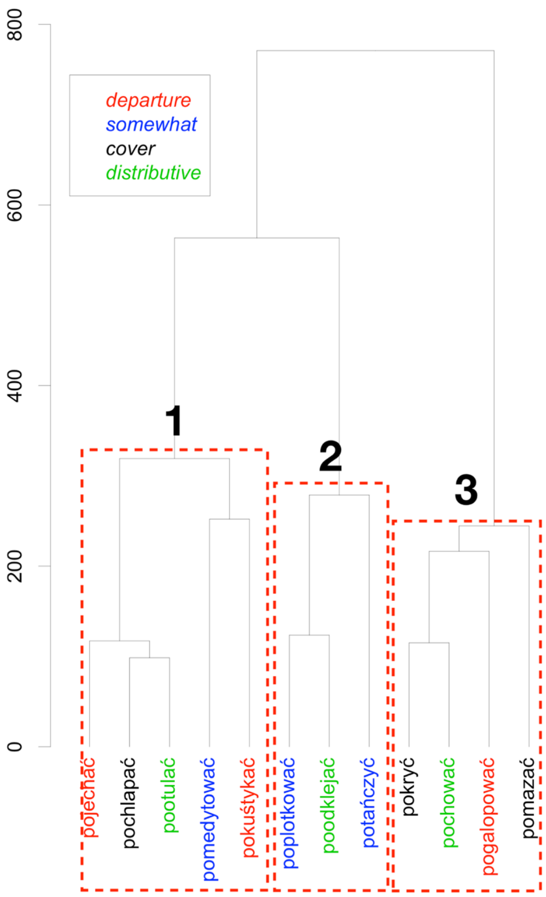


Figure 6.2a. Dendrogram for *po-* (3 sentences per sense; colour-coded for prefix sense)

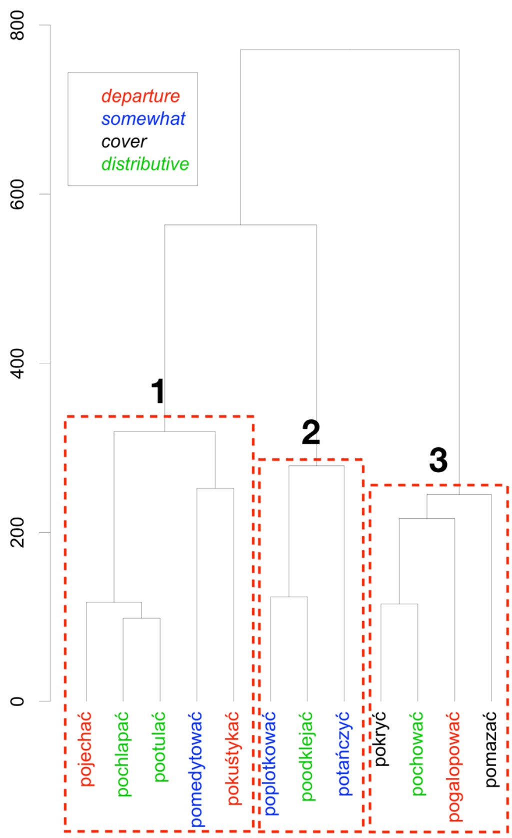


Figure 6.2b. Dendrogram for *po-* (3 sentences per sense; colour-coded for prefix sense, including alternative coding for *pochlapać*)

The verbs in the dendrogram were colour-coded according to the (purported) sense of the prefix they convey; a legend can be found in the upper-right corner of the plot. The best cluster configuration for this dataset (selected on the basis of the highest average silhouette width; see Section 6.3.1.) consisted of three clusters, which are marked with the red broken-line rectangles. This experiment contained one verb that permitted an alternative interpretation – the verb *pochlapać* could be interpreted either as carrying the COVER sense (the default sense given by Śmiech 1986) or the DISTRIBUTIVE sense. Figure 6.2b. shows the same clustering coded for prefix sense but with the verb *pochlapać* labelled as DISTRIBUTIVE.

The three clusters do not exhibit any clear pattern as far as the different senses of the prefix are concerned – sentences with verbs containing the same sense of the prefix were not grouped together in any of the clusters. Even if we take into consideration the alternative sense labelling for the verb *pochlapać* (see Figure 6.2b.), the situation remains the same: the clusters do not consistently group sentences containing verbs with the same sense of the prefix *po*-. The participants must have used other criteria to group the experimental sentences. Since they could not reliably spot similarities as far as prefix senses are concerned, they might have resorted to overtly coded cues such as sentence structure or verbal inflection to group the sentences[50]. Overtly coded or marked cues, for instance, verbal inflection or sentence structure, are prominent and immediately visible. As such, overtly marked cues provide a conspicuous criterion for grouping. As far as sentence structure is concerned, the element that a sentence begins with, for instance, is easy to perceive and might serve as a good criterion for sorting sentences. In the experiment on *po*- involving three sentences per prefix sense, the sentences began with an overt subject (example 6.2a), an implicit subject (i.e. a verb; example 6.2b), or with an adverbial (example 6.2c):

(6.2)	a.	*Po*	*pracy*	***po-****jecha-ł-a*	*do*	*kina*
		After	work	***po-***go-PST-3SG.FEM	to	cinema

[50] Verbal inflection or sentence order can be regarded as more overtly coded cues, because they are perceivable formal properties. Prefix senses, on the other hand, cannot be regarded as overtly coded cues, since they pertain to semantic properties of a verb or a sentence.

‘She went to the cinema after work’

b.	***Po-****plotkowa-ł-y*	*o*	*sąsiadkach.*
	po-gossip-PST-3PL.NON-VIR	about	female.neighbours
	‘They gossiped about their neighbours for a while’		

c.	*Koń*	***po-****galopowa-ł*	*do*	*lasu*
	horse	**po**-gallop-PST.3SG.MASC	to	forest
	‘The horse galloped to the forest’			

Let us now inspect another dendrogram for the same clustering, but this time with verbs colour-coded according to the sentence’s initial element to see whether participants used it as a criterion for grouping (Figure 6.2c):

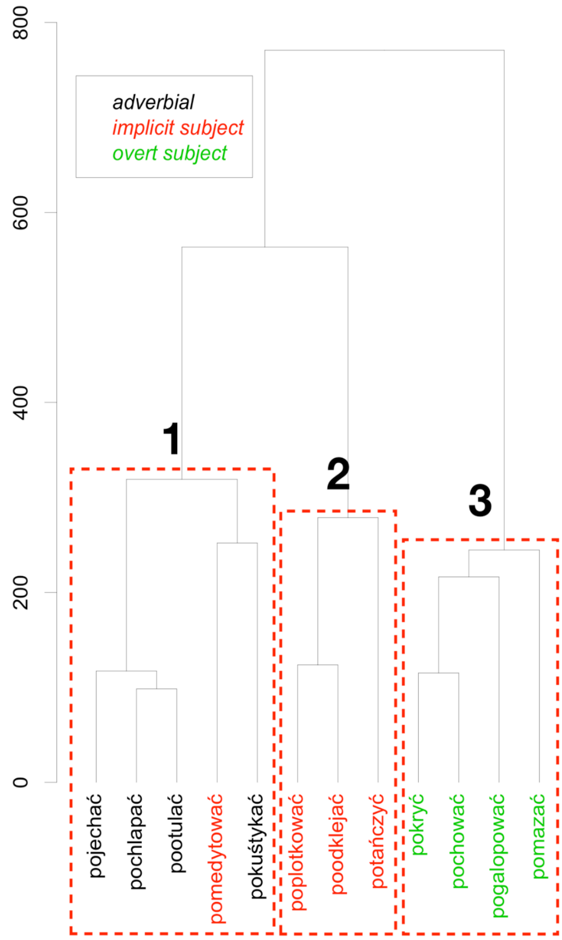


Figure 6.2c. Dendrogram for *po-* (3 sentences per sense; colour-coded for sentence structure)

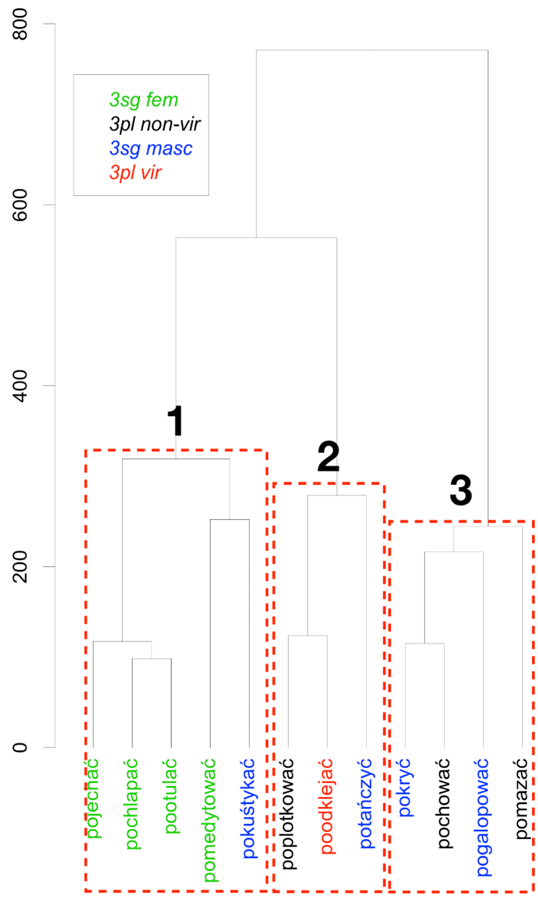


Figure 6.2d. Dendrogram for *po-* (3 sentences per sense; colour-coded for verbal inflection)

The first cluster is comprised of five sentences, out of which four begin with an adverbial, and one sentence begins with an implicit subject. The other two clusters (Cluster 2 and Cluster 3) are composed solely of sentences that begin with an implicit subject (i.e. with a verb) or overt subject respectively. Such a composition of the clusters suggests that participants might have used sentence structure (an overtly coded cue) over prefix senses as a criterion for sorting.

Verbal inflection is also quite a prominent overt and thus could have been used as a criterion for grouping. Let us now compare the clustering coded for sentence structure with another formal criterion: verbal inflection; the dendrogram marked for verbal inflection is presented in Figure 6.2d. Cluster 1 groups four verbs with the same inflection (3SG FEM) and one verb with a different inflection (3SG MASC), which makes verbal inflection an equally good criterion for grouping, as far as this cluster is concerned. Clusters 2 and 3 do not contain verbs carrying the same inflection; this stands in stark contrast to the grouping in terms of sentence structure, where Clusters 2 and 3 grouped sentences that began with identical elements (implicit subject and overt subject respectively). In conclusion, sentence structure appears to be a more likely criterion for grouping than verbal inflection as far as overtly coded cues are concerned.

Overall, we can see that the participants did not judge sentences with verbs containing the same (proposed) sense of the prefix *po-* as similar. It seems more likely that sentence structure constituted the criterion that the participants used for grouping. Consequently, the sorting experiment on *po-* with 3 sentences per prefix sense does not provide evidence supporting the hypothesis that prefix senses for *po-* exist as general usage-based constructions.

The second experiment for *po-* consisted of twenty sentences containing verbs with the following senses of the prefix: COVER, SOMEWHAT, DISTRIBUTIVE, and DEPARTURE (five sentences per prefix sense). The sentences and their translations into English can be found in Appendix 3B. Similarly to the first experiment, a random sample of 160 observations from the data provided by the participants was analysed with Ward's hierarchical agglomerative clustering method on a Manhattan dissimilarity matrix. The dendrogram for the final clustering solution is presented in Figure 6.3a:

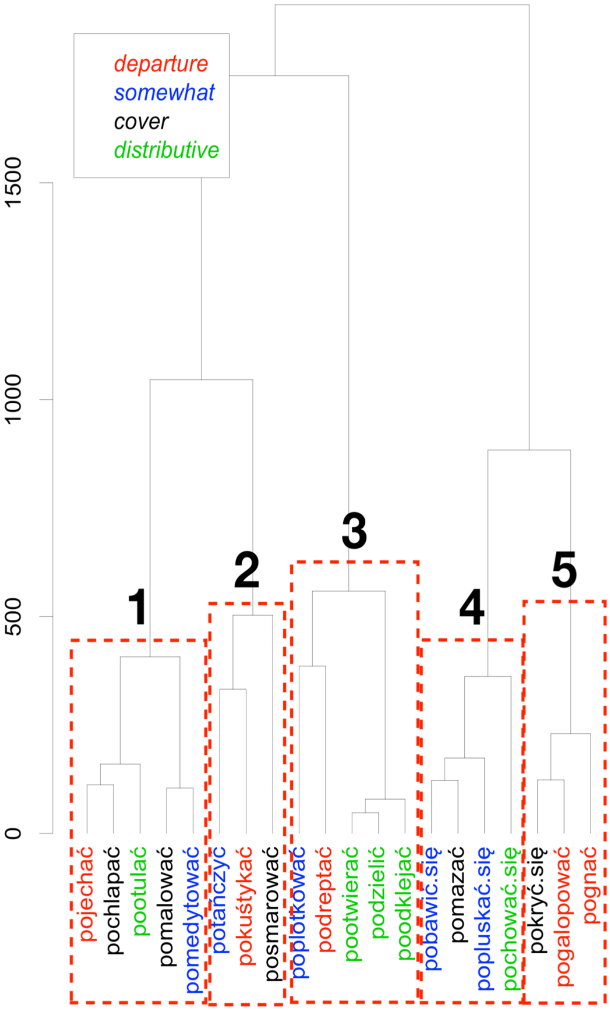


Figure 6.3a. Dendrogram for *po-* (5 sentences per sense; colour-coded for prefix sense)

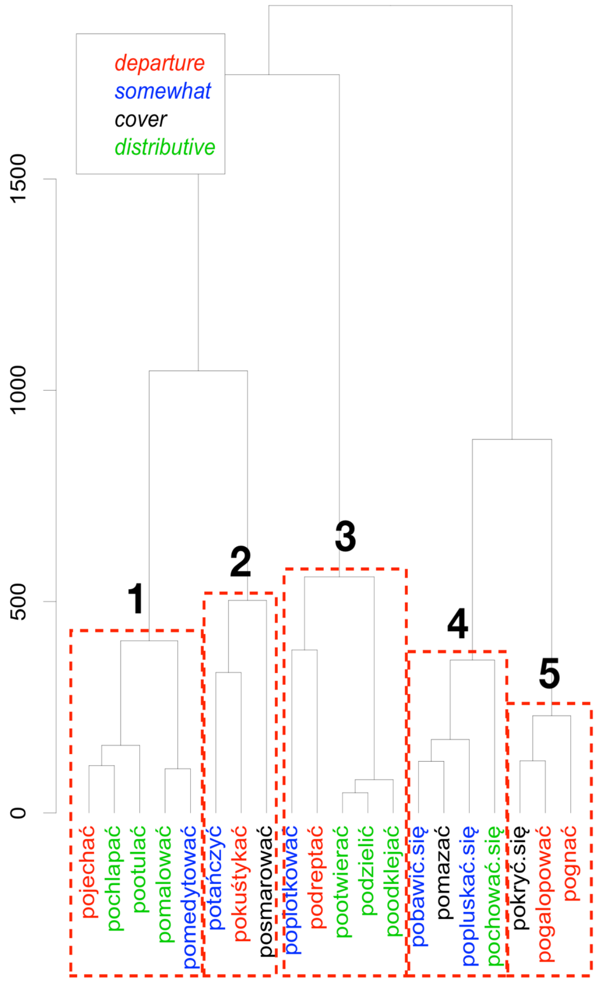


Figure 6.3b. Dendrogram for *po-* (5 sentences per sense; colour-coded for prefix sense, including alternative coding for *pochlapać* and *pomalować*)

Similarly to the previous dendrogram, the verbs were colour-coded according to the sense of the prefix they convey; a legend can be found in the upper-left corner of the plot. The best cluster configuration for this dataset consisted of five clusters, which are marked with the red broken-line rectangles. Figure 6.3b. shows the same clustering coded for prefix sense but with the verbs *pochlapać* and *pomalować* labelled as DISTRIBUTIVE (instead of COVER).

Overall, the clusters do not seem to cluster reliably according to the different senses of the prefix. Only in Cluster 3 (and in Cluster 1 of the grouping with alternative sense coding marked; Figure 6.3b.), we can see three verbs with the same sense of *po-* (DISTRIBUTIVE) grouped together – these verbs form a subcluster of Cluster 3 (Subcluster 3a). Here, similarly to the previous experiment, the participants seem to have used more overt grammatical cues to group the sentences. The two further dendrograms (Figure 6.3c. and 6.3d.) present the same clustering as in the first dendrogram (i.e. Figure 6.3a.). In these dendrograms (Figure 6.3c. and 6.3d.), however, the verbs are colour-coded according to the structure of the sentence they appear in (Figure 6.3c.) and the verbal inflection (Figure 6.3d.) they carry (the legend can be found in the upper-left corner of each dendrogram).

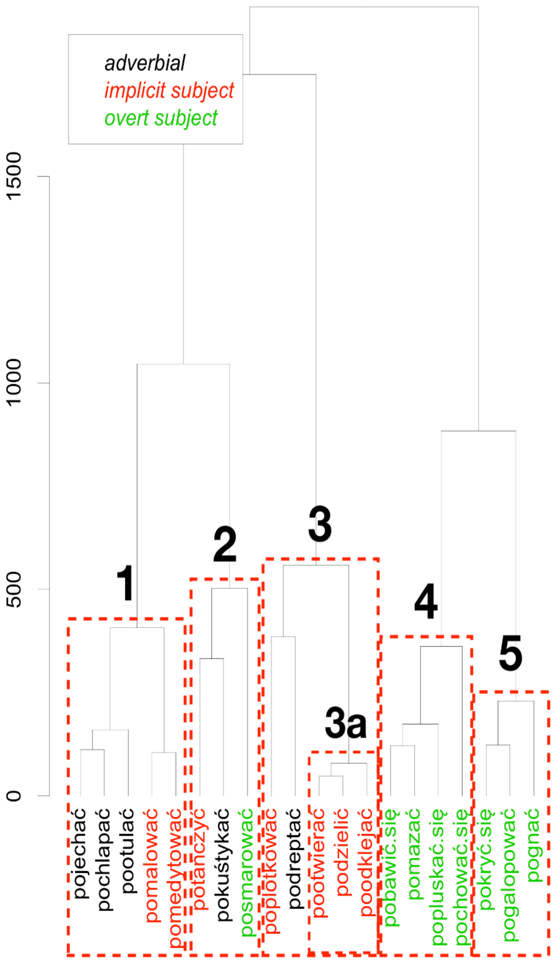


Figure 6.3c. Dendrogram for *po-* (5 sentences per sense; colour-coded for sentence structure)

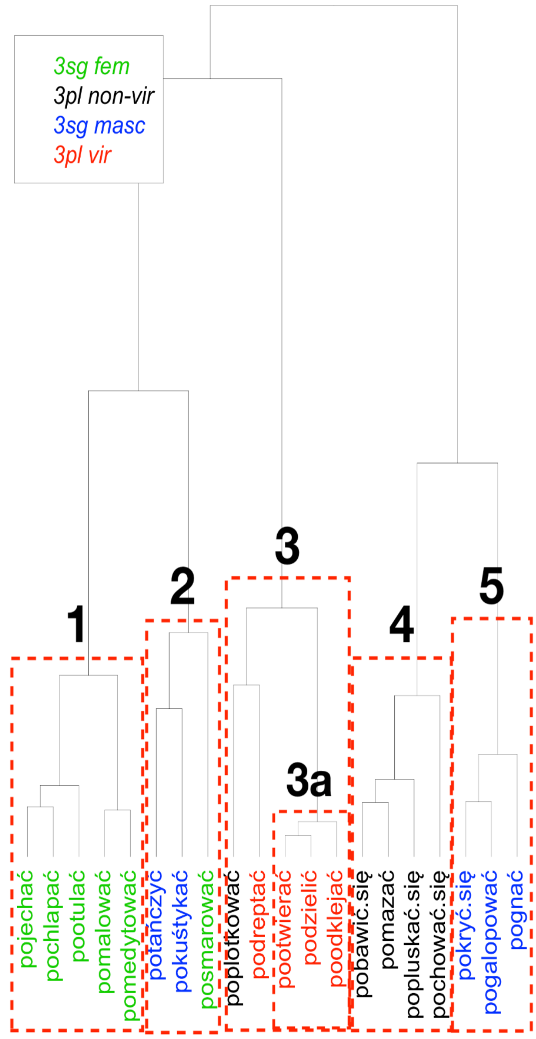


Figure 6.3d. Dendrogram for *po-* (5 sentences per sense; colour-coded for verbal inflection)

Let us inspect verbal inflection first (Fig 6.3d.). Clusters 1, 4, and 5 are made up entirely of sentences with verbs with the same inflection: 3SG FEM (third person singular feminine), 3PL NON-VIR (third person plural non-virile), and 3SG MASC (third person singular masculine). The two remaining clusters (2 and 3) are not entirely composed of verbs with one inflection. Cluster 2 contains two verbs in 3sg masc and one verb in 3sg fem, while Cluster 3 contains four verbs in 3PL VIR (third person plural virile) and one verb in 3PL NON-VIR. Nevertheless, most verbs in Cluster 2 (2 out of 3 verbs) and Cluster 3 (4 out of 5 verbs) carry the same inflection.

When we look at the sentence structure (Figure 6.3c), we can see that Clusters 4 and 5 both consist exclusively of the sentences with an overt subject. In Cluster 3, we have one sentence beginning with an adverbial and four sentences with an implicit subject. We can say that Clusters 3a, 4, and 5 are coherently composed of sentences with the same structure. Cluster 1 contains two sentences beginning with an adverbial and three sentences with an implicit subject, while Cluster 2 contains one sentence of each possible type – we can say that these clusters do not coherently group sentences of the same type. The sentence structure explanation seems to be somewhat worse than the explanation in terms of verbal inflection. Firstly, we have four clusters that coherently group verbs with the same inflection (1, 3a, 4, and 5) in comparison to three clusters in the sentence structure explanation. Additionally, there are four types of verbal inflection in comparison to three types of sentence structure. Since participants could group sentences into five bins, four types of inflection would enable them to consistently fill 4 out 5 bins with the same type of inflection.

Overall, verbal inflection appears to be the most likely criterion the participants used for sorting the sentences – it explains the clustering better than prefix senses or sentence structure. Just like in the case of the experiment on *po-* with three sentences per prefix senses, the results of the sorting experiment with five sentences per prefix do not support the hypothesis that prefix senses for *po-* exist as general usage-based constructions.

6.3.3. Results for the prefix *przy-*

The first experiment for *przy-* consisted of twelve sentences containing verbs with the following senses of the prefix: APPROACH, SOMEWHAT, COVER, and FIT/ATTACH (three sentences per prefix sense). The sentences and their translations into English can be found in Appendix 3C. A random sample of 160 observations from the data provided by the participants was analysed with Ward's hierarchical agglomerative clustering method on a Manhattan dissimilarity matrix. The dendrogram for the final clustering solution is presented in Figure 6.4a. below:

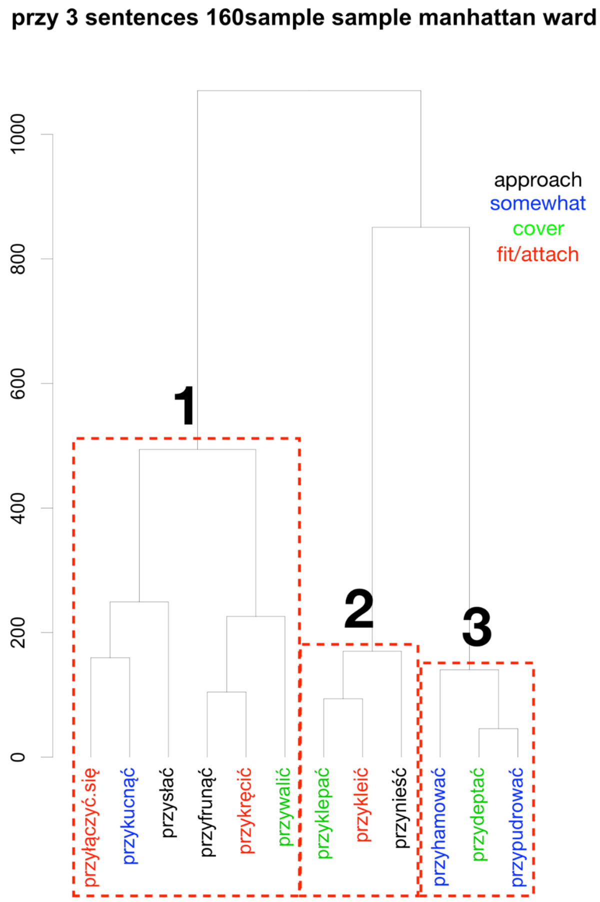


Figure 6.4a. Dendrogram for *przy-* (3 sentences per sense; colour-coded for prefix senses)

The best cluster configuration for this dataset consisted of three clusters, which are marked with the red broken-line rectangles. As previously, the verbs in the dendrogram were colour-coded according to the sense of the prefix they convey; a legend can be found in the upper-right corner of the plot.

Analogously to the experiments on *po-*, the clusters do not exhibit any clear structure reflecting the different senses of the prefix – sentences with verbs containing the same sense of the prefix did not group together. Since the participants seem not to have used the prefix sense as the criterion for grouping in the experiment on *przy-* with 3 sentences per prefix, let us now inspect two other possible grouping criteria: sentence structure and verbal inflection. The dendrograms below present the same clustering as the one rendered in Figure 6.4a, but they are colour-coded for sentence structure (Figure 6.4b.) and verbal inflection (Figure 6.4c.) respectively.

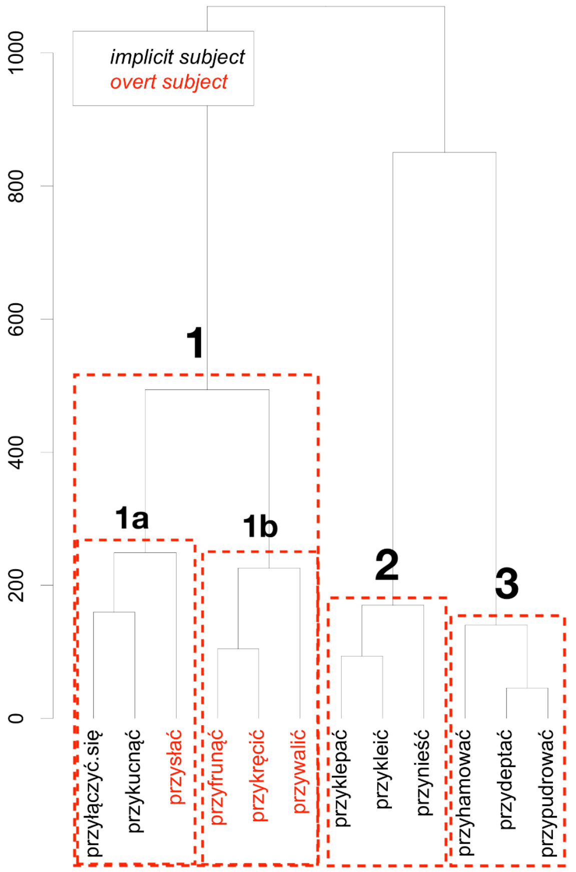


Figure 6.4b. Dendrogram for *przy-* (3 sentences per sense; colour-coded for sentence structure)

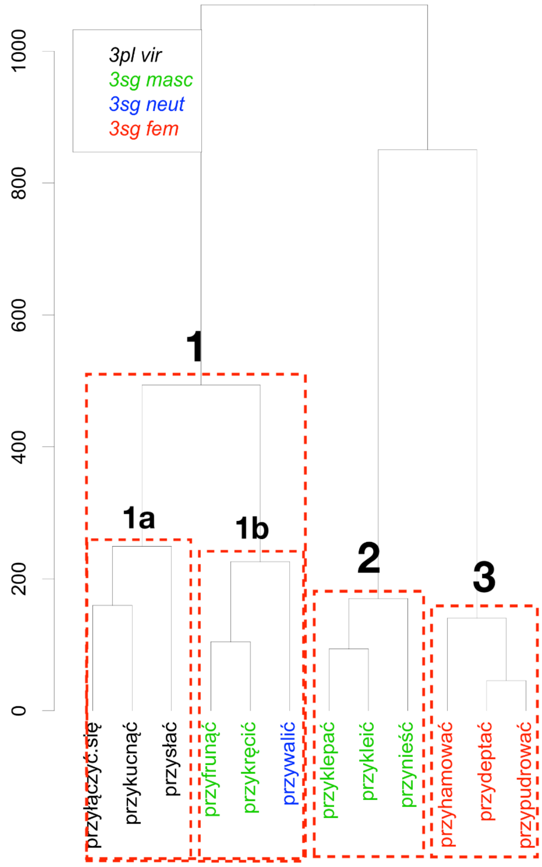


Figure 6.4c. Dendrogram for *przy-* (3 sentences per sense; colour-coded for verbal inflection)

In the dendrogram with sentence structure (Figure 6.4b.), the Clusters 2 and 3, and Subcluster 1b each consist of sentences having the same type of structure; Subcluster 1a has two sentences with an implicit subject and 1 sentence with an overt subject. As we can see, the clusters group sentences similar in terms of sentence structure. The usefulness of the sentence structure criterion for explaining the clustering, however, is diminished by the low variation in the types of sentence structure. The dendrogram shows that eight out of twelve sentences in the experiment began with an implicit subject, while the other four sentences began with an overt subject. Such an arrangement – only two types of sentence structure with one of them significantly more frequent than the other – might not provide enough variation in order to explain the groupings fully.

Let us now inspect the dendrogram colour-coded for verbal inflection (Figure 6.4c) to see if it can provide a better explanation of the clustering. In the case of verbal inflection, we have more variation in the data – the experimental sentences contained four types of marking. The Clusters 2 and 3, and Subcluster 1a each consist of sentences having the same type of structure (3PL VIR, 3SG MASC, and 3SG FEM respectively); Subcluster 1b contains two sentences marked for 3SG MASC, and one sentence marked for 3SG NEUT. Cluster 1 as a whole does not consist of sentences with verbs having identical inflection. If we take a closer look, however, we will see that five out of six sentences contain verbs marked for a "masculine-like" gender, that is either masculine or virile. The virile gender in Polish is similar to the masculine gender in that it usually refers to "all-male groups of people" (Swan 2002: 156)[51] – we could say that the virile gender is the masculine gender for the plural number. Overall, Clusters 1, 2, and 3 group sentences with coherent verbal inflection.

Verbal inflection appears to explain best the clustering obtained on the data for the experiment on *przy*-with three sentences per prefix sense. This indicates that participants might have used more overtly coded cues – that is, verbal inflection instead of prefix senses – as the criterion for grouping sentences. Consequently, the result of the experiment on *przy-* with three sentences per sense does not provide evidence that prefix senses exist as salient general usage-based constructions in the minds of speakers.

[51] It can sometimes alternatively refer to "groups of people with mixed male and female constituency" or "groups of people and other things" (Swan 2002: 156).

As in the case of *po-*, the second experiment for *przy-* consisted of twenty sentences containing verbs with the same prefix senses as the first experiment for the prefix (five sentences per prefix sense). The sentences and their translations into English can be found in Appendix 3D. Similarly to the first experiment, a random sample of 160 observations from the data provided by the participants was analysed with Ward's hierarchical agglomerative clustering method on a Manhattan dissimilarity matrix. The dendrogram for the final clustering solution is presented in Figure 6.5a. below:

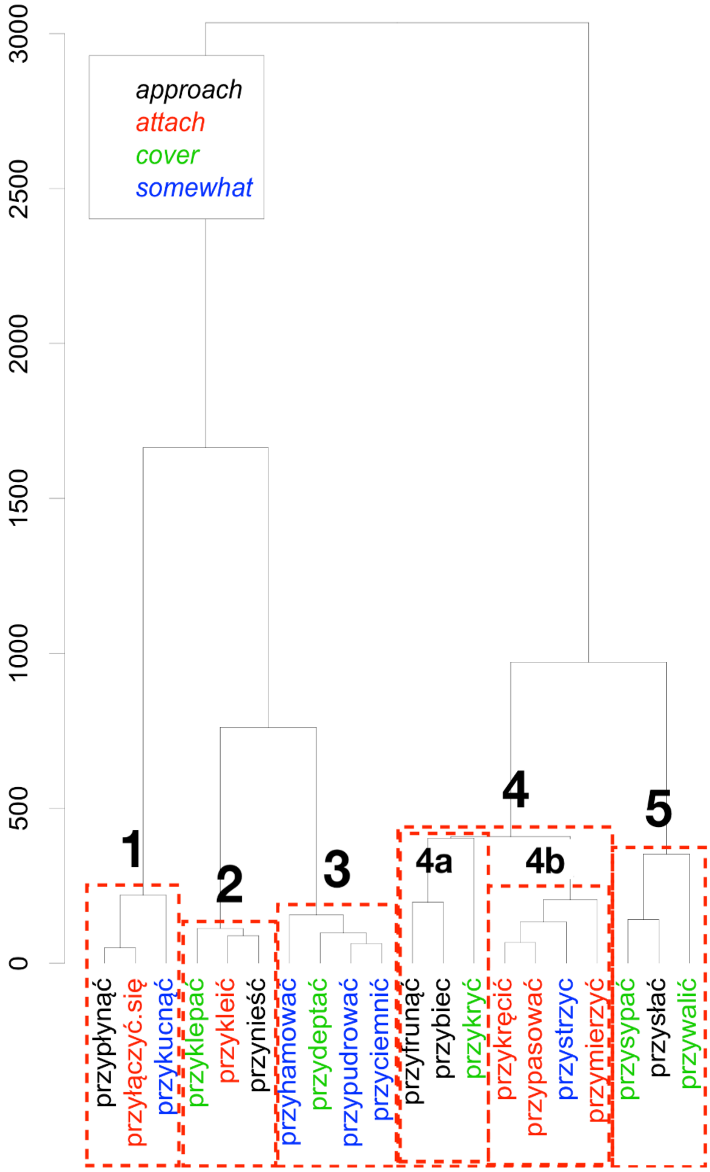


Figure 6.5a. Dendrogram for *przy-* (5 sentences per sense; colour-coded for prefix sense)

The best cluster configuration for this dataset consisted of five clusters, which are marked with the red broken-line rectangles. Similarly to other dendrograms, the verbs were colour-coded according to the sense of the prefix they convey; a legend can be found in the upper-right corner of the plot.

In the dendrogram above, we can see that most of the five clusters do not coherently group verbs with the same sense of the prefix *przy-*. In Cluster 3, we can find three verbs with the SOMEWHAT sense of the prefix and one with the COVER sense. Two subclusters have also been marked on the dendrogram (4a & 4b). In (4a), we can find three verbs with the ATTACH sense of the prefix and one with the SOMEWHAT sense. Subcluster 4b consists of two APPROACH verbs and one COVER verb. It seems that there might be some regularity (in terms of prefix senses) that the participants grouped the sentences in these subclusters. Cluster 4 as a whole, however, contains verbs with all four senses – thus, its structure is not coherently based on prefix senses. If we have a look at Cluster 5, we can see that it consists of two COVER verbs and one APPROACH verb, which might also suggest some regularity in terms of prefix senses (similarly to Subcluster 4b). This regularity, however, would only be a weak one, because the cluster contains only two out of five COVER verbs.

Let us now inspect dendrograms colour-coded for verbal inflection (Figure 6.5b.) and sentence structure (Figure 6.5c.) to see whether the participants might have used more overtly coded cues as the criterion for grouping. The cluster numbers in the dendrograms below have been retained from the first dendrogram (Figure 6.5a.) so that we could compare the different criteria for clustering reliably.

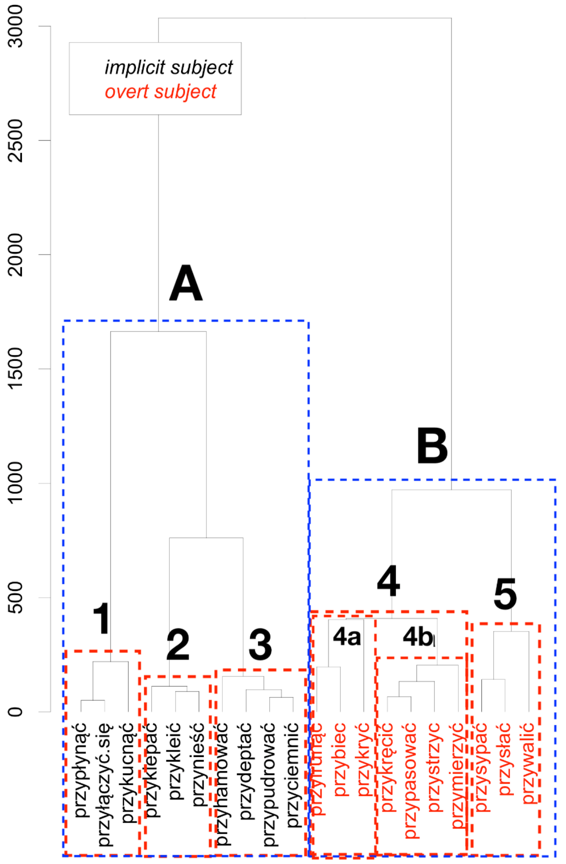


Figure 6.5b. Dendrogram for *przy-* (5 sentences per sense; colour-coded for sentence structure)

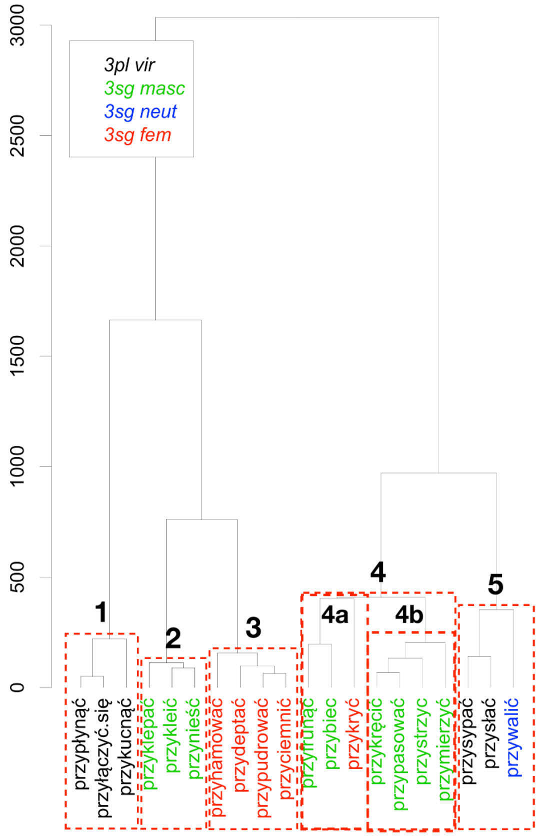


Figure 6.5c. Dendrogram for *przy-* (5 sentences per sense; colour-coded for verbal inflection)

The dendrogram on the left-hand side of Figure 6.5b. was colour-coded according to sentence structure (i.e. what the sentence began with). We can see two large clusters here (marked with blue rectangles): A, which consists only of sentences with an implicit subject; and B, which consists only of sentences with an overt subject. The clustering seems to be very coherent, but does not correspond to the optimal 5-cluster structure determined on the basis of silhouette width. The other dendrogram in Figure 6.5c. is colour-coded according to the verbal inflection a given verb carried. Clusters 1, 2, 3, and Subcluster 4a are comprised of verbs with the same inflection (3PL VIR, 3SG MASC, 3SG FEM, and 3SG MASC, respectively). Subcluster 4b and Cluster 5 consist of two verbs with one marking (3SG MASC and 3PL VIR, respectively) and one verb with another marking (3SG FEM and 3SG NEUT, respectively). In other words, we can say that most clusters coherently group verbs with the same verbal inflection. The verbal-inflection-based grouping seems superior to the grouping based on sentence structure, because there is more variation to the types of inflection the verbs carry (4 types), and the groupings correspond better to the optimal 5-cluster configuration.

If we compare the grouping based on prefix senses and the one based on verbal inflection, we can see that the latter appears to explain the clustering presented on the dendrogram much better. Additionally, Cluster 3 and Subcluster 4a – which contain 75% of verbs with the same prefix sense – are both fully made up of sentences with verbs carrying the same inflection. This suggests that the grouping according to the prefix sense in Cluster 3 and Subcluster 4a might actually have been an epiphenomenon of the verbal inflection. In sum, the result of the experiment on *przy-* with five sentences per sense does not support the hypothesis that Polish speakers have salient general constructions for the different senses of the prefix *przy-*.

6.3.4. Results for the prefix *roz-*

The analysis of the data from the experiments on *roz-* followed exactly the same procedure as in the case of the other two prefixes. A random sample of 160 observations was analysed with Ward's hierarchical agglomerative clustering method on a Manhattan dissimilarity matrix. As previously, the dendrograms were colour-coded based on the sense of the prefix.

The first experiment for *roz-* consisted of nine sentences containing verbs with the following senses of the prefix: DISPERSION, OPPOSITE, and INTENSITY (three sentences per prefix sense). The sentences and their translations into English can be found in Appendix 3E. The dendrogram for the final clustering solution is presented in Figure 6.6a. below:

roz 3 sentences 160sample sample manhattan ward

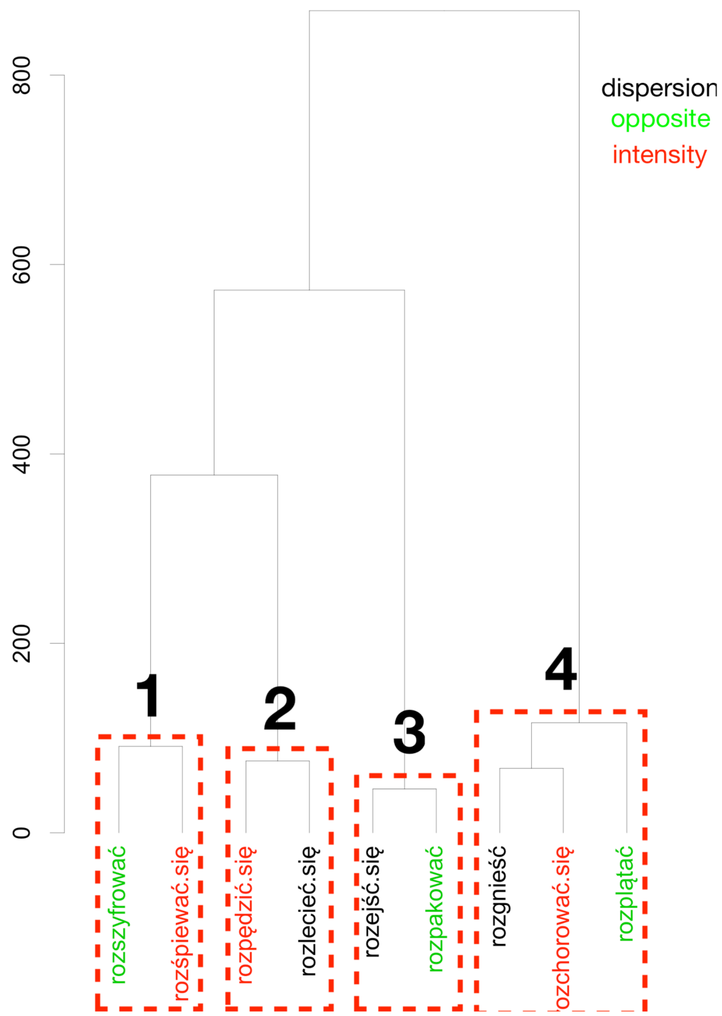


Figure 6.6a. Dendrogram for *roz-* (3 sentences per sense; colour-coded for prefix sense)

The best cluster configuration for this dataset consisted of four clusters, which are marked with the red broken-line rectangles. The verbs in the dendrogram are colour-coded according to the prefix sense that the particular verb conveys. The clusters do not seem to exhibit any clear structure reflecting the different senses of the prefix – in each cluster, there is a mix of different prefix senses without any sense that would be dominant. The participants must have used a different criterion for grouping the experimental sentences.

Analogously to previous analyses, let us now analyse two dendrograms for the same clustering in order to see whether the participants used more overtly coded cues as the criterion for grouping the sentences. The verbs in the two dendrograms below are colour-coded according to verbal inflection (Figure 6.6b.) and sentence structure (Figure 6.6c.). As far as the sentence structure is concerned, most sentences (7 out of 9) began with and adverbial. Such a distribution of sentence structure does not provide enough variation in order to explain the four clusters selected as optimal in this analysis.

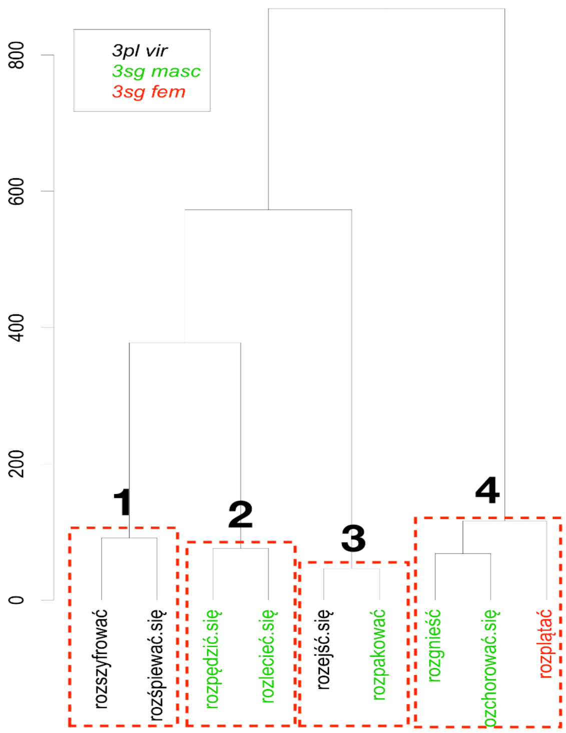


Figure 6.6b. Dendrogram for *roz-* (3 sentences per sense; colour-coded for verbal inflection)

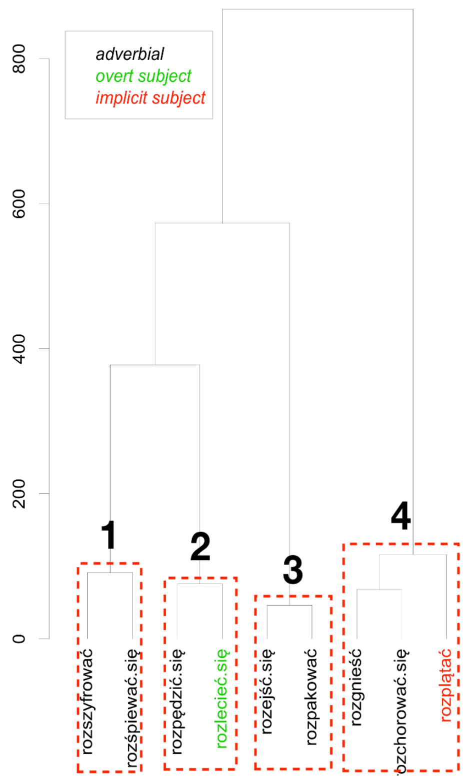


Figure 6.6c. Dendrogram for *roz-* (3 sentences per sense; colour-coded for sentence structure)

The dendrogram for the verbal inflection, on the other hand, appears to provide the best explanation of clusters from all possible criteria for sorting that we have taken into consideration (i.e. prefix senses, sentence structure, and verbal inflection). Clusters 1 and 2 are made up of verbs with the same inflection (3pl vir and 3sg masc respectively), and Cluster 4 contains two 3SG MASC verbs and one 3SG FEM; Cluster 3 is mixed – it contains on verb in 3SG MASC and one verb in 3PL VIR; what those two inflections have in common is that they both correspond to "masculine" subjects. We can see that the structure of most clusters can be effectively explained with verbal inflection – the verbs in clusters either carry the same inflection, or the inflectional forms are related (in Cluster 3, we had two "masculine-like" forms).

Consequently, the participants appear to have used overtly coded cues – verbal inflection in this case – for grouping the sentences instead of prefix senses. The results of the sorting experiment on the prefix *roz-* with three sentences per sense – similarly to the results of the four experiments discussed previously – do not indicate that native speakers of Polish might have general constructions for the different senses of the prefix *roz-*.

The second experiment for *roz-* consisted of seventeen sentences (five or six sentences per sense[52]) with the same prefix senses as the first experiment for the prefix. The sentences and their translations into English can be found in Appendix 3F. Similarly to the first experiment, a random sample of 160 observations from the data provided by the participants was analysed with Ward's hierarchical agglomerative clustering method on a Manhattan dissimilarity matrix. The dendrogram for the final clustering solution is presented in Figure 6.7a. below:

[52] See Section 6.2.2. for the reason for such composition.

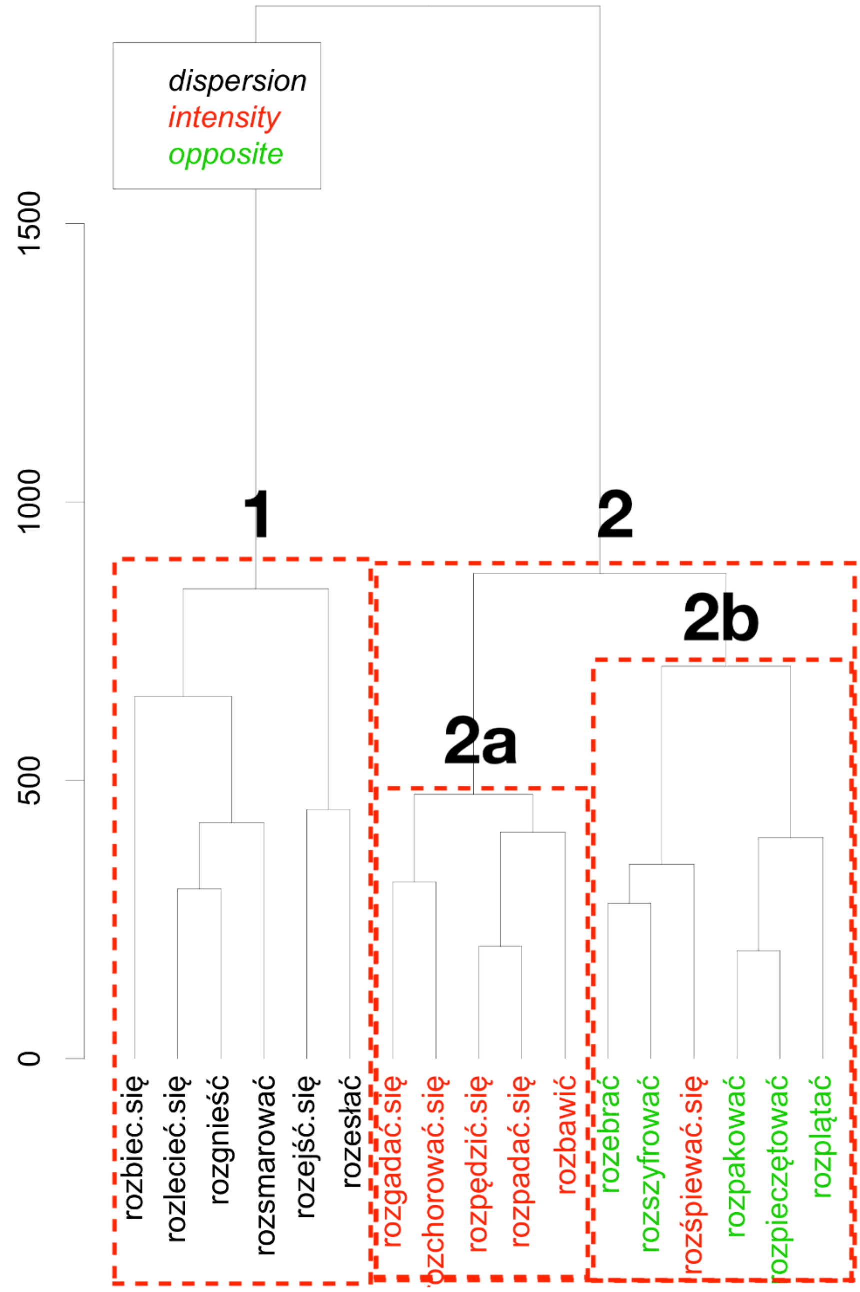


Figure 6.7a. Dendrogram for *roz-* (5 sentences per sense; colour-coded for prefix sense)

The best cluster configuration for this dataset consisted of two clusters, which are marked with the red broken-line rectangles. This clustering differs greatly from the clustering solutions produced for the other five experiments. We can see a very clear structure based on the senses of the prefix: Cluster 1 only contains verbs in which *roz-* conveys DISPERSION, whereas Cluster 2 contains *roz-* in the sense of INTENSITY and OPPOSITE ACTIONS. Furthermore, if we take a closer look at Cluster 2, we will notice that it splits into two more large subclusters – one exclusively consists of verbs with the INTENSITY sense of *roz-* (Subcluster 2a), the other one mainly contains verbs with *roz-* that conveys OPPOSITE *actions* (5 out of 6 verbs in Subcluster 2b). The cluster arrangement seems to correspond to the different sense of the prefix *roz-*, but we need to inspect the other possible criteria for grouping (sentence structure and verbal inflection) to see whether prefix senses did not correlate with particular verbal inflections or sentence structures. Figure 6.7b. and Figure 6.7c. below present two more dendrograms for the clustering presented in Figure 6.7a. above; the dendrograms are colour-coded for verbal inflection (Figure 6.7b.) and sentence structure (Figure 6.7c.).

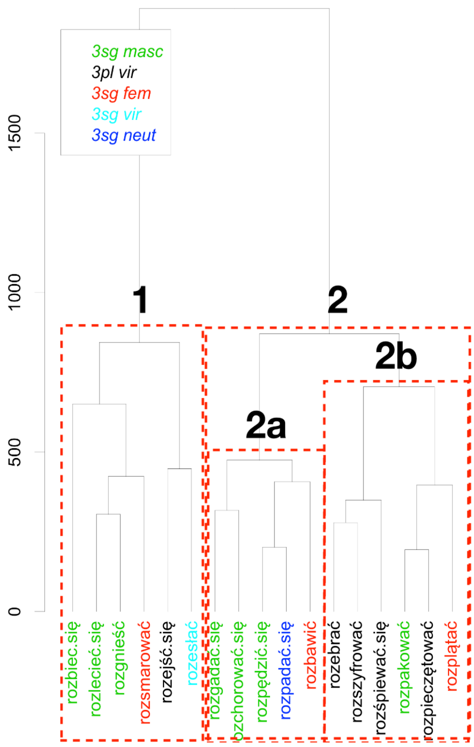


Figure 6.7b. Dendrogram for *roz-* (5 sentences per sense; colour-coded for verbal inflection)

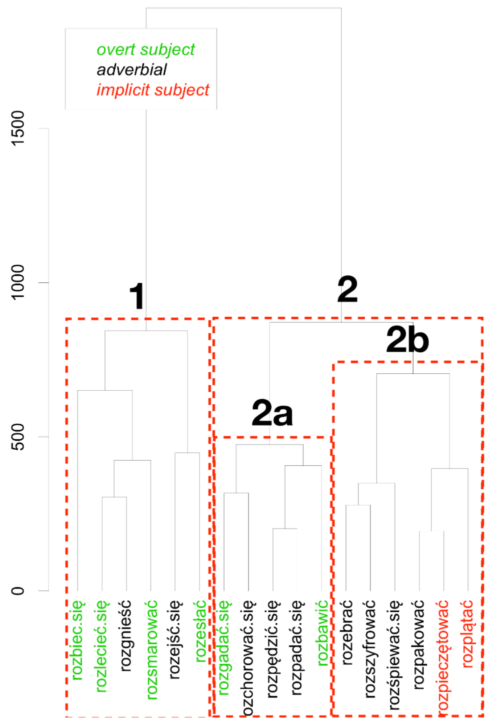


Figure 6.7c. Dendrogram for *roz-* (5 sentences per sense; colour-coded for sentence structure)

In the two dendrograms above, we can see that verbal inflection and sentence structure do not explain the clustering as well as prefix senses. The Clusters 1, 2, 2a, and 2b are not so consistently made up of the same kind of verbal inflection or sentence structure. We can conclude that prefix sense is the most likely criterion the participants used for sorting the sentences.

Even though the data from this experiment exhibit a very clear structure reflecting the different senses of the prefix *roz-*, we need to treat this result with caution, because the average silhouette width for the result – which also happens to be the best clustering solution – equals only 0.4. As a rule of thumb, we can say that silhouette width values between 0.26 and 0.5 suggest that "[t]he structure is weak and could be artificial" (Spector 2011: 172). What it means is that the algorithm "struggled" to find clear clusters, but it nonetheless performed some clustering, because hierarchical agglomerative clustering algorithms do not stop unless they have clustered all data. Let us examine the silhouette plot to determine whether both clusters were of the same (poor) quality, or maybe one cluster was substantially worse than the other and thus affected the average silhouette width:

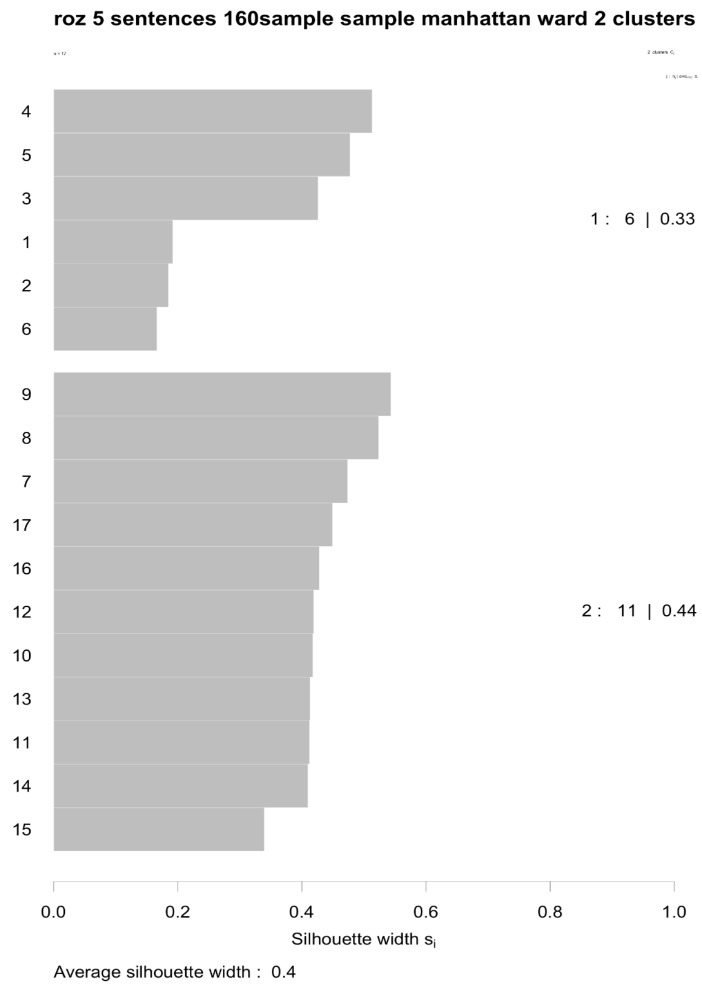


Figure 6.8. Silhouette plot for a two-cluster solution (*roz-* 5/6 sentences per meaning)

Figure 6.8. indicates that the first cluster's quality (average silhouette width = 0.33) is somewhat lower than the second cluster's (average silhouette width = 0.44). We can see that the silhouette width for three observations in cluster 1 is approximately 0.2, which indicates a random structure. Overall, however, none of the cluster's average silhouette width exceeds 0.5, so we cannot treat the analysis as robust.

6.4. Discussion

Five out of six plots for the experiments presented in the previous subsections did not show any clustering in accordance with prefix senses – the experiment on *roz-* with 5/6 sentences was the only exception. By and large, the participants seem to have chosen more prominent and formally marked cues: gender and number. Let us have a look at a breakdown of all clusters in all experiments in terms of the gender and number of verbs included in target sentences:

3 SENTENCES EACH					
PREFIX	**CLUSTER 1**	**CLUSTER 2**	**CLUSTER 3**	**CLUSTER 4**	**CLUSTER 5**
po-	**4x 3sg fem 1x 3sg masc**	2x 3pl non-vir 1x 3sg masc	2x 3sg masc 2x 3pl non-vir		
przy-	**A: 3x 3pl vir** **B: 2x 3sg masc + 3sg neut**	**3x 3sg masc**	**3x 3sg fem**		
roz-	**2x 3pl vir**	**2x 3sg masc**	**2x 3pl vir**	**2x 3sg masc**	

5 SENTENCES EACH					
PREFIX	**CLUSTER 1**	**CLUSTER 2**	**CLUSTER 3**	**CLUSTER 4**	**CLUSTER 5**
po-	**3x 3sg fem**	3sg masc + 3sg fem	**4x 3pl vir + 1x 3pl non-vir**	**4x 3pl non-vir**	**3x 3sg masc**
przy-	**3x 3pl vir**	**3x 3sg masc**	**4x 3sg fem**	**6x 3sg masc 1x 3sg fem**	2x 3pl vir 1x 3sg neut
roz-	3x 3sg masc 1x 3sg fem 2x 3pl vir	4x 3sg masc 1x 3sg neut 2x 3sg fem 3x 3pl vir			

Table 6.5. Gender and number of verbs in target sentences

The data presented in Table 6.5. indicate that 12 out of 22 clusters (or 20, if we do not take into account the last *roz-* experiment) were composed **exclusively** of verbs of the same gender and number (they have been highlighted with light grey shading and a bold font); another 3 clusters consisted predominantly of verbs of the same gender and number (presented in bold, without shading). With Cluster 1 for the three-sentence *przy-* experiment, if we split it into two subclusters (A and B; see Table 6.5.), we will also obtain clusters with verbs of exclusively or predominantly the same gender and number. These results suggest that for the lack of a stronger semantic overarching criterion, the participants needed to resort to other – formally prominent and thus easily perceivable – grouping criteria.

The above data and the clustering results that did not yield any structure based on prefix senses suggest that the participants did not have general prefix constructions, at least for *po-* and *przy-*. Consequently, we can say that the experiment did not yield any evidence that salient general prefix constructions would exist in native speakers' minds. The vast majority of participants had either graduated from a university or still were in university education (96.41% had had at least some tertiary education). Dąbrowska (2008a) found that speakers of different educational backgrounds have categories of different level of generality – only highly educated speakers were able to make use of a very general construction of Polish dative neuter inflection. If we interpret the results of the above experiments in the light of Dąbrowska's (2008a) study, there is no reason to believe that speakers of different background could have grouped the sentences according to general prefix constructions, since the participants already belonged to the demographics that attain the most general categories.

6.5. Choices modulated by frequency

One might wonder why *roz-* was the only "outlier" amongst the prefixes (and only in one experiment) and why the respondents did not sort sentences with *po-* and *przy-* in accordance with the different senses of the prefixes. A possible explanation will be discussed in this section, based on a variable that the data of the corpus study presented in Chapter 5 was tagged for – the sense of each prefix. The senses for each prefix corresponded to the senses discussed in Section 5.1.1. (Table 5.1.):

COVER, DEPARTURE, DISTRIBUTIVE, and SOMEWHAT for *po*-; APPROACH, COVER, FIT, and SOMEWHAT for *przy*-; DISPERSION, INTENSITY, and OPPOSITION for *roz*-. Two additional "senses" (for each prefix), or rather groups, were included: IDIOSYNCRATIC and ASPECTUAL. The former group, IDIOSYNCRATIC, was an umbrella term for all prefixed verbs that did not display any regularity as far as the function/meaning of the prefix was concerned. The ASPECTUAL group contained prefixed verbs in which the only function/meaning of the prefix was transforming the non-prefixed counterpart of a given verb from the imperfective to the perfective aspect. A histogram presenting the frequency of occurrence of each sense of *po-, przy*-, and *roz*- in plTenTen and NKJP is presented in Figure 6.9. below:

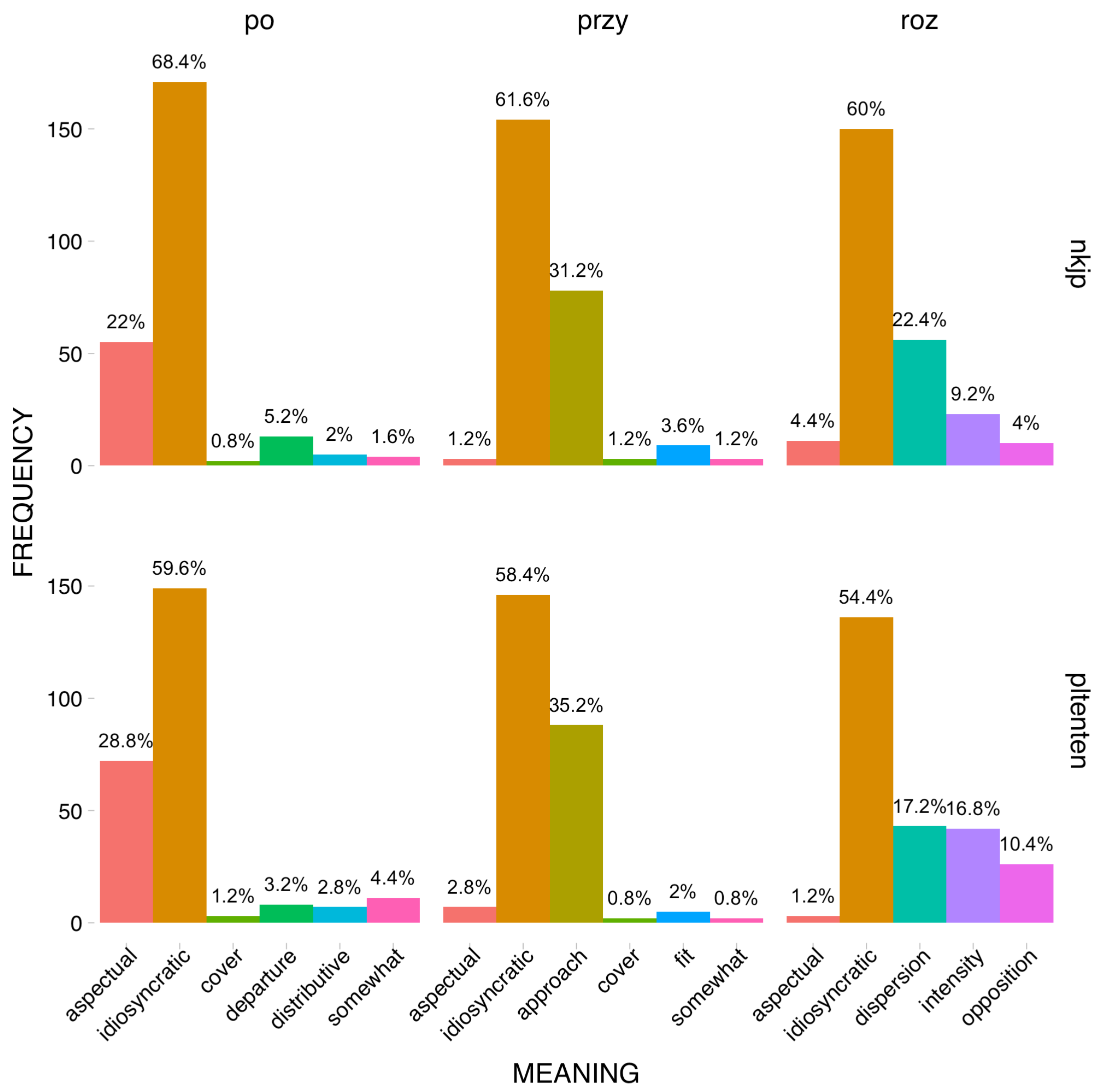


Figure 6.9. Distribution of prefix senses per corpus (for each corpus separately)

Visually, the distributions of prefix senses do not differ substantially between the data from NKJP and the data from plTenTen. The only prefix, for which the distribution of senses differs between the corpora is *roz* – the data from plTenTen seems more evenly distributed among the more concrete senses: DISPERSION, INTENSITY, and OPPOSITION. Since the distributions do not differ substantially between the two corpora (with the exception of *roz-*), in further analysis, we will consider the frequencies tallied over both corpora. Let us now look at a plot displaying the distribution of different senses in the two corpora summed up (Figure 6.10.):

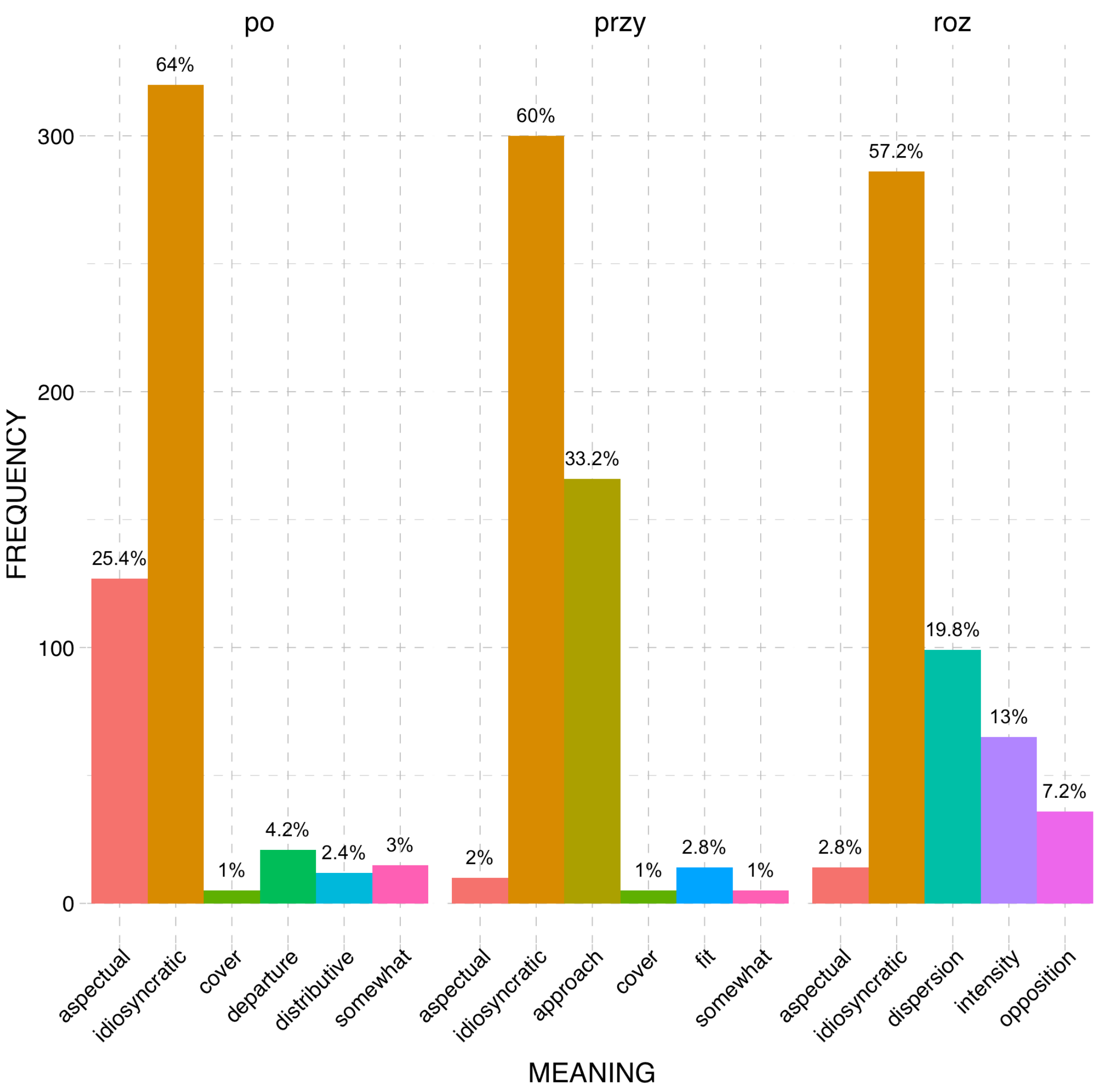


Figure 6.10. Distribution of prefix senses per corpus (summed over corpora)

The first striking property of the data for all prefixes is that the category IDIOSYNCRATIC dominates the dataset in both corpora, constituting from 57.2% to 64% of senses. Such a high proportion of IDIONSYNCRATIC verbs suggests that the majority of uses of prefixes cannot be predicted from a more general category.

In the prefix *po-*, ASPECTUAL is the second most frequent category with 25.4% of all *po*-verbs in the sample. In verbs from the ASPECTUAL group, the prefix does not contribute any meaning beyond transforming the verb from an imperfective verb to a perfective one. If we sum up the percentages, we will see that in 89.4% of examples, *po-* fell either into the category ASPECTUAL or IDIOSYNCRATIC, which means that only around 10% of *po*-verbs carried the other, more "concrete", senses postulated for the prefix. In the case of *przy-*, the categories IDIOSYNCRATIC and APPROACH make up over 93.2% of the distribution, which means that one sense (APPROACH) and one umbrella category (IDIOSYNCRATIC) almost entirely dominate the data, leaving only around 7% for the other senses. When we also factor in the ASPECTUAL category (another 2%), the remaining senses (COVER, FIT, and SOMEWHAT) constitute only 4.8% of the data.

With *roz-*, the senses are somewhat more evenly distributed than in the case of *po-* and *przy-*, and categories other than IDIOSYNCRATIC or ASPECTUAL account for the data. That notwithstanding, IDIOSYNCRATIC still remains the single most frequent category with 57.2% of examples. In Figure 6.9, we can also see that – in contrast to *po-* and *przy-* – the distributions for *roz-* differ between the two corpora. The DISPERSION sense is substantially more frequent than the other two in NKJP, whereas the data from plTenTen does not exhibit this 'bias'.

Already in the simple frequency counts of different prefix senses, we could see that the senses of each prefix are distributed highly unequally. In all prefixes, the IDIOSYNCRATIC category dominated the distribution with a 57- to 64-percent share of the data, which means that approximately half to two-third of occurrences of prefixed verbs in the analysed dataset were lexicalised. In *po-*, this effect was even stronger, because another 25% of examples included verbs that carried the ASPECTUAL sense of the prefix. The four remaining, more concrete, senses of *po-* (DISTRIBUTIVE**,** DEPARTURE**,** SOMEWHAT, and COVER) made up around 10%

of all examples. In the case of *przy-*, apart from the IDIOSYNCRATIC category, the dataset almost exclusively contained prefixed verbs from only one sense: APPROACH (approx. 33% of all *przy-* verbs). Altogether, the two categories – APPROACH and IDIOSYNCRATIC – covered almost 95% of the data for *przy-*. Among the three prefixes, only *roz-* exhibits a relative balance when it comes to the distribution of senses, and the more concrete senses (i.e. other than ASPECTUAL or IDIOSYNCRATIC) share from 7.2% to 19.8% of all examples.

If we extrapolate the results beyond the analysed sample, we could surmise that a native speaker of Polish would not come across usage data of sufficient diversity to be able to develop robust general constructions of most of the senses of *po-* and *przy-* (except for the APPROACH sense of *przy-*). The relative type frequencies of each detailed sense of *po-* and three out of four senses of *przy-* were extremely low. Bybee (2010) suggested that for a construction to be productive, its type frequencies must be high in order to provide enough diverse contexts so that native speakers could extrapolate the construction to new items (see Section 1.7). Most likely, the number of different lexemes that contain the different senses of *po-* or *przy-* (except the APPROACH sense of *przy-*) would not suffice to make the constructions postulated for the prefixes productive. The overwhelming dominance of the IDIOSYNCRATIC and ASPECTUAL categories would make it even less probable for speakers to encounter verbs with one of the four more concrete senses, preventing them from developing robust constructions for those senses. The results of the sentence-sorting experiment reflect the sense distributions of the prefixes – the participants were unable to sort the sentences for *po-* and *przy-* according to the prefix sense.

In *przy-*, we could see that one sense – APPROACH – has a very high type frequency, with 33.2% (166 out of 500) of all occurrences of the prefix *przy-* in the sample used for this study (including the examples from both NKJP and plTenTen). Such high type frequency of the APPROACH sense of *przy-* makes it a much more likely candidate for a construction that the speakers could develop and use. This sense of *przy-* occurred with 27 different verbs in total, but as few as four verbs (*przyjść* 'come$_{\text{PFV}}$', *przynieść* 'bring$_{\text{PFV}}$', *przychodzić* 'come$_{\text{IMPF}}$', *przyjechać* 'drive to$_{\text{PFV}}$') accounted for around 66% of occurrences of the entire sense. Moreover, if we look at a bar plot of frequencies of different verbs conveying the sense APPROACH (Figure 6.11. below), we can see a very clear Zipfian[53] distribution – there are only a few very frequent verbs, and the rest of the distribution is dominated by verbs that did not occur very often (frequency ≤ 5).

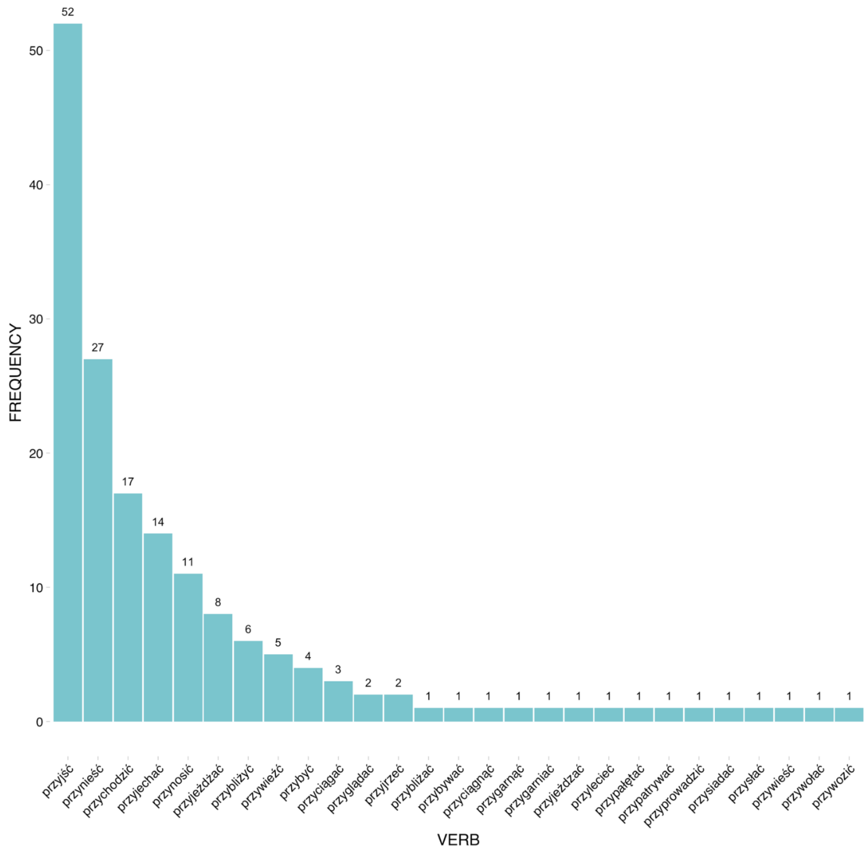


Figure 6.11. The distribution of przy- verbs in the sample conveying the sense approach

Such a distribution has been shown to facilitate the acquisition of constructions (Ellis & Ferreira-Junior 2009; see Section 1.7. for discussion). Lastly, we could argue that the four most frequent verbs in the APPROACH category (*przyjść*, *przynieść*, *przychodzić*, and *przyjechać*) are also very general and schematic, and thus prototypical. The prototypicality and generality of the most frequent lexical items associated with a construction has also been shown to facilitate the acquisition of a construction. To sum up, the APPROACH sense of *przy-*

[53] A Zipfian distribution is a distribution in which "there are only a few words of very high frequency and large numbers of words of low frequency" (Divjak: in press).

makes a very good candidate for a construction that could be built and used by native speakers. The design of this sentence-sorting experiment did not allow for investigating this issue – it might be an attractive avenue for future research.

Roz- exhibits a somewhat different kind of distribution, one in which all senses received substantially more coverage in the data than in the other two prefixes. Additionally, the distributions among the senses of *roz-* differed from that of the APPROACH sense of *przy-*. All 166 occurrences of APPROACH in *przy-* were distributed among 27 different verbs (which translates into 6.15 occurrences per verb), while the 99 occurrences of the DISPERSION sense of *roz-* were distributed among 65 different verbs (1.52 occurrences per verb for DISPERSION). The lower the number of occurrences per verb, the more "evenly" distributed a prefix sense is. If the number of occurrences per verb amounted to one, it would mean that each verb occurs only once. Conversely, if the number of occurrences of a verb was equal to the number of occurrences of a given sense, it would mean that this sense of a prefix occurs with only one verb in the sample. Following that, the DISPERSION sense of *roz-* is much more "evenly" distributed than the distribution of APPROACH in *przy-*. It would also be quite difficult to find one prototypical verb for DISPERSION amongst the verbs that occurred in the dataset. Goldberg (2006) and Ellis & Ferreira-Junior (2009) argue that if a construction has a strong and frequently occurring prototype, native speakers can acquire it more easily. On the other hand, it has also been argued that a wide variety of lexemes occurring with the same construction would provide many different contexts in which this construction could be used and thus increase its productivity. If a construction can be applied in many different contexts, it would be easier for native speakers to extend it beyond the lexical items with which they have already encountered the construction (Wonnacott, Newport & Tanenhaus 2008: 201). Since the DISPERSION sense of *roz-* occurs with so many different verbs, speakers encounter it in many different contexts; thus, speakers might be more likely to build a construction for this sense and, perhaps, apply it to new lexical items.

When it comes to the two remaining senses of *roz-*, it would be somewhat more difficult to draw robust conclusions, because the data for them are sparser: the INTENSITY sense of *roz-* occurred 65 times (over 32 different verbs; 2.03

occurrences per verb), whereas OPPOSITION occurred only 36 times (over 16 different verbs; 2.25 occurrence per verb). Among much fewer sentences (65), the INTENSITY sense of *roz-* occurred with more verbs (32) than the APPROACH sense of *przy-*, which had 27 different verbs distributed among 166 occurrences. This gives potentially many different lexical items that INTENSITY could pair with. We could surmise that – provided that the distribution holds for the entire "population" of *roz-* verbs – these two senses of *roz-* are also much more "evenly" distributed than the APPROACH sense of *przy-*. Even though the amount of data for INTENSITY and OPPOSITION is low, the high number of verbs in comparison to the number of occurrences gives a reason to think that all the senses of *roz-* could form coherent usage-based constructions in the minds of native speakers of Polish.

To sum up, in the case of *po-* and *przy-,* the sense distribution dominated by unpredictable (IDIOSYNCRATIC) verbs or a very abstract (ASPECTUAL) sense of prefix combined with extremely low numbers for the more concrete senses might make it difficult for language users to build constructions for the concrete senses of those prefixes. Users would need much more exposure to language in order to obtain enough input to be able construct those categories. The only exception in those two prefixes was the APPROACH sense of the prefix *przy-*, which makes a good candidate for a construction that could be built and used by native speakers of Polish. The design of the sentence-sorting experiment did not allow for investigating whether a construction for only one sense of a prefix could be built by native speakers; this issue constitutes an avenue for further research. The distribution of senses in *roz-* is much more 'even', in the sense that the more concrete senses occur much more often in relation to IDIOSYNCRATIC verbs, when compared to *po-* or *przy-*. Consequently, it seems much more likely that speakers could build categories for the different senses of the prefix *roz-*, which was reflected in the results of the sentence-sorting experiment.

6.6. Interim conclusions

The results of the corpus study on three Polish prefixes, *po-, przy-,* and *roz-*, presented in Chapter 5 suggested that speakers might not be able to build maximally general categories for prefixes as a whole. The aim of the sentence sorting experiment discussed in the present chapter (Chapter 6) was to investigate the question of whether native speakers of Polish could build (less general) categories for the different senses of the prefixes. The primary hypothesis was that if speakers group the sentences containing verbs with the same sense of a prefix, they are able to perceive similarities between different verbs containing the same sense. Consequently, if they are able to perceive similarities between those verbs, it is likely that they build categories for the different senses of prefixes. The results of the experiment indicate that speakers are likely to build categories for prefix senses only in the case of the prefix *roz-,* because they consistently grouped verbs with the same sense of the prefix together. With the other two prefixes, participants most likely used other criteria for the grouping of sentences, which indicates that they might not have built categories for the senses of *po-* and *przy-*.

The discrepancy between *roz-* and *po-* and *przy-* can be explained by the frequency distribution of the prefix senses. For *po-,* each prefix sense occurred only in around 1-4% of cases, which is a very low number. Almost 90% of examples with the prefix *po-* contained verbs that belonged either to the IDIOSYNCRATIC group (no coherent prefixal meaning could be discerned) or to the PURELY ASPECTUAL group (the only "function" of the prefix was to make an imperfective verb perfective). To put it differently, in an overwhelming majority of *po*-verbs, the prefix either did not carry any coherent meaning or it performed a maximally general function of making the verb perfective. Speakers would not have enough input to generalise from and create general categories for the prefix senses. In the case of *przy-*, the situation was similar but with one difference – one sense, APPROACH, occurred significantly more frequently than the other senses. The APPROACH sense and the IDIOSYNCRATIC and ASPECTUAL groups jointly accounted for more than 90% of all examples in the dataset, which suggests that, apart from APPROACH, speakers might not be able to build general usage-based categories for the senses of *przy-*. Unfortunately, the nature of the sentence-sorting experiment presented in this chapter did not allow for investigating whether speakers could build a construction for only one sense of a prefix. In contrast, the distribution of *roz-* differed substantially from the distribution of senses of the other two prefixes. Even though the IDIOSYNCRATIC group accounted for more than a half of examples, the other prefix senses occurred in between 7% to 19% of cases. Such frequency might constitute enough input for the speakers to be able to abstract from to create general constructions for the senses of the prefix *roz-*.

Overall, the sentence sorting experiment showed that speakers might not have salient categories for the senses of all prefixes. In the case of *po-* and *przy-*, speakers might not abstract from prefixed verbs to create general categories for prefix senses – instead, they might rely on more specific categories for each prefixed verb. For *roz-*, speakers are more likely to build general categories for the different senses of the prefix. Another possibility is that the categories for the senses of *po-* and *przy-* are too weak for the speakers to use them in their daily experience with language. Hence, the prefix senses might have been too weak as a cue for the participants to use them as a criterion for sorting in the experiment. What if we could **force** speakers to make generalisations about prefix senses? In that case, perhaps, they would be able to use even the weakest cues for prefix senses. The next chapter will discuss a nonce-verb experiment that will expose speakers to previously unseen invented prefixed verbs and compel them to come up with interpretations of those verbs.

Chapter 7: Polish prefixes: nonce-verb experiment

7.1. Introduction

As we could see in the results of the sentence sorting task (presented in Chapter 6), the participants did not seem to have high-level general constructions for the different senses of the prefixes *po-* and *przy-*, while there was an indication that they might have more general constructions for the senses of *roz-*. The sorting task involved relatively frequent verbs that actually exist in the Polish language, which were used in plausible everyday contexts. What if we pushed native speakers to the extreme and asked them to try and interpret words they have never encountered before? Would they use (or come up with) more general categories for the prefix senses to make sense of novel linguistic stimulus in such an unusual situation? These questions were investigated with an experiment that consisted of a forced choice task involving Polish nonce verbs (i.e. verbs invented solely for the purpose of this experiment).

The participants in the experiment were presented with sentences with a verb that does not exist in the Polish language (i.e. a nonce verb) to which one of the studied prefixes was attached and asked to choose one out of three possible interpretations. They were not given the definition of the prefixed verb – only the meaning of its non-prefixed "base" counterpart was explained with a comprehensive contextual definition. The interpretations of the sentences were either based on the senses of a given prefix proposed in the literature (for an overview of the senses of *po-, przy-*, and *roz-,* see Section 5.1.1.) or – as a control condition – on senses of other prefixes. The aim of the experiment was to see whether participants robustly choose the interpretations based on the senses of the prefixes present in the experimental sentences ("primary interpretations") or the senses of other prefixes ("control interpretations"), for instance, whether in sentences with verbs containing the prefix *po-*, the participants choose interpretations based on the senses of *po-* or the interpretations based on senses of other prefixes. If they consistently choose primary interpretations over control interpretations, we could conclude that native speakers of Polish are able to refer to (or come up with) general constructions for prefix senses in situations where this is crucial for the understanding of linguistic stimulus.

7.2. A few words on nonce words

The very concept of nonce words deserves explanation. According to Hartmann & James, a *nonce word* is "[a] word or phrase coined for a particular occasion" (2002: 100). In linguistic experiments, a *nonce word* is a previously non-existent word invented for the purposes of a given experiment. Berko's famous (1958) study introduced *nonce words* into the arsenal of experimental methods in linguistics as well as the nonce word *wug*, known by almost every student of English linguistics. Berko (1958) asked English-speaking children to supply the plural forms of invented nouns (such as *wug*) and past and progressive forms of invented verbs (such as *zib*) in order to see whether they would be able to extrapolate the ending from other regular words. The participants of the experiment were able to systematically supply the endings for the nouns, which was taken by Berko (1958) as evidence that children already possess morphological rules. To sum up, in a typical nonce-word experiment, participants inflect invented words, and if they inflect the words with a given ending more frequently than chance, the ending is regarded as regular.

According to Dąbrowska (2004: 237), "the ability to inflect nonce words in an experimental setting has been regarded as the gold standard of productivity". In other words, if participants robustly use an affix to inflect non-existing words, it means that the rule (or schema) encoded by this affix is productive. In the present thesis, the nonce-word experimental paradigm will be altered in order to make it suitable for the investigation of the semantics of prefixed verbs. Instead of being asked to inflect nonce verbs or nouns, the participants of the experiment were asked to choose between three possible meanings of a nonce verb with a prefix attached to it. Rather than investigating the formal productivity of an inflection, this study tested whether the participants could extrapolate the **meaning** of a prefix to words that they had not encountered prior to the experiment.

7.3. Structure of the experiment

Participants were presented with 20 experimental contexts. Each context included (1) a non-prefixed nonce verb with a comprehensive definition; (2) a sentence with a prefixed verb derived from the non-prefixed nonce verb explained in the definition; (3) three possible interpretations of the sentence. Subsequently, participants were asked to choose the interpretation of the sentence that they found the most plausible out of the three given options. No time limits were imposed, and the participants could go back to already answered questions if they felt the need to reconsider their choices. Full instructions for the task can be found in Appendix 1B. Apart from the experimental questions, participants were also asked to provide some supplementary information: demographics, reading habits, professional situation, and foreign language competencies (not all questions were mandatory). There were three versions of the experiment – each contained a different set of experimental contexts.

The study was given ethics approval from the University of Sheffield. The questionnaires were delivered electronically via the Qualtrics[54] platform and distributed through an anonymous link over the period from 19 April 2017 to 29 April 2017. The system randomly assigned one version of the experiment to each participant, in roughly equal numbers.

7.3.1. Stimuli

First, 20 nonce-verbs were generated with the use of Wuggy, a nonce-word generator[55]; subsequently, a definition was created for each verb. Next, 60 sentences with prefixed derivates of the previously generated verbs were created (20 per prefix). Each of the 60 generated sentences was given three possible interpretations (1) an interpretation based on the most plausible sense of the prefix in the experimental sentence (primary choice); (2) an interpretation based on another sense of the prefix in the sentences (secondary choice); and (3) an

[54] http://www.qualtrics.com
[55] http://crr.ugent.be/programs-data/wuggy

interpretation based on the sense of a prefix different from the prefix in the sentence. Overall, 60 different contexts of the following fashion were created:

DEFINITION: Jak każde inne stworzenie w całym wszechświecie, Konstrulianie lubią spędzać czas na robieniu fajnych rzeczy**. Robić coś fajnego to *hączyć***.

'Just like any other creature in the entire universe, Construlians like to spend their time doing cool things. To do cool things is ***hączyć.***'

TASK: Co znaczy poniższe zdanie? 'What does the sentence below mean?'

SENTENCE: *Pohączyli i byli bardzo zadowoleni.* ['***po****hączyć*$_{3SG.MASC}$ and were very satisfied']

ANSWERS:

1. Hączyli przez chwilę i byli bardzo zadowoleni. 'They *hączyć*$_{3SG.MASC}$ for a while and were very satisfied'
2. Co i raz hączyli, dzięki czemu byli bardzo zadowoleni. 'They *hączyć*$_{3SG.MASC}$ once every now and then and were very satisfied'
3. Hączyli długo. Tak długo, aż poczuli się zadowoleni. 'They *hączyć*$_{3SG.MASC}$ long. So long, until they felt satisfied'

The definition in the experimental context always contained one or two sentences of introduction, whose function was to set the scene and explain the behaviour of the "alien population". The main definition – in which the meaning of the nonce verb was explained – followed the introduction. Each experimental context also contained a question 'What does the sentence below mean?', which was the same for each context and indicated to participants what they needed to do. The question was followed by a sentence that contained the verb explained in the definition with one of the studied prefixes attached. The sentences contained one or two clauses at most – they were short enough to avoid biasing participants in favour of any interpretation, but not too short so as to remain plausible in Polish.

After the definition, the question, and the example sentence, participants were presented with three options to choose from. These answers were designed in such a way as to reflect one of the prefix senses, but they always contained an **unprefixed** version of the verb. The prefix sense was reflected, for instance, with an adverbial – *for a while* in the first answer in the above context reflects the SOMEWHAT sense of the prefix *po-*. Two out of three options were designed to reflect one of the senses of the prefix used in the example sentence (in the context above, it was the prefix *po-*). One of those two options, the one which best fitted the context, was coded as the primary (expected) answer, while the other, less fitting, was coded as the secondary answer. The coding of the answers was based on which answer best fitted the context according to the researcher's intuition. The third option was designed to reflect a sense of a different prefix, for instance, a verb with the prefix *po-* was given an answer that would rather fit the prefix *na-* – this answer was the confounder. In the example context above, (1) is the primary expected answer, (2) is the secondary answer, whereas (3) is the confounder.

The stimuli were divided into three different versions of experiment (20 stimuli per version). The versions contained roughly equal number of verbs with each prefix, however, since there were three prefixes and 20 stimuli per version, one prefix always received one stimulus fewer in each version. Overall, however, each prefix had the same number of stimuli (20) when summed over all three versions. All stimuli can be found in Appendix 4 (and their translations in Appendix 5).

7.3.2. Participants

Participants were recruited on the official Facebook groups for the University of Warsaw and the Warsaw School of Economics. Facebook viewers were encouraged to share the link to the study with their peers. Eventually, the Facebook post was shared on other Facebook pages, which resulted in a large number of participants.

Overall, the study was completed by 2498 participants (2050 females; 82.06%). In the questions about reading habits, the overwhelming majority of participants (91.79%) declared that they either read more than their peers (1169; 46.80%) or the same amount of time (1124; 44.99%). The mean participant age was 31.96 years (median = 30, sd = 9.38, min = 16, max = 77).

7.4. Results

The frequencies of all participants' answers to each question were tallied and summarised as a table – a table with all frequencies broken down by prefix can be found below (Table 7.1.). Participants' answers are also presented as a parallel line plot (Figure 7.1. below).

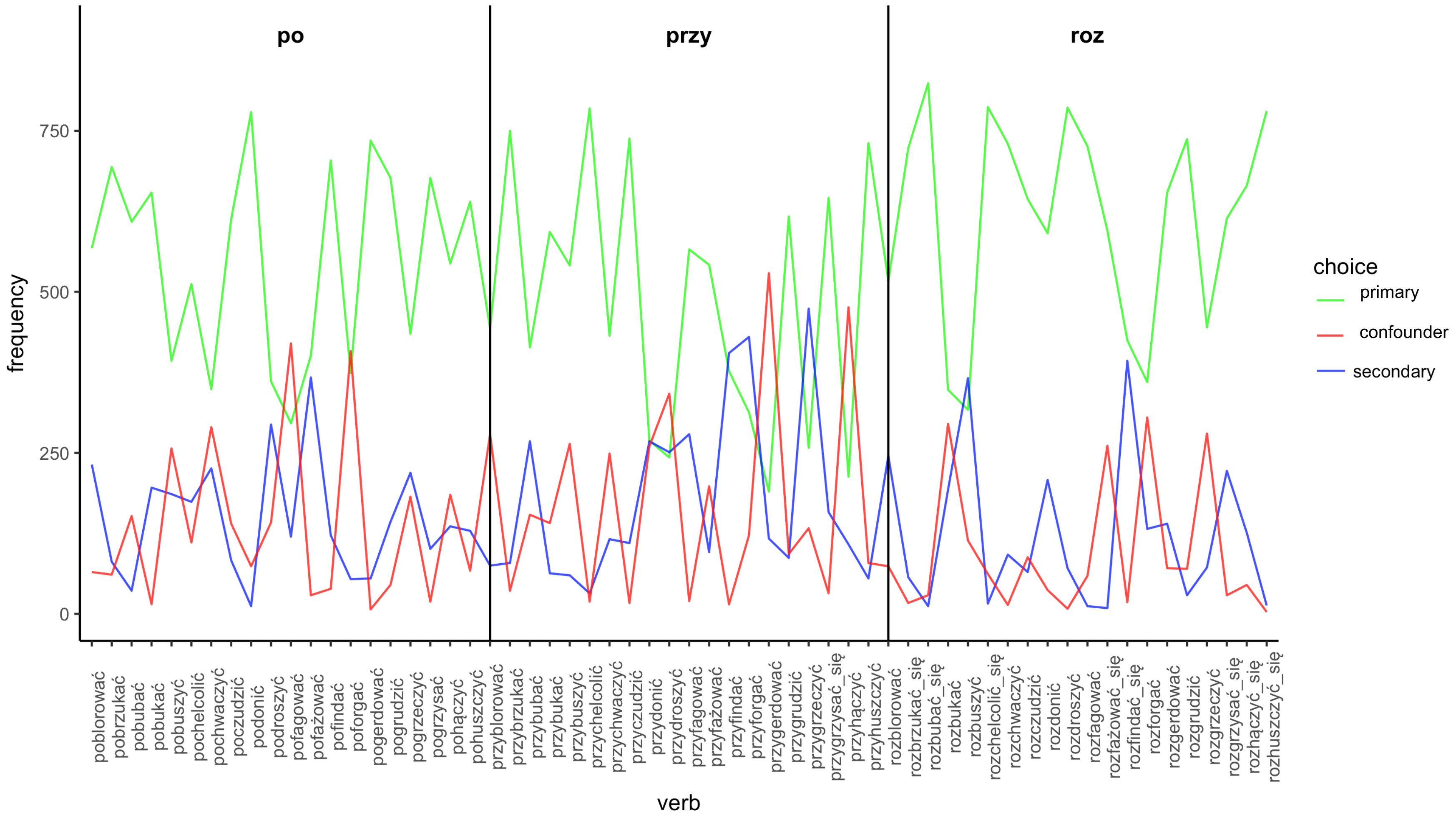


Figure 7.1. Parallel line plot of participant choices (all prefixes)

PO-											
CHOICE	**poblorować**	**pobukać**	**pochwaczyć**	**podonić**	**pofindać**	**pogrudzić**	**pohączyć**	**pobrzukać**	**pobuszyć**	**poczudzić**	
primary	568	654	349	779	704	677	544	694	393	613	
secondary	232	196	226	12	122	143	136	81	186	83	
confounder	65	15	290	74	39	45	185	61	257	140	
Choice	**pofagować**	**poforgać**	**pogrzeczyć**	**pohuszczyć**	**pobubać**	**pochelcolić**	**podroszyć**	**pofażować**	**pogerdować**	**pogrzysać**	**TOTAL**
primary	296	374	435	640	609	512	**361**	**401**	735	677	11015
secondary	120	54	219	129	36	174	**294**	**367**	55	101	2966
confounder	420	408	182	67	152	111	142	29	7	19	2708

PRZY-											
CHOICE	**przybrzukać**	**przybuszyć**	**przyczudzić**	**przyfagować**	**przyforgać**	**przygrzeczyć**	**przyhuszczyć**	**przybubać**	**przychelcolić**	**przydroszyć**	
primary	750	541	738	566	313	258	731	414	785	243	
secondary	79	60	110	279	430	474	55	268	32	251	
confounder	36	264	17	20	122	133	79	154	19	342	
Choice	**przyfażować**	**przygerdować**	**przygrzysać się**	**przyblorować**	**przybukać**	**przychwaczyć**	**przydonić**	**przyfindać**	**przygrudzić**	**przyhączyć**	**TOTAL**
primary	542	190	646	443	593	432	**269**	377	617	213	9661
secondary	96	117	158	75	63	116	**268**	**405**	87	108	3531
confounder	198	529	32	279	141	249	**260**	15	93	476	3458

ROZ-											
CHOICE	**rozbubać się**	**rozchelcolić się**	**rozdroszyć**	**rozfażować się**	**rozgerdować**	**rozgrzysać się**	**rozblorować**	**rozbukać**	**rozchwaczyć**	**rozdonić**	
primary	824	787	786	595	654	614	517	348	730	591	
secondary	12	16	71	9	140	222	245	193	92	208	
confounder	29	62	8	261	71	29	74	295	14	37	
Choice	**rozfindać się**	**rozgrudzić**	**rozhączyć się**	**rozbrzukać się**	**rozbuszyć**	**rozczudzić**	**rozfagować**	**rozforgać**	**rozgrzeczyć**	**rozhuszczyć się**	**TOTAL**
primary	**425**	737	665	723	317	644	726	360	445	781	12269
secondary	393	29	126	57	**366**	65	12	132	72	13	2473
confounder	18	70	45	17	114	88	59	305	280	3	1879

Table 7.1. Frequencies of participant choices (by prefix)

Before we analyse the results of the questionnaire, a brief explanation of the parallel line plot in Figure 7.1. is in order. All verbs used in the questionnaire were plotted on the X axis, whereas the frequencies of different answers were plotted on the Y axis. Each line corresponds to one type of interpretation (answer): green represents the primary interpretation, blue represents the interpretation involving the secondary sense of the prefix, while red stands for the interpretation involving the sense of a different prefix. One caveat about this type of plot is that the lines **do not** represent trends, because it is not a time-series plot – the data points were connected with lines for the sake of convenience and easy interpretation of the plot. The orthodox plotting choices for a categorical variable (verb is categorical) would have been a dot plot or a stacked bar plot, however, the interpretability would have suffered dramatically.

The lines in Figure 7.1. show that the answer with the primary interpretation outstripped the other answers by an order of magnitude in most cases – the green line remains above the blue and red lines in most parts of the plot. We can also see, however, that with some verbs the frequency of the non-primary answer was higher than that of the primary answer. The raw frequencies of answers are presented in Table 7.1. above. The highest frequency for each verb is highlighted with a colour – if the primary interpretation was the most frequent answer, the colour is green; the secondary interpretation is blue; while the confounding interpretation is red. The participants chose the primary answer more frequently than any other answer in 51 out of 60 experimental conditions (85%) – the ratio of the primary answer to other answers varied from 22.73% (*przygrzysać*) to 97.99% (*przygrudzić*). The verbs for which the frequency of the secondary answer was higher than the frequency of the primary answer are as follows: *przyforgać, przygrzeczyć, przyfindać, rozbuszyć*; the verbs for which the frequency of the other-prefix answer was higher than the frequency of the primary answer are: *pofagować, poforgać, przydroszyć, przygerdować, przyhączyć*. To test the statistical significance of the differences, chi-squared tests were run verb-wise, and one verb missed statistical significance: *przydroszyć* ($\chi2 = 2.11952$, df = 2, $p = 0.3465$, $\alpha = 0.000833$)[56]. The

[56] The 'standard' α level of 0.05 has been corrected for multiple comparisons using Bonferroni correction, thus $\alpha = 0.000833$.

primary answer was chosen more frequently than the other answers – and the result was statistically significant – in 51 out of 60 verbs.

Let us also look at some of the experimental contexts in more detail. The two contexts in which the participants chose the other-prefix answer more frequently (and the difference was statistically significant) were the contexts based around the verbs *przygerdować* and *przyhączyć*. In the case of *przygerdować,* the target sentence was *Przygerdowała drzwi* 'She locked the doors' and the most frequently chosen answer was *Zamknęła na zasuwki wszystkie drzwi w domu* 'She locked all the doors in the house'. The plural object *wszystkie drzwi* 'all (the) doors' was designed to evoke a DISTRIBUTIVE reading, compatible with the prefix *po-* rather than *przy-*, and since this sentence contains a verb with *przy-*, rather than *po-*, participants were expected to reject this answer in favour of the other two answers. Apart from the DISTRIBUTIVE reading, the ASPECTUAL reading would also be compatible with this answer, because an ASPECTUAL reading is potentially neutral as to the type of object it receives, and it can well accept the plural object without the need for a DISTRIBUTIVE reading of the prefix. The participants must have regarded the ASPECTUAL reading as the most plausible reading of the prefixed verb in this example, and they, consequently, chose this answer as the preferred answer.

For *przyhączyć,* the target sentence was *Miała chwilę, więc przyhączyła* 'She had a while so she *przyhączyć*' and the most frequently chosen answer was *Zaczęła haczyć, ponieważ akurat miała chwilę* 'She started *haczyć* because she had a while'. This answer contains the meaning of BEGINNING something, which is more characteristic for the prefix *za-* than *przy-* (the prefix in the target sentence). The participants were expected to reject it as incompatible with the target sentence, because, to the best of my knowledge, BEGINNING has not been postulated as one of the possible senses of the prefix *przy-*. That notwithstanding, beginning an action is a complete act in itself (an act of beginning something) and carries a perfective, i.e. ASPECTUAL function. In this case – similarly to the previous contexts with the *przygerdować* – the participants must have considered the ASPECTUAL reading as the most plausible option in this context, while the readings (1) SOMEWHAT and (2) INTENSELY would likely have required more supporting context for the readers to infer the expected prefixal sense.

The other two contexts where the other-prefix answer was more (or approximately equally) frequent than the other two answers are also interesting and we will now inspect them in detail. With the verb *pofagować,* the target sentence was *Pofagowali całą rodziną* 'The entire family cooked beans', while the most frequently chosen answer was *Pofagowali i wspólnie skończyli* 'They were cooking beans and finished together'. The two other answers, more directly related to the postulated senses of the prefix *po-,* must have been regarded by the participants as not plausible enough. Most likely, the word *skończyli* 'finished' caused the ASPECTUAL reading of the verb with *po-* in the most frequently chosen answer. The target sentence did not have any contextual hints as to which reading the respondent should select, so they defaulted on the ASPECTUAL reading. The context also has no object – in order for the DISTRIBUTIVE reading of the prefix *po-* (the primary answer) to be inferred, the verb should be transitive and the context should probably have a plural object. The SOMEWHAT reading would have appeared if the sentence had an adverbial of time that would indicate a very short duration, e.g. *przez chwilę* 'for a while'. For *poforgać,* the target sentence was *Poforgał ogródek*, while the most frequently chosen answer was *Forgał ogródek tak długo, aż skończył*. The situation with *forgać* is very similar to the one with *fagować – aż skończył* 'until he finished' has most likely induced an ASPECTUAL reading. The long duration, *tak długo* 'so long' in conjunction with *aż skończył* was predicted to induce a COMPLETENESS or THOROUGH reading, more compatible with the prefixes *wy-* or *prze-*. Hence, the expectation was that respondents would reject this sense in favour of one of the other two senses in this condition.

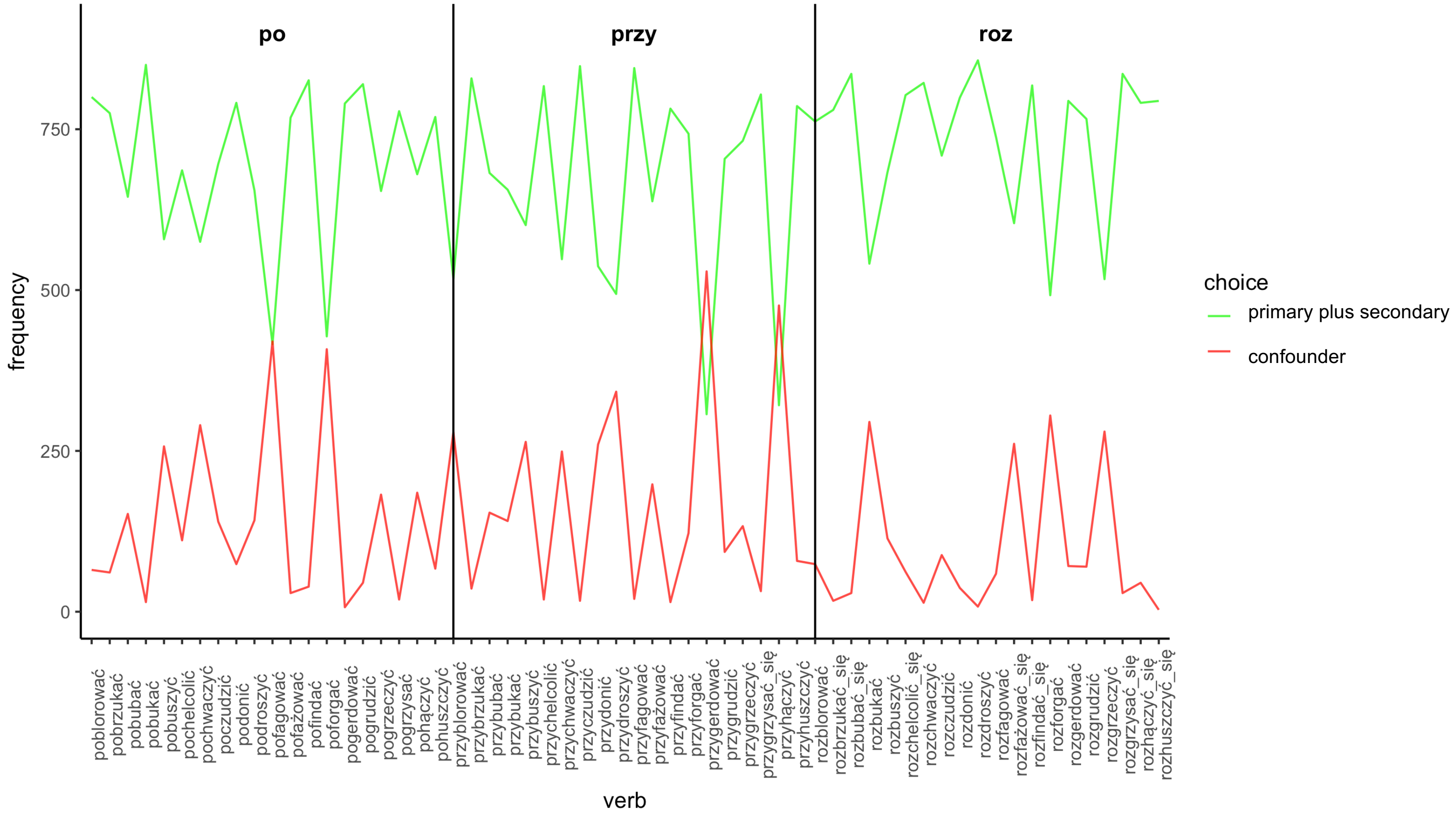


Figure 7.2. Parallel line plot of participant choices (all prefixes); primary and secondary answers summed

Let us look at one more plot (Figure 7.2.), which – similarly to Figure 7.1. – presents the frequencies of participants' answers plotted against different verbs. In contrast to Figure 7.1., the plot in Figure 7.2. presents only two distributions: the sum of the primary and secondary answers (green line), and the other-prefix answers (red line). In other words, the graph shows the difference between how often the participants chose answers that included **any** (either primary or secondary) sense related to the prefix in question versus answers containing a sense of another prefix. Collapsing the primary and secondary categories takes into account the possibility that the prefix sense chosen as primary was not, in fact, the most likely meaning in a given context – either the option that the secondary meaning was more plausible or the option that the two senses were equally plausible. This graph shows quite clearly that the answers with senses related to the postulated prefix senses (i.e. the primary and the secondary answers) were dominant across the overwhelming majority of experimental contexts. Only in the case of two verbs was the other-prefix answer more frequent: *przygerdować* and *przyhączyć*. Similarly to the non-combined data, chi-squared tests were run to verify whether the differences between the distributions of the two types of answers were statistically significant. The differences missed statistical significance in two verbs: *pofagować* ($\chi2 = 0.00733$, df = 1, p = 0.932, $\alpha = 0.000833$), and *poforgać* ($\chi2 = 0.4785$, df = 1, p = 0.261, $\alpha = 0.000833$)[57], all other differences were statistically significant. Overall, answers with the interpretation based on a sense of the prefix used in a given experimental sentence (primary or secondary answer) were preferred – and the difference was statistically significant – in 56 out of 60 cases. This result indicates that the participants might have some general constructions for the prefix sense, because they overwhelmingly chose answers that contained either of the senses of the target prefix over answers that contained a sense of a non-target prefix (i.e. the confounding prefix).

[57] The 'standard' α level of 0.05 has been corrected for multiple comparisons using Bonferroni correction, thus $\alpha = 0.000833$.

7.5. Discussion

As we could see in the results above, the participants picked the answer containing the primary interpretation of the verb with remarkable consistency – in 47 out of 60 cases, and if we count any sense related to the prefix in a given sentence (either primary or secondary) the number will reach 56 out of 60 cases. The most important conclusion that can be drawn from this outcome is that the participants – and, potentially, native speakers in general – can take advantage of some sort of a general prefix meaning to infer the meaning of a prefixed verb that they have never seen before. When exposed to an extreme situation (i.e. novel lexical items), the participants needed to refer back to their previous linguistic experience and search for any hints that would help them interpret the unfamiliar linguistic input. The nature of those 'hints', however, cannot be determined by the methods used in this experiment – we still do not know whether the participants had already developed abstract prototypes of prefixes before taking part in the experiment or whether they conjured up one-off categories based on the previous exemplars of prefixed verbs they had accumulated so far.

If we take into consideration all of the above, we can say that native speakers of Polish might have (or can come up with) some generalisations about the different sense of prefixed verbs of which they can take advantage when faced with unfamiliar linguistic input. Those generalisations were used by the participants of this experiment to rule out implausible interpretations and to select the interpretations that were compatible with the context and the prefix a given verb received. The context allowed speakers to hone in on a specific interpretation, and when the context was insufficient or the provided interpretations were implausible, the participants would default on the maximally general 'purely' aspectual perfectivising function. The nature of the generalisations could not be assessed with this experiment, so it is still an open question whether speakers use those generalisations in day-to-day linguistic interactions or whether they only conjure them up in extreme situations such as this experimental task, based on the exemplars they have accumulated over time.

7.6. Interim conclusions

The results of the sentence-sorting task presented in Chapter 6 provided evidence in support of speakers being able to build general constructions only in one prefix: *roz-*. In the other two prefixes (*po-* and *przy-*), the evidence suggested that native speakers of Polish might not be able to build general constructions for the different senses of the prefixes. The sentence-sorting experiment only contained language that the participants knew and used daily – it did not contain any invented words or constructions. The results of the sentence-sorting experiment gave rise to a question: would speakers of Polish be able to come up with generalisations about the senses of each of the three prefixes (not only *roz-*) if presented with language that they have not experienced before? The ultimate aim of the nonce-verb experiment discussed in this chapter was to investigate this question.

The nonce-word experiment discussed in the present chapter exposed the participants to extraordinary language conditions – they were asked to interpret sentences containing invented (nonce) verbs combined with the prefixes *po-*, *przy-*, and *roz-*. The results of the experiment demonstrate that the participants consistently chose the primary sense in all prefixes under investigation, which means that they needed to use at least some generalisations about the prefixes. These generalisations, however, were used only in an extreme language situation – the participants were forced to make generalisations about prefix senses (or take a wild guess), because they would fail to understand the linguistic input otherwise. This means that such generalisations might not be used by speakers in their everyday communicative situations and, hence, they might have been too weak to serve as a criterion for grouping in the sentence-sorting experiment, which was based on everyday language. The participants might already have had categories for the different prefix senses, but those categories might not be strong (or useful) enough to be used during everyday production and interpretation of language. Alternatively, the participants might not have had categories for some prefix senses prior to the experiment and only formed ad hoc generalisations based on salient exemplars of prefixed verbs that they had encountered throughout their experience with language – this issue might be an attractive avenue for further investigation.

Chapter 8: Conclusions

Linguists thrive on structures and generalisations. A natural tendency of a linguist adopting a semasiological approach would be to seek to develop a maximally general category for a linguistic form if this form occurs sufficiently frequently for the linguist to discern it and consider it a candidate for a unit of language. For instance, the Polish 'light' reflexive marker *się* occurs with many verbs, and the tendency of at least some linguists studying the Polish **language** has been to establish a category general enough to fit all occurrences of *się* in Polish reflexive verbs (e.g. Klemensiewicz 1946; Nagórko 2007; Dancygier 1997; Tabakowska 2003a). Linguists might be 'naturally' inclined to look for maximally general categories, but do native speakers build and use those general categories? If the primary aim of linguistics is to describe language as used by real language speakers, answering the above question appears to be of paramount importance.

Two linguistic phenomena were studied for the purposes of exploring the questions outlined above: Polish reflexive verbs and Polish prefixed verbs (containing prefixes *po-, przy-*, and *roz-*). Each of them was investigated using the same methodology so that the results could be compared with each other. First, behavioural profiles based on corpus data were built to check whether the properties of contexts in which Polish reflexive verbs and prefixed verbs appear would enable native speakers of Polish to build maximally general categories for the Polish reflexive markers *się* and *siebie*, and the prefixes *po-, przy-*, and *roz-* respectively. Then, experimental studies were conducted to investigate whether speakers could build categories not for the reflexive marker or prefixes as a whole but for each of the senses postulated in the literature for each of the phenomena. Thanks to the combined empirical approach, different levels of generality were explored for reflexive and prefixed verbs. It allowed for establishing the highest level of generality at which native speakers of Polish might build categories for each of the studied phenomena.

The behavioural profiles built for Polish prefixed verbs and reflexive verbs do not warrant postulating one maximally general category for the 'light' reflexive marker *się* or any of the prefixes studied (i.e. *po-, przy-,* and *roz-*); the only maximally general construction supported by the behavioural profile data was the 'heavy' reflexive marker *siebie*. Corpus data cannot fully replace a longitudinal acquisition study, but since there are no contextual properties that would facilitate the formation of general categories for *się* and the prefixes, we could conclude that it would be difficult, if not impossible, for a native speaker to acquire such categories. As far as the less general categories are concerned, the experimental studies provided evidence that native speakers of Polish might be able to build separate categories for the different senses of the marker *się* and the prefix *roz-* and use them in their day-to-day experience with language. In contrast, the results of experiments for *po-* and *przy-* indicate that speakers might not be able to build categories for the different senses of those prefixes. That notwithstanding, speakers might use some generalisations about the senses of those prefixes in extreme communicative situations, for instance, when they need to infer the meaning of a prefixed verb they do not know based on the meaning of a prefix and the meaning of an unprefixed verb they already know.

The results of the studies on Polish reflexive verbs and prefixed verbs presented in this thesis have implications for a number of areas of research within usage-based linguistics and descriptive linguistics of the Polish language. First and foremost, the results tie in with previous research suggesting that native speakers might not always build and use on a daily basis the constructions proposed by linguists (see e.g. Dąbrowska 2008a; Perek 2015). One of the fundamental assumptions of usage-based linguistics is that speakers gradually build more general language categories (including grammatical constructions) by generalising over many less general categories. For instance, to build the construction for relative clauses in English, children first learn how to use concrete phrases such as *Here's the*… and then gradually develop the more general construction (Diessel & Tomasello 2000). If speakers develop general categories from more specific ones, we must entertain the possibility that they will stop at some point on the generality 'scale' and not develop any more general constructions, even if they could be postulated by linguists. The sole fact that a general language category can be proposed does not always mean that speakers will build the same category and use it when producing or processing language. The reflexive marker *się* is a case in point here – very general constructions for the marker have been proposed (Dancygier 1997; Tabakowska 2003a), but the research presented in this thesis did not provide evidence that would indicate that speakers could build those constructions. We must not assume that speakers have a given language category unless we provide evidence from actual language **usage** in favour of the category's existence. The discrepancy between linguists' grammars and speakers' grammars is by no means typical only for non-usage-based approaches to language research (be it descriptive, generative, or structural). Studies that subscribe to usage-based theories of language (e.g. Cognitive Grammar or Construction Grammar) will often stop at formulating a cognitively plausible description of a grammatical construction without attempting to produce any substantial empirical evidence (Dąbrowska 2016: 483–484). That was the case with Dancygier's (1997) and Tabakowska's (2003a) Cognitive Linguistic accounts of *się*, which, to my best knowledge, have not been investigated empirically prior to the studies presented in this thesis.

Should speakers be unable to build one general category for the marker *się* or the different senses of prefixes *po-* and *przy-*, the descriptive practice in Polish linguistics might need rethinking. Describing *się* as a 'defective' form of the pronoun *siebie*$_{\mathrm{ACC}}$ (Nagórko 2007: 155), for instance, implies that *się* is a single category. Similarly, listing the different senses for the prefixes *po-* or *przy-* (see e.g. Swan 2002 or Śmiech 1986) also implies that categories for those senses exist in the minds of speakers – of course, if we aim for the linguistic descriptions to be a reflection of actual speakers' knowledge of language. The evidence collected in the experiments and the corpus studies discussed in this thesis suggest that a single category postulated for *się* and the categories for the different senses of *po-* and *przy-* might have no corollary in the minds of native speakers of Polish. What it means in terms of linguistic descriptions is that, firstly, more emphasis should be put on the idiosyncratic nature of prefixed verbs with prefixes *po-* and *przy-*. Secondly, authors of descriptive grammars of Polish could refrain from postulating high-level generalisations about *się* as a whole, because those generalisations might not have any corollary in the minds of native speakers. These recommendations apply especially to the authors of pedagogical grammars, as the chief aim of such publications is to help non-native speakers acquire the language – including categories that do not exist in the minds of native speakers (e.g. one general category for *się*) could eventually hamper the acquisition of the language.

So far we have chiefly spoken about linguists' categories that speakers might not be able to build, but certainly, not **all** categories proposed by linguists will fail to converge with those used by native speakers. This thesis provided evidence in support of speakers having the categories for the different senses of the prefix *roz-* and the 'light' reflexive marker *się*[58] or a general category for the 'heavy' reflexive marker *siebie*. Linguists' expert intuitions are not wrong by default, but we can never know whether they reflect native speakers' knowledge of language unless we subject them to empirical tests. If we discover that speakers do not build general, or even less general categories, for a given phenomenon contrary to linguists' intuitions – for instance one category for the prefix *po-* or categories for each of the prefix's senses – what is the level of generality at which they stop? Do they only build individual categories for each prefixed verb or do they, perhaps, also build local categories for a few verbs with similar meanings? The limits of generality and speakers' ability to build and use more local generalisations are questions that need further research.

Usage-based linguistics advocates the use of converging evidence, because converging evidence can ensure that "the skewing effects of any one method will be canceled out by the other methods" (Lakoff & Johnson 1999: 80). Converging evidence coming from many different sources can also make stronger the argument in support of a given hypothesis. An option not frequently considered, however, is when evidence from different sources diverges. What then? Diverging evidence has the potential to be just as informative as converging evidence, and the results of the experimental studies on Polish prefixes discussed in this thesis are a case in point. The results of the sentence-sorting experiment and the nonce-verb experiment seem to diverge when it comes to the prefixes *po-* and *przy-*. The former experiment (sentence sorting) does not provide evidence supporting the

[58] Bear in mind that the different senses of the reflexive marker are not tantamount to one general category for the entire marker.

hypothesis that speakers build categories for the different senses of the prefixes, while the latter (the nonce verb task) indicates that speakers could use some categories for those senses in certain extraordinary situations. If we look closely at the divergent evidence from the two experiments, we will see that it shows when speakers might use categories for the different senses of *po-* and *przy-*: they might only be able to use them in extreme communicative situations when they could fail to understand novel language. To conclude, diverging evidence can motivate the linguist to look for alternative explanations for a given phenomenon and, in consequence, explore the matter further.

The diverging evidence from the experiments on *po-* and *przy-* raises questions about the nature of linguistic categories in general and how we should test their existence in speakers. Since the experiment participants could make generalisations about prefix senses in the task involving invented language (an unusual communicative situation) but not in the task involving everyday language, they might not have had those generalisations 'ready' when they took part in the study. In other words, they might not have had categories for those prefix senses prior to the experiment. It is therefore likely that the participants constructed ad hoc categories (Barsalou 1983) for the prefix senses to be able to fulfil the task. In usage-based linguistics, ad hoc categories have so far been usually considered in the context of figurative language (see e.g. Gibbs 1992; Gibbs 2007), but the results of the experiments on Polish prefixed verbs discussed in this thesis demonstrate that ad hoc language categories might also be relevant to grammatical constructions. It appears that speakers might not build and use in normal communicative situations some general categories proposed by linguists, but they might be able to construct ad hoc categories should such a communicative need arise. If speakers are indeed able to build ad hoc linguistic categories only for the purpose of fulfilling a particular communicative task, linguists must exert caution when designing experiments and drawing conclusions. We need to ascertain whether the constructions we are investigating could be built by speakers and used regularly or whether they are rather generated on an ad hoc basis. More research into the nature of linguistic categories and their permanence is necessary – as studies on ad hoc categories are scant in usage-based linguistics, it is an avenue certainly worth exploring.

A question that begs investigation is how linguistic categories are built, stored, and used – regardless of whether they are created ad hoc for a specific communicative task or stored permanently and used in speakers' daily language experience. The results obtained in the studies discussed in this thesis can be explained with reference to the two major theories of categorisation: the prototype theory and the exemplar theory. Even though the two theories have usually been considered incompatible, some studies (e.g. Vanpaemel & Storms 2008; Divjak & Arppe 2013) suggest that exemplars and prototypes might in fact be two opposite ends of a generality (or abstraction) spectrum. If we take the constructions for different senses of *się* and *roz-* into consideration, it seems likely that speakers will have formed prototypes for those senses, and they have a general concept of what the typical context for each sense would be. Having prototypes for the senses of *się* and *roz-* could have enabled the experimental participants to consistently group stimuli containing the same sense of *się* and *roz-* together, because the prototype would serve as a template for grouping the sentences. When it comes to the senses of *po-* and *przy-*, it is likely that the participants did not have prototypes for them, because they did not group the experimental sentences containing verbs with the same sense of a prefix together. Nevertheless, participants of the nonce-verb experiment consistently chose the relevant sense of a given prefix when interpreting the experimental sentences. As we mentioned earlier, they might have created ad hoc categories just for the purpose of completing the task. But how could those ad hoc categories be formed? The exemplar theory of categorisation could provide some answers to this question. A likely explanation is that the participants retrieved a number of exemplars of prefixed verbs from their memory similar to the nonce verbs presented in the experiment and selected the most probable meaning for each experimental context based on similarity to those exemplars. One must bear in mind that the above explanations are hypotheses and establishing the exact nature of the linguistic categories discussed in this thesis will require more research.

We could see that speakers might be able to build and use categories for the different senses of the prefix *roz-* but not the prefixes *po-* and *przy-*. Why would that be so? The most likely answer to this question lies in the frequency distributions of the prefix senses. Bybee (2010) suggested that a construction must exhibit a sufficiently high type frequency (i.e. occur with many different lexical items) for speakers to be able build the construction and extend it to new items. If a construction occurs with many different lexical items, it means that it occurs in many different contexts – this ensures that speakers can generalise over the particular tokens of a construction and come up with a more general category. In the sample used in the corpus study, the different senses of *roz-* occurred much more often in relation to all occurrences of the prefix than the senses of the other two prefixes, and they also occurred with a larger number of different verbs – thus, speakers of Polish might encounter *roz-* verbs with a greater variety of verbs than it is the case with *po-* and *przy-*. The different sense distributions found their reflection in the experimental results, where the participants grouped the experimental sentences according to the different senses of the prefix *roz-*, which was not the case with the remaining two prefixes. This result provides evidence in support of the hypothesis that in order for speakers to build and use a (general) construction in everyday communicative situations, the construction must occur frequently and in many different contexts. If a purported construction occurs infrequently and with a handful of different lexical items, it might not generalise well and remain confined to the limited number of lexical items (e.g. verbs) with which it occurs.

The low frequencies of the different senses of *po-* an *przy-* compared to the very high frequencies of idiosyncratic verbs or verbs in which the prefix only marks the perfective aspect suggest that these prefixes might have undergone semantic bleaching. In many cases, the only meaning import of the prefix in a prefixed verb as compared to the unprefixed verb is marking the perfective aspect (e.g. *bić* 'hit$_{\text{IMPF}}$' → ***po**bić* 'hit$_{\text{PFV}}$'). In other cases, there is no clear semantic import of the prefix, because the relation between the prefixed verb and the unprefixed verb may have been semantically transparent in the past, but no longer is (e.g. *znać* 'know' and ***przy**znać* 'admit'). The meaning of the prefix in most prefixed verbs has bleached to such an extent that it is no longer clear and transparent, and, as a consequence, speakers cannot build categories for the different senses of the prefix. This result shows that coupled with diachronic studies, the experimental and corpus studies employed in the research discussed in this thesis could be used for the investigation of loss of semantic meaning. If we could ascertain with diachronic studies that speakers might have been able to build a certain linguistic category in the past, we could also investigate with corpus and experimental research (of the sort discussed in this thesis) whether speakers could still build this category synchronically.

The final question one might ask is whether speakers' social backgrounds affect the types of linguistic categories they build. Dąbrowska's (2008a) studies indicated that this might be true when it comes to the educational attainments of speakers – highly educated speakers might develop different grammatical categories than their less educated counterparts. In the experiments discussed in this thesis, most speakers have had at least some tertiary education, which bears upon the interpretation of the experimental results. When it comes to the negative results – that is, the results that suggest that native speakers might not be able to build a given category – such a demographic composition of the participants makes the evidence even stronger. Usually, speakers with more years in education have had contact with more diverse linguistic input and thus they might be more likely to develop general linguistic categories (Dąbrowska 2008a: 947). Following that, if more educated speakers are unlikely to build certain categories (e.g. the different senses of the prefixes *po-* and *przy-*), it is even less likely that less educated speakers will be able to do it. The relationship is exactly opposite in the case of evidence supporting speakers' ability to build a given construction (e.g. the different senses of *się* and *roz-*). If the evidence suggests that more educated speakers might be able to build a construction, it does not follow automatically that speakers with fewer years of education will be able to do the same. Consequently, the finding that speakers might be able to build the categories for *się* and *roz-* should be further investigated with studies having subjects with fewer years of education as participants. Another social dimension that needs to be taken into account is gender. In this study, approximately 80 percent of the participants were women. Women have been shown to perform slightly better on verbal aptitude tests (Priess & Hyde 2010: 305), which might also mean that they could be slightly more likely to develop more general linguistic categories. The data collected as part of the studies discussed in this thesis does not allow for the effective investigation of this issue – the differences in linguistic categorisation between men and women could be an attractive avenue for future studies.

In sum, the studies discussed in this thesis have demonstrated that even though they are undoubtedly elegant and economical, big general categories postulated by linguists might not always be cognitively realistic. The sole fact that a general category can be postulated does not automatically imply that speakers will be able to build and use the category regularly for everyday language tasks. In other words, speakers' grammars do not always coincide with linguists' grammars. This result is of major importance to usage-based linguistics and even linguistics in general – it demonstrates that linguists must not tacitly assume that the linguistic categories they postulate will also be present in the minds of speakers. We rarely question the existence of the big generalisations, perpetuating categories and labels that have been used within the discipline for many years, for instance, the label reflexive pronoun (Pol. *zaimek zwrotny*) for Polish *się.* Unless we study those 'traditional' categories empirically, we can never know whether the categories we use for describing natural languages have any corollary in speakers' minds. Empirical research into linguistic categories as used by native speakers will make usage-based linguistics truly based in real language usage.

The studies discussed in this thesis also showed that it is crucial to explore different levels of generality for language phenomena – we cannot know a priori which level(s) will be the most relevant to language speakers. Here, carefully designed frequency counts could serve as a useful heuristic for assessing the likelihood that speakers build a given category. The results of the studies on prefixes showed that this likelihood correlates positively with the type frequency of a given construction – speakers need a large number of different contexts to generalise from to be able to build a general category.

Last but not least, the research presented in this thesis has demonstrated the power of **diverging** evidence. Certainly, in comparison to converging evidence, diverging evidence does not sound so impressive as the ultimate objective linguists should be aiming at. Nevertheless, it can have a very informative function, and it can compel a linguist to explore a given issue from different angles. Overall, the study has demonstrated the importance of empirical evidence in the research on language categories. In order to develop more realistic descriptions of the categories that speakers use, we need to adopt comprehensively empirical approaches – such as the one presented in this thesis – and see whether the categories we postulate as linguists withstand the scrutiny of real usage data.

References

Aitchison, Jean. 2004. The rhinoceros's problem: the need to categorize. In Christian J. Kay & Jeremy J. Smith (eds.), *Categorization in the History of English*, 1–18.

Barsalou, Lawrence W. 1983. Ad hoc categories. *Memory & cognition* 11(3). 211–227.

Beilock, Sian L., Thomas H. Carr, Clare MacMahon & Janet L. Starkes. 2002. When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied* 8(1). 6–16.

Bencini, Giulia M. L. & Adele E. Goldberg. 2000. The Contribution of Argument Structure Constructions to Sentence Meaning. *Journal of Memory and Language* 43(4). 640–651.

Berez, Andrea L. & Stefan Thomas Gries. 2009. In defense of corpus-based methods: A behavioral profile analysis of polysemous get in English. *Proceedings of the 24th NWLC*, 157–166. Seattle, WA: University of Washington.

Berko, Jean. 1958. The child's acquisition of English morphology. *Word Journal Of The International Linguistic Association* 14(March). 150–177.

Bogusławski, Andrzej. 1977. Polskie 'się' – słowo nie do końca poznane. *International Review of Slavic Linguistics* 2(1).

Boryś, Wiesław. 2008. *Słownik etymologiczny języka polskiego*. Kraków: Wydawnictwo Literackie.

Bybee, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form*. (Typological Studies in Language). Amsterdam, Philadelphia: John Benjamins.

Bybee, Joan. 2010. *Language, Usage and Cognition*. Cambridge, UK: Cambridge University Press.

Bybee, Joan & Clay Beckner. 2010. Usage-Based Theory. In Bernd Heine & Heiko Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*. Oxford, UK: Oxford University Press.

Bybee, Joan & Sandra A. Thompson. 1997. Three frequency effects in syntax. *Berkeley Linguistic Society* 23. 65–85.

Chomsky, Noam. 1995. *The Minimalist Program*. (Current Studies in Linguistics Series). Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 2002. *Syntactic structures*. 2nd ed. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2007. Approaching UG from Below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*. Berlin: Mouton de Gruyter.

Christensen, Jason Heath. 2011. The Prefix PO- and Aspect in Russian and Polish: A Cognitive Grammar Account. University of Kansas.

Coxon, Anthony Peter Macmillan. 1999. *Sorting data: collection and analysis*. Thousand Oaks, CA, USA: Sage Publications.

Dąbrowska, Ewa. 1997. The spatial structuring of events: A study of Polish perfectivizing prefixes. In Rene Dirven & Martin Pütz (eds.), *The Construal of Space in Language and Thought*, 467–490. Berlin: Mouton de Gruyter.

Dąbrowska, Ewa. 2004. Rules or schemas? Evidence from Polish. *Language and Cognitive Processes* 19(2). 225–271.

Dąbrowska, Ewa. 2008a. The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language* 58(4). 931–951.

Dąbrowska, Ewa. 2008b. The later development of an early-emerging system: the curious case of the Polish genitive. *Linguistics* 46. 629–650.

Dąbrowska, Ewa. 2008c. Questions with long-distance dependencies: A usage-based perspective. *Cognitive Linguistics* 19(3). 391–425.

Dąbrowska, Ewa. 2014. Recycling utterances: A speaker's guide to sentence processing. *Cognitive Linguistics* 25(4). 617–653.

Dąbrowska, Ewa. 2016. Cognitive Linguistics' seven deadly sins. *Cognitive Linguistics* 27(4). 479–491.

Dąbrowska, Ewa, Caroline Rowland & Anna L. Theakston. 2009. The Acquisition of Questions with Long Distance Dependencies. *Cognitive Linguistics* 20. 571–596.

Dancygier, Barbara. 1997. Reflexive Markers in Polish: Participants, Metaphors and Constructions. In Marjolijn H. Verspoor, Kee Dong Lee & Eve Sweetser (eds.), *Lexical and Syntactical Constructions and the Construction of Meaning*, 311–327. Amsterdam, Philadelphia: John Benjamins.

Diessel, Holger. 2013. Construction Grammar and First Language Acquisition. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*. Oxford, UK: Oxford University Press.

Diessel, Holger & Michael Tomasello. 2000. The development of relative clauses in spontaneous child speech. *Cognitive Linguistics* 11(1–2). 131–151.

Divjak, Dagmar. 2006. Ways of intending: delineating and structuring near-synonyms.

In Anatol Stefanowitsch & Stefan Thomas Gries (eds.), *Corpora in cognitive linguistics: corpus-based approaches to syntax and lexis*, 19–56. Berlin: Mouton de Gruyter.

Divjak, Dagmar. 2015. Exploring the grammar of perception: A case study using data from Russian. *Functions of Language* 22(1). 44–68.

Divjak, Dagmar. In press. *Frequency in Language. Context, Memory and Attention*. Cambridge University Press.

Divjak, Dagmar & Antti Arppe. 2013. Extracting prototypes from exemplars What can corpus data tell us about concept representation? *Cognitive Linguistics* 24(2). 221–274.

Divjak, Dagmar & Nick Fieller. 2014. Cluster analysis: Finding structure in linguistic data. In D. Glynn & J. A. Robinson (eds.), *Corpus Methods for Semantics: Quantitative studies in polysemy and synonymy.*, 405–441. Amsterdam, Philadelphia: John Benjamins.

Divjak, Dagmar & Stefan Thomas Gries. 2006. Ways of trying in Russian: clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2(1). 23–60.

Divjak, Dagmar & Stefan Thomas Gries. 2008. Clusters in the mind?: Converging evidence from near synonymy in Russian. *The Mental Lexicon* 3(2). 188–213.

Divjak, Dagmar & Stefan Thomas Gries. 2009. Corpus-based cognitive semantics: A contrastive study of phrasal verbs in English and Russian. In Barbara Lewandowska-Tomaszczyk & Katarzyna Dziwirek (eds.), *Studies in Cognitive Corpus Linguistics*, 273–296. Frankfurt am Main: Peter Lang.

Divjak, Dagmar, Natalia Levshina & Jane Klavan. 2016. Cognitive Linguistics: Looking back, looking forward. *Cognitive Linguistics* 27(4). 447–463.

Ellis, Nick C. 2012. What can we count in language, and what counts in language. In Stefan Th. Gries & Dagmar Divjak (eds.), *Frequency Effects in Language Learning and Processing*. Berlin: Mouton de Gruyter.

Ellis, Nick C. & Fernando Ferreira-Junior. 2009. Construction Learning as a Function of Frequency, Frequency Distribution, and Function. *The Modern Language Journal* 93(3). 370–385.

Everitt, Brian S., Sabine Landau, Morven Leese & Daniel Stahl. 2011. *Cluster analysis*. *Wiley series in probability and statistics*. New York: John Wiley & Sons.

Firth, John Rupert. 1957. *Papers in linguistics, 1934-1951*. Oxford, UK: Oxford University Press.

Frajzyngier, Zygmunt. 2000. Coding of the reciprocal function. In Zygmunt Frajzyngier & Traci S Curl (eds.), *Reciprocals: Forms and Functions*, 179–194.

Amsterdam, Philadelphia: John Benjamins.

Frajzyngier, Zygmunt & Traci S Curl (eds.). 2000. *Reflexives: Forms and Functions*. Amsterdam, Philadelphia: John Benjamins.

Freedman, David A. 2010. *Statistical Models and Causal Inference*. Cambridge, UK: Cambridge University Press.

Gerz, Daniela, Ivan Vulić, Edoardo Ponti, Jason Naradowsky, Roi Reichart & Anna Korhonen. 2018. Language Modeling for Morphologically Rich Languages: Character-Aware Modeling for Word-Level Prediction. *Proceedings of TACL 2018* 6. 451–465.

Gibbs, Raymond W. 1992. Categorization and metaphor understanding. *Psychological Review* 99(3). 572–577.

Gibbs, Raymond W. 2007. Why cognitive linguists should care more about empirical methods. In Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson & Michael J Spivey (eds.), *Human Cognitive Processing*, vol. 18, 2–18. Amsterdam, Philadelphia: John Benjamins.

Giermak-Zielińska, Teresa. 1979. *Polskie czasowniki przedrostkowe o znaczeniu przestrzennym i ich odpowiedniki w języku francuskim*. Wrocław: Zakład Narodowy im. Ossolińskich.

Gilquin, Gaëtanelle & Stefan Thomas Gries. 2009. Corpora and experimental methods: A state-of-the-art review. *Corpus Linguistics and Linguistic Theory* 5(2009). 1–26.

Glynn, Dylan. 2014. Correspondence analysis: Exploring data and identifying patterns. In Dylan Glynn & Justyna Robinson (eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy*, 443–485. Amsterdam, Philadelphia: John Benjamins.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at work: the nature of generalization in language*. Oxford, UK: Oxford University Press.

Goldberg, Adele E. 2009. The nature of generalization in language. *Cognitive Linguistics* 20(1). 93–127.

Górski, Rafał & Marek Łaziński. 2012. Typologia tekstów w NKJP. In Adam Przepiórkowski, Mirosław Bańko, Rafał Górski & Barbara Lewandowska-Tomaszczyk (eds.), *Narodowy Korpus Języka Polskiego*. Warszawa: Wydawnictwo Naukowe PWN.

Greenacre, Michael. 2007. *Correspondence Analysis in Practice*. London: Chapman and Hall/CRC.

Greenacre, Michael. 2010. Correspondence analysis. *Wiley Interdisciplinary Reviews: Computational Statistics* 2(5). Wiley-Blackwell. 613–619.

Greenacre, Michael & Jörg Blasius. 2006. Correspondence Analysis and Related Methods in Practice. In Michael Greenacre & Jörg Blasius (eds.), *Multiple Correspondence Analysis and Related Methods*, 3–40. London: Chapman and Hall/CRC.

Gries, Stefan Thomas. 2006. Corpus-based methods and cognitive semantics: The many senses of to run. *Gries, S. T., & Stefanowitsch, A. (Eds.). (2006). Corpora in Cognitive Linguistics: Corpus-Based Approaches to Syntax and Lexis. Berlin: Mouton de Gruyter.* 57–99.

Gries, Stefan Thomas. 2010. Behavioral profiles: A fine-grained and quantitative approach in corpus-based lexical semantics. *The Mental Lexicon* 5(3). 323–346.

Gries, Stefan Thomas & Naoki Otani. 2010. Behavioral profiles: A corpus-based perspective on synonymy and antonymy. *ICAME Journal* 34. 121–150.

Gries, Stefan Thomas & Anatol Stefanowitsch. 2004. Extending collostructional analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9(1). 97–129.

Gries, Stefan Thomas & Stefanie Wulff. 2009. Psycholinguistic and corpus-linguistic evidence for L2 constructions. *Annual Review of Cognitive Linguistics* 7. 163–186.

Gyongyi, Zoltan & Hector Garcia-Molina. 2005. Web spam taxonomy. *First International Workshop on Adversarial Information Retrieval on the Web*.

Harrell, Frank E. 2015. *Regression Modeling Strategies*. New York: Springer.

Harris, Zellig S. 1954. Distributional structure. *Word* 10(23). 146–162.

Hartmann, R. R. K. & Gregory James. 2002. *Dictionary of Lexicography*. New York and London: Routledge.

Hothorn, Torsten, Kurt Hornik & Achim Zeileis. 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics* 15(3). 651–674.

Iwata, Seizi. 2008. *Locative Alternation*. . Vol. 6. (Constructional Approaches to Language). Amsterdam, Philadelphia: John Benjamins.

Jakubíček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý & Vít Suchomel. 2013. The TenTen Corpus Family. *7th International Corpus Linguistics Conference CL 2013*, 125–127. Lancaster.

Janda, Laura A., Anna Endresen, Iulia Kuznetsova, Olga Lyashevskaya, Anastasiia Makarova, Tore Nesset & Svetlana Sokolova. 2013. *Why Russian aspectual* *prefixes aren't empty: prefixes as verb classifiers*. Bloomington (Indiana): Slavica Publishers.

Jansegers, Marlies, Clara Vanderschueren & Renata Enghels. 2015. The polysemy of the Spanish verb sentir: a behavioral profile analysis. *Cognitive Linguistics* 26(3). 381–421.

Kabacoff, Robert I. 2015. *R in Action: Data analysis and graphics with R*. 2nd ed. Shelter Island: Manning Publications.

Kemmer, Suzanne. 1993. *The Middle Voice*. Amsterdam, Philadelphia: John Benjamins.

Kemmer, Suzanne. 1995. Emphatic and reflexive-self: expectations, viewpoint and subjectivity. In Dieter Stein & Susan Wright (eds.), *Subjectivity and Subjectivisation: Linguistic Perspectives*. Cambridge, UK: Cambridge University Press.

Klemensiewicz, Zenon. 1946. *Gramatyka współczesnej polszczyzny kulturalnej w zarysie.* Wrocław, Warszawa: Książnica Atlas.

Krupianka, Aleksandra. 1969. O tzw. dopełnieniowej funkcji przedrostków czasownikowych. *Poradnik Językowy* 10. 560–570.

Kubiński, Wojciech. 1982. Polish 'się' constructions and their English counterparts. *PSiCL* 15. 55–65.

Lakoff, George. 1987. *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*. Chicago: The University of Chicago Press.

Lakoff, George & Mark Johnson. 1999. *Philosophy in the Flesh: the Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.

Lakoff, George & Mark Turner. 1989. *More Than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago: The University of Chicago Press.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar, Volume I: Theoretical Prerequisites*. Stanford: Stanford University Press.

Langacker, Ronald W. 1990. *Concept, image, and symbol: the cognitive basis of grammar*. Berlin: Mouton de Gruyter.

Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar, Volume II: Descriptive Application*. Stanford: Stanford University Press.

Langacker, Ronald W. 1999. *Grammar and conceptualization*. Berlin: Mouton de Gruyter.

Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne. Kemmer (eds.), *Usage-based Models of Language*, 1–63. Chicago: The

University of Chicago Press.

Langacker, Ronald W. 2003. Constructions in cognitive grammar. *English Linguistics* 20(1). The English Linguistic Society of Japan. 41–83.

Langacker, Ronald W. 2004. Grammar as Image: The Case of Voice. *Imagery in Language: Festschrift in Honour of Professor Ronald W. Langacker*, 63–114. Frankfurt am Main: Peter Lang.

Langacker, Ronald W. 2008. *Cognitive grammar: A Basic Introduction*. Oxford, UK: Oxford University Press.

Lê, Sébastien, Julie Josse & François Husson. 2008. FactoMineR: A Package for Multivariate Analysis. *Journal of Statistical Software* 25(1). 1–18.

Leaper, Campbell. 2014. Gender Similarities and Differences in Language. In Thomas M. Holtgraves (ed.), *The Oxford Handbook of Language and Social Psychology*. Oxford, UK: Oxford University Press.

Lewandowska-Tomaszczyk, Barbara, Mirosław Bańko, Rafał Górski, Marek Łaziński, Piotr Pęzik & Adam Przepiórkowski. 2012. Narodowy Korpus Języka Polskiego: geneza i dzień dzisiejszy. In Aadam Przepiórkowski, Mirosław Bańko, R L Górski & Barbara Lewandowska-Tomaszczyk (eds.), *Narodowy Korpus Języka Polskiego*, 3–10. Warszawa: Wydawnictwo Naukowe PWN.

Lewandowski, Wojciech. 2014. The locative alternation in verb-framed vs. satellite-framed languages: A corpus study of Spanish and Polish. *Studies in Language* 38(1). 864–895.

Lewandowski, Wojciech. 2016. Verbal prefixation, construction grammar, and semantic compatibility: Evidence from the locative alternation in Polish. *Folia Linguistica* 50(1). 175–206.

Mos, Maria, Antal van den Bosch & Peter Berck. 2012. The predictive value of word-level perplexity in human sentence processing: A case study on fixed adjective-preposition constructions in Dutch. In S Gries & D Divjak (eds.), *Frequency effects in language learning and processing*, 207–239. Berlin: Mouton de Gruyter.

Murphy, Gregory L. 2004. *The big book of concepts*. Cambridge, Massachusetts: MIT Press.

Nagórko, Alicja. 2007. *Zarys gramatyki polskiej*. Warszawa: Wydawnictwo Naukowe PWN.

Nenadić, Oleg & Michael Greenacre. 2007. Correspondence Analysis in R, with two-and three-dimensional graphics: The ca package. *Journal of Statistical Software* 20(3). 1–13.

Newman, Nic & David A. L. Levy. 2013. *The Reuters Institute Digital Report 2013:*

Tracking the Future of News. Oxford, UK.

Niedzielski, Henry. 1976. Polish and English Pseudo-Reflexives. *PSiCL* 4. 167–198.

Nosofsky, Robert M. 1986. Attention, Similarity, and the Identification-Categorization Relationship. *Journal of Experimental Psychology* 115(1). 39–57.

Nosofsky, Robert M. 1987. Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Nosofsky, Robert M. 1992. Exemplars, prototypes, and similarity rules. In Alice F. Healy, Stephen M. Kosslyn & Richard M. Shiffrin (eds.), *From Learning Theory to Connectionist Theory: Essays in Honor of William K. Estes, Vol. 1*, 149–167.

Ouhalla, Jamal. 1994. *Introducing transformational grammar: from rules to principles and parameters*. London: Hodder Education Publishers.

Perek, Florent. 2015. *Argument Structure in Usage-Based Construction Grammar*. . Vol. 17. (Constructional Approaches to Language). Amsterdam, Philadelphia: John Benjamins.

Pęzik, Piotr. 2012. Język mówiony w NKJP. In Adam Przepiórkowski, Mirosław Bańko, Rafał Górski & Barbara Lewandowska-Tomaszczyk (eds.), *Narodowy Korpus Języka Polskiego*. Warszawa: Wydawnictwo Naukowe PWN.

Priess, Heather A. & Janet Shibley Hyde. 2010. Gender and Academic Abilities and Preferences. *Handbook of Gender Research in Psychology*, 297–316. New York, NY: Springer New York.

R Core Team. 2017. R: A Language and Environment for Statistical Computing. Vienna, Austria.

Rosch, Eleanor. 1975. Cognitive representations of semantic categories. *Journal of experimental psychology: General* 104(3). 192–233.

Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch & Barbara Bloom Lloyd (eds.), *Cognition and categorization*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rousseeuw, Peter J. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* 20(C). 53–65.

Sadowska, Iwona. 2012. *Polish. A Comprehensive Grammar*. New York and London: Routledge.

Sandra, Dominiek & Sally Rice. 1995. Network analyses of prepositional meaning: Mirroring whose mind – the linguist's or the language user's. *Cognitive Linguisticst* 6(1). 89–130.

Sardá-Espinosa, Alexis, Subanatarajan Subbiah & Thomas Bartz-Beielstein. 2017. Conditional inference trees for knowledge extraction from motor health condition data. *Engineering Applications of Artificial Intelligence* 62. 26–37.

Saussure, Ferdinand de. 1959. *Course in General Linguistics*. New York: The Philosophical Library.

Sawicki, Lea. 1988. *Verb-valency in Contemporary Polish: A Study of the Major Valency Types*. Tübingen: Gunter Narr Verlag.

Shibatani, Masayoshi. 1985. Passives and Related Constructions: A Prototype Analysis. *Language* 61(4). 821–848.

Shimodaira, Hidetoshi. 2004. Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. *Annals of Statistics* 32(6). 2616–2641.

Shimodaira, Hidetoshi & Ryota Suzuki. 2017. pvclust. http://stat.sys.i.kyoto-u.ac.jp/prog/pvclust/ (13 September, 2017).

Siyanova-Chanturia, Anna. 2014. On the 'holistic' nature of formulaic language. *Corpus Linguistics and Linguistic Theory* 11(2). 285–301.

Slobin, Dan. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In Sven Strömqvist & Ludo Th. Verhoeven (eds.), *Relating Events in Narrative, Volume 2: Typological and Contextual Perspectives*. New York and London: Taylor & Francis.

Słoń, Anna. 2007. The 'impersonal' impersonal construction in Polish. A Cognitive Grammar analysis. In Dagmar Divjak & Agata Kochańska (eds.), *Cognitive Paths into Slavic Domain*, 257–287. Berlin: Mouton de Gruyter.

Śmiech, Witold. 1986. *Derywacja prefiksalna czasowników polskich*. Wrocław: Zakład Narodowy im. Ossolińskich.

Sokolova, Svetlana & Wojciech Lewandowski. 2010. Constructional Profile of the Verbal Prefix 'za-': a Comparative Study of Russian and Polish. *Oslo Studies in Language* 2(2). 365–391.

Spector, Phil. 2011. Stanford Stat 133 Class Notes. Stanford: Stanford University. www.stat.berkeley.edu/classes/s133/all2011.pdf.

Speelman, Dirk. 2014. Logistic regression: A confirmatory technique for comparisons in corpus linguistics. In Dylan Glynn & Justyna Robinson (eds.), *Corpus Methods for Semantics: Quantitative studies in polysemy and synonymy*, 487–533. Amsterdam, Philadelphia: John Benjamins.

Stefanowitsch, Anatol & Stefan Thomas Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.

Stefanowitsch, Anatol & Stefan Thomas Gries. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1(1). 1–43.

Striekałowa, Zofia. 1962. Budowa słowotwórcza czasowników ruchu we współczesnej polszczyźnie. *Poradnik Językowy* 5–6. 205–235.

Strobl, Carolin, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin & Achim Zeileis. 2008. Conditional variable importance for random forests. *BMC Bioinformatics* 9(1). 307–317.

Struyf, Anja, Mia Hubert & Peter J. Rousseeuw. 1997. Integrating robust clustering techniques in S-PLUS. *Computational Statistics & Data Analysis* 26(1). 17–37.

Suzuki, Ryota & Hidetoshi Shimodaira. 2006. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 22(12). 1540–1542.

Swan, Oscar E. 2002. *A Grammar of Contemporary Polish*. Bloomington, Indiana: Slavica Publishers.

Szałkiewicz, Łukasz & Adam Przepiórkowski. 2012. Anotacja morfoskładniowa. In Adam Przepiórkowski, Mirosław Bańko, Rafał Górski & Barbara Lewandowska-Tomaszczyk (eds.), *Narodowy Korpus Języka Polskiego*. Warszawa: Wydawnictwo Naukowe PWN.

Tabakowska, Elżbieta. 2003a. Those notorious Polish reflexive pronouns: a plea for Middle Voice. *Glossos*(4). 1–18.

Tabakowska, Elżbieta. 2003b. Space and time in Polish: The preposition 'za' and the verbal prefix 'za-'. In Hubert Cuyckens, Thomas Berg, Rene Dirven & Klaus-Uwe Panther (eds.), *Motivation in Language: Studies in honor of Günter Radden*, 153–177. Amsterdam, Philadelphia: John Benjamins.

Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (ed.), *Language typology and lexical description, Volume 3: Grammatical categories and the lexicon*, 36–149. Cambridge, UK: Cambridge University Press.

Talmy, Leonard. 1991. Path to Realization: A Typology of Event Conflation. *Annual Meeting of the Berkeley Linguistics Society* 17(1). 480–519.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics: Volume II.* Cambridge, Massachusetts: MIT Press.

Tomasello, Michael. 2003. *Constructing a Language*. Cambridge, Massachusetts: Harvard University Press.

Tomasello, Michael. 2009. The usage-based theory of language acquisition. In Edith L. Bavin (ed.), *Cambridge Handbook of Child Language*, 69–88. Cambridge, UK: Cambridge University Press.

Tummers, José, Kris Heylen & Dirk Geeraerts. 2005. Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1(2). 225–261.

Vanpaemel, Wolf & Gert Storms. 2008. In search of abstraction: The varying abstraction model of categorization. *Psychonomic Bulletin and Review* 15(4). 732–749.

Vanpaemel, Wolf & Gert Storms. 2010. Abstraction and model evaluation in category learning. *Behavior research methods* 42(2). 421–437.

Verbeemen, Timothy, Wolf Vanpaemel, Sven Pattyn, Gert Storms & Tom Verguts. 2007. Beyond exemplars and prototypes as memory representations of natural concepts: A clustering approach. *Journal of Memory and Language* 56. 537–554.

Wilczewska, Krystyna. 1966. Czasowniki zwrotne we współczesnej polszczyźnie. Toruń: Wydawnictwa Towarzystwa Naukowego w Toruniu.

Wonnacott, Elizabeth, Elissa L. Newport & Michael K. Tanenhaus. 2008. Acquiring and processing verb argument structure: distributional learning in a miniature language. *Cognitive Psychology* 56(3). 165–209.

Wray, Alison. 2002. *Formulaic language and the lexicon*. Cambridge, UK: Cambridge University Press.

Wray, Alison & Michael R. Perkins. 2000. The functions of formulaic language: an integrated model. *Language & Communication* 20(1). 1–28.

Yim, Odilia & Kylee T. Ramdeen. 2015. Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data. *The Quantitative Methods for Psychology* 11(1). 8–21.

Zipf, George. 1935. *The Psychobiology of Language: An Introduction to Dynamic Philology*. Cambridge, Massachusetts: MIT Press.

APPENDIX 1: Task instructions for experiments

A: Sorting experiments

Proszę podzielić poniższe zdania na trzy/cztery/pięć grup – w każdej grupie należy umieścić zdania, które wydają się Państwu podobne do siebie. Zdania z lewej strony proszę przeciągnąć do jednej z grup po prawej stronie.

'Please sort the sentences below into three/four/five groups. In each group, please put sentences that you find similar. Drag the sentences from the left into the bins on the right'

NOTE: The number of groups for sorting depended on the given experiment.

B. Nonce-verb experiment

Proszę przeczytać definicję słowa oraz podany kontekst. Następnie, proszę wybrać spośród trzech możliwości to znaczenie ***wytłuszczonego*** *zdania, które wydaje się najbardziej prawdopodobne. Czynności te należy powtórzyć dla wszystkich zdań.*

'Please read the word's definition and the context given. Subsequently, please choose the suggested meaning of the sentence in **bold** that you find most fitting in this context. Do the same for all sentences'

APPENDIX 2: Stimuli for sorting experiments on reflexives

A. Three sentences per meaning

STIMULUS	TRANSLATION
Nie perfumuj się przed kolacją.	Don't use too much perfume before dinner.
Ten komputer szybko się zepsuje.	This computer will break really soon.
Rodzice kłócą się bardzo rzadko.	My parents don't argue very often.
Świątynia budowała się bardzo długo.	It took a very long time to build the temple.
Dużo się o tym mówi w mediach.	There's a lot of discussion about this in the media.
Z zimna otuliła się płaszczem.	She was cold, so she wrapped herself with her coat.
Skończyła nam się mąka.	We've run out of flour.
Dwaj koneserzy licytują się o słynny obraz.	Two art collectors are bidding against each other for a famous picture.
Jacek wychował się w dobrym domu.	Jacek was raised in a good home.
„Gżegżółka" pisze się przez samo 'ż'.	One spells "gżegżółka" with 'ż'.
On się chyba zabije.	I think he's going to kill himself.
Na rynku utworzyło się zbiegowisko.	A crowd gathered at the marketplace.
Politycy przerzucali się obelgami.	The politicians kept insulting each other.
Dziadek leczy się tylko u znachorów.	My grandad only gets treatment from alternative medicine practitioners.
Aktualne bezrobocie szacuje się na 8%.	The current unemployment rate is estimated at 8 per cent.

B. Five sentences per meaning

STIMULUS	TRANSLATION
Starannie uczesała się do pracy.	She carefully did her hair for work.
Drzwi nagle się otworzyły.	The door opened suddenly.
Chłopcy pobili się o koleżankę.	The boys had a fight over their female friend.
Ta książka dobrze się czyta.	This book reads well.
Jutro się coś wymyśli.	We'll think something out tomorrow.
Z satysfakcją pogładzili się po brzuchach.	They rubbed their tummies with pleasure.
Stan pacjenta się pogarsza.	The patient's condition is getting worse.
Przytulili się mocno.	They gave each other a strong hug.
Nasze produkty kiepsko się sprzedają.	Our products don't sell well.
Kiedyś jadło się u nas obiad o 16.	We used to have dinner at 4 pm.
Nie perfumuj się przed kolacją.	Don't use too much perfume before dinner.
Ten komputer szybko się zepsuje.	This computer will break really soon.
Rodzice kłócą się bardzo rzadko.	My parents don't argue very often.
Świątynia budowała się bardzo długo.	It took a very long time to build the temple.
Dużo się o tym mówi w mediach.	There's a lot of discussion about this in the media.
Z zimna otuliła się płaszczem.	She was cold, so she wrapped herself with her coat.
Skończyła nam się mąka.	We've run out of flour.
Dwaj koneserzy licytują się o słynny obraz.	Two art collectors are bidding against each other for a famous picture.
Jacek wychował się w dobrym domu.	Jacek was raised in a good home.
„Gżegżółka" pisze się przez samo 'ż'.	One spells "gżegżółka" with 'ż'.
On się chyba zabije.	I think he's going to kill himself.
Na rynku utworzyło się zbiegowisko.	A crowd gathered at the marketplace.
Politycy przerzucali się obelgami.	The politicians kept insulting each other.
Dziadek leczy się tylko u znachorów.	My grandad only gets treatment from alternative medicine practitioners.
Aktualne bezrobocie szacuje się na 8%.	The current unemployment rate is estimated at 8 per cent.

APPENDIX 3: Stimuli for prefix sorting experiments

A: the prefix *po-* (three sentences per meaning)

Three-sentences-per-meaning version	TRANSLATION
Po pracy pojechała do kina.	She went to the cinema after work.
Poplotkowały o sąsiadkach.	They gossiped about their neighbours.
Dach pokrył się śniegiem.	The roof got covered with snow.
Dzieci pomazały ścianę flamastrami.	The kids covered the wall with doodles.
Skacząc do wody, pochlapała wszystkich wokół basenu.	Jumping into the pool, she splashed water on everyone around.
Pomedytowała, by się uspokoić.	She meditated a bit to calm down.
Zwierzęta pochowały się ze strachu przed myśliwym.	The animals scattered away and hid from the hunter.
Koń pogalopował do lasu.	The horse galloped into the forest.
Przed zimą pootulała wszystkie drzewa.	She wrapped all her trees before winter.
Potańczył chwilę do swojej ulubionej piosenki.	He danced a bit to his favourite song.
Z bolącym kolanem pokuśtykał do szkoły.	He hobbled to school, his knee hurting.
Poodklejali ze ścian plakaty wyborcze.	They took the political campaign posters off walls.

B: the prefix *po-* (five sentences per meaning)

Five-sentences-per-meaning version	TRANSLATION
Po pracy pojechała do kina.	She went to the cinema after work.
Poplotkowały o sąsiadkach.	They gossiped about their neighbours.
Dzieci pobawiły się klockami.	The kids played with blocks.
Dach pokrył się śniegiem.	The roof got covered with snow.
Pomalowała wszystkie paznokcie na śliwkowo.	She painted all her nails plum-purple.
Dzieci pomazały ścianę flamastrami.	The kids covered the wall with doodles.
Skacząc do wody pochlapała wszystkich wokół basenu.	Jumping into the pool, she splashed water on everyone around.
Pomedytowała, by się uspokoić.	She meditated a bit to calm down.
Zwierzęta pochowały się ze strachu przed myśliwym.	The animals scattered away and hid from the hunter.
Pootwierali wszystkie okna.	They opened all the windows.
Koń pogalopował do lasu.	The horse galloped into the forest.
Po porażce powoli podreptali zrezygnowani do domu.	After the defeat, they dawled home dispiritedly.
Przed zimą pootulała wszystkie drzewa.	She wrapped all her trees before winter.
Potańczył chwilę do swojej ulubionej piosenki.	He danced a bit to his favourite song.
Z bolącym kolanem pokuśtykał do szkoły.	He hobbled to school, his knee hurting.
Poodklejali ze ścian plakaty wyborcze.	They took the political campaign posters off walls.
Podzielili tort na 16 kawałków.	They cut the cake into 16 pieces.
Mama posmarowała chleb dżemem.	Mum spread jam over the bread.
Dziewczyny popluskały się w rzece.	The girls had some fun splashing water around in the river.
Policjant pognał za złodziejem.	The policeman chased after the thief.

C: the prefix *przy-* (three sentences per meaning)

Three-sentences-per-meaning version	TRANSLATION
Przyłączyli się do zabawy.	They joined other kids playing.
Ptak przyfrunął do karmnika.	The bird flew to the feeder.
Przyklepał zmierzwione włosy.	He patted his hair straight.
Przykleił plasterek na ranę.	He put a plaster on the cut.
Rodzice przysłali dzieciom pieniądze.	The parents sent their children some money.
Monter przykręcił półkę do ściany.	The builder fastened the shelf to the wall with screws.
Złamane drzewo przywaliło dom.	A tree broke and fell onto a house.
Przyhamowała lekko przed skrzyżowaniem.	She slowed down a little bit before the crossroads.
Zmęczeni przykucnęli pod drzewem.	They sat down under the tree to rest for a while.
Przydeptała niedopałek papierosa.	She put out the cigarette butt with her foot.
Przypudrowała sińce pod oczami.	She put some rouge over her swollen eyes.
Przyniósł gościom kawę.	He brought his guests some coffee.

D: the prefix *przy*- (five sentences per meaning)

STIMULUS	TRANSLATION
Przypłynęli do brzegu.	They sailed up to the shore.
Przyłączyli się do zabawy.	They joined other kids playing.
Ptak przyfrunął do karmnika.	The bird flew to the feeder.
Przyklepał zmierzwione włosy.	He patted his hair straight.
Grabarze przysypali trumnę piaskiem.	The cemetery workers threw sand over the coffin.
Przykleił plasterek na ranę.	He put a plaster on the cut.
Rodzice przysłali dzieciom pieniądze.	The parents sent their children some money.
Monter przykręcił półkę do ściany.	The builder fastened the shelf to the wall with screws.
Złamane drzewo przywaliło dom.	A tree broke and fell onto a house.
Marek przybiegł do domu.	Marek run over to his house.
Przyhamowała lekko przed skrzyżowaniem.	She slowed down a little bit before the crossroads.
Zmęczeni przykucnęli pod drzewem.	They sat down under the tree to rest for a while.
Przydeptała niedopałek papierosa.	She put out the cigarette butt with her foot.
Przypudrowała sińce pod oczami.	She put some rouge over her swollen eyes.
Przyniósł gościom kawę.	He brought his guests some coffee.
Hydraulik przypasował uszczelkę do rury.	The plumber fit a gasket to the pipe.
Przyciemniła lekko włosy.	She dyed her hair a bit darker.
Sławek przymierzył nowe spodnie.	Sławek tried on new trousers.
Ogrodnik przystrzygł trawnik.	The gardener trimmed the lawn.
Mama przykryła dzieci kocem..	Mum covered the kids with a blanket.

E: the prefix *roz-* (three sentences per meaning)

STIMULUS	TRANSLATION
Matematycy rozszyfrowali skomplikowany kod.	The mathematicians cracked a difficult code
Samochód rozpędził się do setki.	The car accelerated to 60 mph.
Przez nieuwagę rozgniótł stopą ślimaka.	He accidentally squashed a snail with his foot.
Stół rozleciał się ze starości.	The table fell apart because it was old.
Po kolacji rozeszli się do swoich pokojów.	After dinner, they all went to their rooms
Goście weselni rozśpiewali się na dobre.	The wedding guests started singing very loud.
Po powrocie do domu rozpakowali walizki.	After they got back home, they unpacked their bags.
Nie mogła rozplątać swoich sznurówek	She couldn't disentangle her shoelaces.
Pierwszy raz się tak rozchorował.	It was the first time he got ill like this.

F: the prefix *roz-* (five sentences per meaning)

STIMULUS	TRANSLATION
Tłum rozbiegł się na wszystkie strony.	The crowd scattered around.
Po kolacji wszyscy rozeszli się do swoich pokojów.	After dinner, they all went to their rooms.
Stół rozleciał się ze starości.	The table fell apart because it was old.
Przez nieuwagę rozgniótł stopą ślimaka.	He accidentally squashed a snail with his foot
Mama rozsmarowała starannie masło na kanapce.	Mum carefully spread butter on bread.
Ola i Jarek rozesłali zaproszenia ślubne do wszystkich gości.	Ola and Jarek sent out wedding invitations to all their guests.
Zazwyczaj cichy Marek strasznie się rozgadał.	Marek, who usually stays quiet, started talking like a chatterbox.
W kilka sekund samochód rozpędził się do setki.	Within a few seconds, the car accelerated to 60 mph.
Wieczorem strasznie się rozpadało.	It rained very heavily in the evening.
Sztuka rozbawiła widzów do łez.	The theatrical play made the audience cry with laughter.
Pierwszy raz w życiu tak się rozchorował.	It was the first time he got ill like this.
Po wejściu do domu od razu rozpakował walizki.	After he got back home, he unpacked his bags.
W styczniu rozebrali nasz spalony dom.	They took apart our burnt down house in January.
Po wielu próbach rozszyfrowali skomplikowany kod.	After many attempts, they managed to crack a difficult code.
Nie mogła rozplątać swoich sznurówek.	She couldn't disentangle her shoelaces.
Rozpieczętowała list od razu po wyjęciu go ze skrzynki.	She opened the letter immediately after she took it out of the mailbox.
Przy swojej ulubionej piosence rozśpiewali się na dobre.	They started singing aloud when they heard their favourite song.

APPENDIX 4: Stimuli for prefix nonce-verb experiment

A. Version 1

Q1 *Konstrulianie mają bardzo uporządkowane życie. W wieku od 18 do 25 lat zawsze mieszkają w bloku – to ich sposób na poznawanie innych Konstrulian. Mieszkać w bloku to* ***blorować****.*

Co znaczy poniższe zdanie?

Alojzy poblorował i przeniósł się do domu jednorodzinnego.

- ○ Alojzy blorował chwilę i przeniósł się do domu jednorodzinnego.
- ○ Alojzy blorował w różnych miejscach, aż przeniósł się do domu jednorodzinnego.
- ○ Alojzy blorował dłużej niż zwykle i dopiero potem przeniósł się do domu jednorodzinnego.

Q2 *Konstrulianie są bardzo dobrze wychowani. Gdy czkają, starają się robić to całkowicie bezgłośnie. Czknąć bezgłośnie to* ***buknąć****.*

Co znaczy poniższe zdanie?

Pobukał i mu przeszło.

- ○ Chwilę bukał i mu przeszło.
- ○ Bukał co chwila przez jakiś czas i mu przeszło.
- ○ Tak mocno bukał, że mu przeszło.

Q3 *Konstrulia to przepięknie zielona planeta – wszystko rośnie tam jak na drożdżach. Niestety oznacza to, że wszystkie pola i grządki szybko zapełniają się chwastami. Konstrulianie wyrywają je bez wytchnienia. Wyrywać chwasty to* ***chwaczyć****.*

Co znaczy poniższe zdanie?

Pochwaczyli działkę, a potem poszli na piwo.

- ○ Skończyli chwaczyć kilka grządek i poszli na piwo.
- ○ Trochę chwaczyli, ale nie skończyli i poszli na piwo.
- ○ Chwaczyli działkę bardzo mocno aż skończyli i wtedy poszli na piwo.

Q4 *Konstrulianie to urodzeni filozofowie. Każdy regularnie oddaje się rozmyślaniom o sensie życia. Zastanawiać się nad sensem życia to **donić**.*

Co znaczy poniższe zdanie?

Podonił przed snem.

- ◯ Donił chwilkę przed snem.
- ◯ Donił przed snem, robiąc na przemian inne rzeczy.
- ◯ Tak długo donił przed snem, że aż się zmęczył i zasnął.

Q5 Zjeżdanie po poręczy schodów to sport chętnie uprawiany przez młodych Konstrulian. Tak jak w ziemskich dyscyplinach typu skoki narciarskie, czy łyżwiarstwo, oceniany jest styl zjazdu po poręczy. Jedną z najwyżej punktowanych figur jest zjeżdżanie z dyndającymi nogami. Zjeżdżać po poręczy z dyndającymi nogami to ***findać***.

Co będzie oznaczać poniższe zdanie?

Pofindali we trójkę.

- ◯ Findali trochę razem dla przyjemności.
- ◯ Findali we trójkę na różnych poręczach.
- ◯ Bardzo szybko razem findali.

Q6 *Na Konstrulii często pada deszcz i w związku z tym przez większą część roku na ulicach jest pełno błota. Mieszkania brudzą się wtedy niemiłosiernie. Brudzić w mieszkaniu błotem to **grudzić**.*

Co znaczy poniższe zdanie?

Przyszła do domu w brudnych butach i pogrudziła.

- ◯ Chodziła po domu i grudziła w różnych miejscach.
- ◯ Grudziła tak bardzo, że cały dom jest teraz brudny.
- ◯ Zanim zdjęła buty, to grudziła i przez to dom jest troszkę brudny.

Q7 *Jak każde inne stworzenie w całym wszechświecie, Konstrulianie lubią spędzać czas na robieniu fajnych rzeczy. Robić coś fajnego to **hączyć**.*

Co znaczy poniższe zdanie?

Pohączyli i byli bardzo zadowoleni.

- ○ Hączyli przez chwilę i byli bardzo zadowoleni.
- ○ Co i raz hączyli, dzięki czemu byli bardzo zadowoleni.
- ○ Hączyli długo. Tak długo, aż poczuli się zadowoleni.

Q8 *Konstrulianie nie znoszą lenistwa, a tym bardziej lenistwa w dzień powszedni, gdy wszyscy inni idą do pracy. Lenić się w dzień powszedni to* ***bubać****.*

Co znaczy poniższe zdanie?

Rozbubał się.

- ○ Wcześniej tak nie było, ale ostatnio zaczął bardzo bubać.
- ○ Bubał, ale wziął się za siebie i przestał bubać.
- ○ Tak bubał, że aż rozpadł się na kawałki.

Q9 *Obrażanie Króla Konstrulii to ciężka zbrodnia karana więzieniem. Obrażać Króla to* ***chelcolić****.*

Co znaczy poniższe zdanie?

Naród się rozchelcolił.

- ○ Naród od jakiegoś czasu chelcolił coraz bardziej, tak że teraz już chelcoli na całego.
- ○ Naród tak chelcolił, że nic z niego nie zostało.
- ○ Naród zaczął troszkę chelcolić.

Q10 *Dorsz to jedyny gatunek ryby na Konstrulii. Konstrulianie przywiązują ogromną wagę do jego jakości i potrafią spędzać godziny na wybieraniu najlepszych okazów. Czasem robi się z tego rodzinna wycieczka na bazar. Kupować dorsza to* ***droszyć****.*

Co oznacza poniższe zdanie?

Rozdroszyli wszystko w sklepie.

- ○ Wiele osób droszyło, aż nie został żaden dorsz.
- ○ Wszystko w sklepie zamieniło się w dorsze.
- ○ Przywieźli dorsze i sprzedali je wszystkie w sklepie.

Q11 *W przeciwieństwie do Ziemian, Konstrulianie to zapaleni podróżnicy międzyplanetarni. Spędzają mnóstwo czasu na wyobrażaniu sobie życia na innych planetach. Fantazjować o życiu na innych planetach to* ***fażować****.*

Co oznacza poniższe zdanie?

Rozfażował się.

- ○ Fażował co raz to intensywniej aż całkowicie się w tym zatracił.
- ○ Tak fażował, że aż zniknął.
- ○ Zaczął fażować.

Q12 *Poziom przestępczości na Konstrulii nie jest wysoki, bo Konstrulianie bardzo dbają o swoje bezpieczeństwo. Drzwi, na przykład, zawsze zamykają na kilka zasuwek i kłódek. Zamykać drzwi to* ***gerdować****.*

Co oznacza poniższe zdanie?

Rozgerdował drzwi wejściowe.

- ○ Drzwi były zamknięte na wiele zasuwek, a on je otworzył.
- ○ Zamknął drzwi wejściowe.
- ○ Otworzył drzwi wejściowe na oścież.

Q13 *Większość Konstrulian nosi bardzo długie i niesforne grzywki, których układanie to prawdziwy ceremoniał rozciągnięty na kilkanaście minut każdego dnia. Układać grzywkę to* ***grzysać****.*

Co oznacza poniższe zdanie?

Rozgrzysał się na bok.

- ◯ Grzysał się tak, że grzywka była uczesana na bok.
- ◯ Zaczął się grzysać na bok.
- ◯ Przestał się grzysać na bok.

Q14 *Konstrulianie mają wyjątkowo delikatny układ trawienny i często cierpią na ból brzucha. Mają nawet specjalne słowo które oznacza "narzekać na ból brzucha". To słowo to **brzukać**.*

Co oznacza poniższe zdanie?

Próbował przybrzukać mamie.

- ◯ Próbował trochę brzukać mamie.
- ◯ Podwójnie brzukał mamie.
- ◯ Próbował brzukać tak mocno, jak mama.

Q15 *Konstrulianie uwielbiają ponowne przetwarzanie materiałów (recykling). KAŻDA butelka jest oddawana do skupu, a przed oddaniem każdy Konstrulianin czyści ją i suszy. Suszyć butelki to **buszyć**.*

Co oznacza poniższe zdanie?

Przybuszył butelki.

- ◯ Buszył butelki, ale nie do końca.
- ◯ Buszył butelki i zrobił to dokładnie z każdą butelką.
- ◯ Zaczął buszyć butelki.

Q16 *Cuda i magia to nieodłączna część życia Konstrulian. Każdy z nich może raz lub dwa razy w życiu czynić cuda. Czynić cuda to **czudzić**.*

Co oznacza poniższe zdanie?

Karol potrafi przyczudzić.

- ○ Karol czudzi naprawdę dobrze.
- ○ Karoli potrafi czudzić, ale tylko trochę.
- ○ Karol potrafi uczyć czudzić.

Q17 *Fasola to podstawa żywienia Konstrulian. Gotowanie fasoli to prawdziwa ceremonia, w którą angażują się całe rodziny. Gotować fasolę to **fagować**.*

Co oznacza poniższe zdanie?

Przyfagowali i zjedli.

- ○ Ugotowali wszystko, co mieli w domu i zjedli to.
- ○ Ugotowali porządną porcję i zjedli.
- ○ Ugotowali trochę fasoli i ją zjedli.

Q18 *Na Konstrulii drzewa są zawsze zielone, liście nigdy nie opadają. Za to trawa więdnie średnio raz na dwa tygodnie, schnie i opada. Trzeba wtedy ją starannie zagrabić i oddać do Centrum Przetwarzania Trawy. Grabić trawnik to **forgać**.*

Co oznacza poniższe zdanie?

Przyforgał do ogrodzenia.

- ○ Forgając dotarł do ogrodzenia.
- ○ Forgał, ale nie chciało mu się za bardzo, więc skończył przy ogrodzeniu.
- ○ Forgał, aż zagrabił całą trawę z ogródka do ogrodzenia.

Q19 *Konstrulianie bardzo cenią grzeczność, więc nawet, gdy się z kimś nie zgadzają, powinni robić to grzecznie. Grzecznie komuś przeczyć, to **grzeczyć**.*

Co oznacza poniższe zdanie?

Przygrzeczyła jej.

○ Niezbyt zdecydowanie jej grzeczyła, ponieważ była bardzo dobrze wychowana.

○ Grzeczyła jej z wielką siłą.

○ Zaczęła jej grzeczyć.

Q20 *Na Konstrulii przez kilka miesięcy wieją silne wiatry i po całej planecie roznosi się ogromny huk. Gdy wiatr wieje i bardzo mocno huczy, Konstrulianie mówią, że* ***huszczy****.*

Co oznacza poniższe zdanie?

Wiatr przyhuszczył dobrą pogodę.

○ Wiatr huszczył i razem z nim przyszła dobra pogoda.

○ Wiatr huszczył i zabrał dobrą pogodę w siną dal.

○ Wiatr huszczył tak mocno, że niby dobra pogoda wydawała się bardzo nieprzyjemna.

B. Version 2

Q1 *Konstrulianie nie znoszą lenistwa, a tym bardziej lenistwa w dzień powszedni, gdy wszyscy inni idą do pracy. Lenić się w dzień powszedni to **bubać**.*

Co znaczy poniższe zdanie?

Przybubał z rana.

○ Lekko bubał o poranku.

○ Bardzo bubał rano.

○ Zaczął bubać rano.

Q2 *Obrażanie Króla Konstrulii to ciężka zbrodnia karana więzieniem. Obrażać Króla to **chelcolić**.*

Co znaczy poniższe zdanie?

Tak przychelcolił, że go zamknęli.

○ Lekko chelcolił, ale i tak go zamknęli.

○ Chelcolił bardzo mocno, więc go zamknęli.

○ Chelcolił bardzo długo, aż go zamknęli.

Q3 *Dorsz to jedyny gatunek ryby na Konstrulii. Konstrulianie przywiązują ogromną wagę do jego jakości i potrafią spędzać godziny na wybieraniu najlepszych okazów. Czasem robi się z tego rodzinna wycieczka na bazar. Kupować dorsza to **droszyć**.*

Co znaczy poniższe zdanie?

Przydroszyły, ale nic nie kupiły.

○ Droszyły bardzo długo, ale nic nie kupiły.

○ Droszyły porządnie i dokładnie, ale nic nie kupiły.

○ Droszyły chwilkę i nic nie kupiły.

Q4 *W przeciwieństwie do Ziemian, Konstrulianie to zapaleni podróżnicy międzyplanetarni. Spędzają mnóstwo czasu na wyobrażaniu sobie życia na innych planetach. Fantazjować o życiu na innych planetach to **fażować**.*

Co będzie oznaczać poniższe zdanie?

Przyfażował do Plutona.

◯ Tak fażował, że aż w swoich fantazjach dotarł do Plutona.

◯ Fażował o podróży na Plutona.

◯ Fażował trochę, ale skończył, gdy w fantazjach dotarł do Plutona.

Q5 *Poziom przestępczości na Konstrulii nie jest wysoki, bo Konstrulianie bardzo dbają o swoje bezpieczeństwo. Drzwi, na przykład, zawsze zamykają na kilka zasuwek i kłódek. Zamykać drzwi to* ***gerdować****.*

Co będzie oznaczać poniższe zdanie?

Przygerdowała drzwi.

◯ Zamknęła drzwi, ale nie na wszystkie zasuwki.

◯ Zamknęła na zasuwki wszystkie drzwi w domu.

◯ Zamknęła drzwi tak, że przylegały jedno do drugiego.

Q6 *Większość Konstrulian nosi bardzo długie i niesforne grzywki, których układanie to prawdziwy ceremoniał rozciągnięty na kilkanaście minut każdego dnia. Układać grzywkę to* ***grzysać****.*

Co będzie oznaczać poniższe zdanie?

Przygrzysał się na lewą stronę.

◯ Zaczesał grzywkę tak, że przykryła lewą stronę czoła.

◯ Ułożył grzywkę w lewą stronę.

◯ Miał grzywkę ułożoną w prawą stronę, ale zmienił stronę na lewą.

Q7 *Konstrulianie mają wyjątkowo delikatny układ trawienny i często cierpią na ból brzucha. Mają nawet specjalne słowo które oznacza "narzekać na ból brzucha". To słowo to* ***brzukać****.*

Co będzie oznaczać poniższe zdanie?

Pobrzukał i brzuch przestał go boleć.

- ◯ Brzukał chwilę i brzuch przestał go boleć.
- ◯ Brzukał wielu osobom, a potem brzuch przestał go boleć.
- ◯ Tak mocno brzukał, że aż brzuch przestał go boleć.

Q8 *Konstrulianie uwielbiają ponowne przetwarzanie materiałów (recykling). KAŻDA butelka jest oddawana do skupu, a przed oddaniem każdy Konstrulianin czyści ją i suszy. Suszyć butelki to* ***buszyć****.*

Co znaczy poniższe zdanie?

Irena pobuszyła w kuchni.

- ◯ Irena buszyła przez jakiś czas, ale nie skończyła.
- ◯ Irena buszyła w kuchni butelki na stole, w szafkach, itd.
- ◯ Irena buszyła i skończyła buszyć wszystko, co miała.

Q9 *Cuda i magia to nieodłączna część życia Konstrulian. Każdy z nich może raz lub dwa razy w życiu czynić cuda. Czynić cuda to* ***czudzić****.*

Co znaczy poniższe zdanie?

Poczudził z nimi i poszedł do domu.

- ◯ Czudził chwilę z nimi i poszedł do domu.
- ◯ Czudził z nimi w kilku miejscach i poszedł do domu.
- ◯ Uczynił z nimi jeden cud od początku do końca i poszedł do domu.

Q10 *Fasola to podstawa żywienia Konstrulian. Gotowanie fasoli to prawdziwa ceremonia, w którą angażują się całe rodziny. Gotować fasolę to* ***fagować****.*

Co znaczy poniższe zdanie?

Pofagowali całą rodziną.

- ○ Jakiś czas fagowali, ale niekoniecznie skończyli wszystko fagować.
- ○ Fagowali i wspólnie skończyli.
- ○ Fagowali całą rodziną w kilku garnkach.

Q11 *Na Konstrulii drzewa są zawsze zielone, liście nigdy nie opadają. Za to trawa więdnie średnio raz na dwa tygodnie, schnie i opada. Trzeba wtedy ją starannie zagrabić i oddać do Centrum Przetwarzania Trawy. Grabić trawnik to* ***forgać****.*

Co znaczy poniższe zdanie?

Poforgał ogródek.

- ○ Forgał ogródek tak, by chociaż trochę uprzątnąć.
- ○ Forgał ogródek tak długo, aż skończył.
- ○ Forgał ogródek, ale nie mógł się zdecydować czy cały, czy nie i w końcu zrobił to w kilku miejscach.

Q12 *Konstrulianie bardzo cenią grzeczność, więc nawet, gdy się z kimś nie zgadzają, powinni robić to grzecznie. Grzecznie komuś przeczyć, to* ***grzeczyć****.*

Co znaczy poniższe zdanie?

Pogrzeczył mu.

- ○ Grzeczył przez chwilę.
- ○ Nieśmiale mu grzeczył.
- ○ Skończył grzeczyć.

Q13 *Na Konstrulii przez kilka miesięcy wieją silne wiatry i po całej planecie roznosi się ogromny huk. Kiedy wiatr wieje i bardzo głośno huczy, Konstrulianie mówią, że* ***huszczy****.*

Co znaczy poniższe zdanie?

Po burzy jeszcze pohuszczało.

- ○ Po burzy jeszcze chwilę huszczało.
- ○ Burza się skończyła, ale potem huszczało jeszcze bardzo długo.
- ○ Po burzy jeszcze co i raz to huszczało z przerwami.

Q14 *Konstrulianie mają bardzo uporządkowane życie. W wieku od 18 do 25 lat zawsze mieszkają w bloku – to ich sposób na poznawanie innych Konstrulian. Mieszkać w bloku to* ***blorować****.*

Co znaczy poniższe zdanie?

Rozblorowali ich.

- ○ Umieścili ich w blokach.
- ○ Wyrzucili ich z bloków
- ○ Umieścili ich w tymczasowych blokach.

Q15 *Konstrulianie są bardzo dobrze wychowani. Gdy czkają, starają się robić to całkowicie bezgłośnie. Czknąć bezgłośnie to* ***buknąć****.*

Co będzie oznaczać poniższe zdanie?

Rozbukał cały obiad.

- ○ Bukał, bukał aż mu obiad z brzucha wyparował.
- ○ Przez to, że bukał, zepsuł gościom cały obiad.
- ○ Bukał w trakcie całego obiadu

Q16 *Konstrulia to przepięknie zielona planeta – wszystko rośnie tam jak na drożdżach. Niestety oznacza to, że wszystkie pola i grządki szybko zapełniają się chwastami. Konstrulianie wyrywają je bez wytchnienia. Wyrywać chwasty to* ***chwaczyć****.*

Co znaczy poniższe zdanie?

Rozchwaczyła ogród.

- ○ Wyrwała chwasty w całym ogrodzie.
- ○ Wyrywała chwasty tak mocno, że poryła cały ogród.
- ○ Zaczęła chwaczyć w ogrodzie.

Q17 *Konstrulianie to urodzeni filozofowie. Każdy regularnie oddaje się rozmyślaniom o sensie życia. Zastanawiać się nad sensem życia to **donić**.*

Co znaczy poniższe zdanie?

Rozdoniła całą filozofię.

- ○ Doniła, aż zrozumiała całą filozofię.
- ○ Doniła i wymyśliła lepsze rzeczy niż cała dotychczasowa filozofia.
- ○ Przekonała całą filozofię, by też doniła.

Q18 *Zjeżdanie po poręczy schodów to sport chętnie uprawiany przez młodych Konstrulian. Tak jak w ziemskich dyscyplinach typu skoki narciarskie, czy łyżwiarstwo, oceniany jest styl zjazdu po poręczy. Jedną z najwyżej punktowanych figur jest zjeżdżanie z dyndającymi nogami. Zjeżdżać po poręczy z dyndającymi nogami to **findać**.*

Co znaczy poniższe zdanie?

Rozfindała się.

- ○ Findała coraz więcej, aż to bardzo polubiła.
- ○ Findała coraz to szybciej i szybciej, tak że już szybciej nie mogła.
- ○ Skończyła findać.

Q19 *Na Konstrulii często pada deszcz i w związku z tym przez dużą część roku na ulicach jest pełno błota. Mieszkania brudzą się wtedy niemiłosiernie. Brudzić w mieszkaniu błotem to **grudzić**.*

Co oznacza poniższe zdanie?

Rozgrudziła błoto na korytarzu.

- ○ Grudziła i rozmazała błoto po korytarzu.
- ○ Sprzątnęła błoto na korytarzu.
- ○ Zaczęła grudzić błotem na korytarzu.

Q20 *Jak każde inne stworzenie chyba w całym wszechświecie, Konstrulianie lubią spędzać czas na robieniu fajnych rzeczy. Robić coś fajnego to* ***hączyć****.*

Co oznacza poniższe zdanie?

Rozhączyły się.

- ○ Zaczęły hączyć.
- ○ Jak już raz zrobiły coś fajnego, to zaczęły to robić częściej i częściej. Teraz robią to regularnie.
- ○ Przestały hączyć.

C. Version 3

Q1 *Konstrulianie mają wyjątkowo delikatny układ trawienny i często cierpią na ból brzucha. Mają nawet specjalne słowo które oznacza "narzekać na ból brzucha". To słowo to **brzukać**.*

Co znaczy poniższe zdanie?

Rozbrzukała się.

- ○ Zaczęła strasznie brzukać ostatnio.
- ○ Brzukała, aż się wyleczyła z bólu brzucha.
- ○ Brzukała, aż miała dosyć i przestała.

Q2 *Konstrulianie uwielbiają ponowne przetwarzanie materiałów (recykling). KAŻDA butelka jest oddawana do skupu, a przed oddaniem każdy Konstrulianin czyści ją i suszy. Suszyć butelki to **buszyć**.*

Co znaczy poniższe zdanie?

Rozbuszyli butelki.

- ○ Tak buszyli butelki, że aż się rozpadły.
- ○ Buszyli butelki i pozanosili je do skupów.
- ○ Buszyli butelki tak długo, aż te same zaczęły się buszyć.

Q3 *Konstrulianie mają bardzo uporządkowane życie. W wieku od 18 do 25 lat zawsze mieszkają w bloku – to ich sposób na poznawanie innych Konstrulian. Mieszkać w bloku to **blorować**.*

Co znaczy poniższe zdanie?

Przyblorował u kolegi.

- ○ Przyszedł blorować do kolegi.
- ○ Blorował u kolegi tylko przez chwilę.
- ○ Blorował u kolegi od czasu do czasu.

Q4 *Cuda i magia to nieodłączna część życia Konstrulian. Każdy z nich może raz lub dwa razy w życiu czynić cuda. Czynić cuda to **czudzić**.*

Co znaczy poniższe zdanie?

Rozczudził nieuleczalną chorobę.

- ○ Uleczył chorobę poprzez czudzenie.
- ○ Czudził, ale przez to choroba stała się jeszcze poważniejsza.
- ○ Czudził, aż "wyczarował" nieuleczalną chorobę.

Q5 *Konstrulianie są bardzo dobrze wychowani, gdy czkają to starają się robić to całkowicie bezgłośnie. Czknąć bezgłośnie to **buknąć**.*

Co znaczy poniższe zdanie?

Przybukała z przejedzenia.

- ○ Bukała już wcześniej, ale z przejedzenia buknęła teraz jeszcze bardziej.
- ○ Lekko buknęła po wielkim jedzeniu.
- ○ Zaczęła strasznie bukać po jedzeniu.

Q6 *Fasola to podstawa żywienia Konstrulian. Gotowanie fasoli to prawdziwa ceremonia, w którą angażują się całe rodziny. Gotować fasolę to **fagować**.*

Co znaczy poniższe zdanie?

Rozfagowały garnek białej fasoli.

- ○ Tak fagowały, że z fasoli została tylko kleista maź.
- ○ Ugotowały garnek białej fasoli i dały po porcji wielu ludziom.
- ○ Zjadły cały garnek fasoli.

Q7 *Konstrulia to przepięknie zielona planeta – wszystko rośnie tam jak na drożdżach. Niestety oznacza to, że wszystkie pola i grządki szybko zapełniają się chwastami. Konstrulianie wyrywają je bez wytchnienia. Wyrywać chwasty to **chwaczyć**.*

Co znaczy poniższe zdanie?

Przychwaczył działkę.

- ◯ Trochę chwaczył na działce, ale na pewno nie skończył.
- ◯ Bardzo mocno chwaczył na działce.
- ◯ Chwaczył działkę tak, że nie został na niej ani jeden chwast.

Q8 *Na Konstrulii drzewa są zawsze zielone, liście nigdy nie opadają. Za to trawa więdnie średnio raz na dwa tygodnie, schnie i opada. Trzeba wtedy ją starannie zagrabić i oddać do Centrum Przetwarzania Trawy. Grabić trawnik to* ***forgać****.*

Co znaczy poniższe zdanie?

Rozforgali działkę.

- ◯ Zgrabili całą trawę na działce.
- ◯ Tak forgali, że aż zostawili na działce wielkie bruzdy i doły.
- ◯ Forgali tak, że na całej działce leży teraz trawa.

Q9 *Konstrulianie to urodzeni filozofowie. Każdy regularnie oddaje się rozmyślaniom o sensie życia. Zastanawiać się nad sensem życia to* ***donić****.*

Co znaczy poniższe zdanie?

Przydonił do Boga.

- ◯ Tak donił, że aż zrozumiał Boga.
- ◯ Donił aż skontaktował się z Bogiem.
- ◯ Zaczął donić z Bogiem.

Q10 *Konstrulianie bardzo cenią grzeczność, więc nawet, gdy się z kimś nie zgadzają, powinni robić to grzecznie. Grzecznie komuś przeczyć to* ***grzeczyć****.*

Co znaczy poniższe zdanie?

Rozgrzeczył jego teorię.

- ○ Grzeczył mu tak długo i skutecznie, że obalił jego teorię.
- ○ Podał jego teorię w wątpliwość.
- ○ Sprawił, że teoria zaczęła grzeczyć samemu autorowi.

Q11 *Zjeżdanie po poręczy schodów to sport chętnie uprawiany przez młodych Konstrulian. Tak jak w ziemskich dyscyplinach typu skoki narciarskie, czy łyżwiarstwo, oceniany jest styl zjazdu po poręczy. Jedną z najwyżej punktowanych figur jest zjeżdżanie z dyndającymi nogami. Zjeżdżać po poręczy z dyndającymi nogami to **findać**.*

Co znaczy poniższe zdanie?

Przyfindał do drzwi.

- ○ Findał tak mocno, że aż dotarł do drzwi.
- ○ Findał ale tylko do drzwi i przestał.
- ○ Findał do drzwi i z powrotem, i tak wiele razy.

Q12 *Na Konstrulii przez kilka miesięcy wieją silne wiatry i po całej planecie roznosi się ogromny huk. Gdy wiatr wieje i bardzo mocno huczy, Konstrulianie mówią, że **huszczy**.*

Co znaczy poniższe zdanie?

Rozhuszczyło się.

- ○ Zaczęło huszczyć i teraz huszczy na całego.
- ○ Przestało huszczyć i nastała ładna pogoda.
- ○ Huszczyło z lekka.

Q13 *Na Konstrulii często pada deszcz i w związku z tym przez dużą część roku na ulicach jest pełno błota. Mieszkania brudzą się wtedy niemiłosiernie. Brudzić w mieszkaniu błotem to **grudzić**.*

Co znaczy poniższe zdanie?

Przygrudził podłogę w przedpokoju.

○ Zabrudził lekko błotem podłogę w przedpokoju.

○ Grudził podłogę to tu, to tam i była brudna w kilku miejscach.

○ Pozostawił na podłodze grubą warstwę błota, która ją całą przykryła.

Q14 *Jak każde inne stworzenie chyba w całym wszechświecie, Konstrulianie lubią spędzać czas na robieniu fajnych rzeczy. Robić coś fajnego to* ***hączyć****.*

Co znaczy poniższe zdanie?

Miała chwilę, więc przyhączyła.

○ Miała chwilę, ale wystarczyło jej to, żeby porządnie hączyć.

○ Lekko hączyła, bo miała tylko chwilę.

○ Zaczęła hączyć, ponieważ akurat miała chwilę.

Q15 *Konstrulianie nie znoszą lenistwa, a tym bardziej lenistwa w dzień powszedni, gdy wszyscy inni idą do pracy. Lenić się w dzień powszedni to* **bubać***.*

Co znaczy poniższe zdanie?

Pobubał rano, a potem pracował.

○ Chwilę bubał rano, a potem pracował cały dzień.

○ Rano trochę bubał, trochę pracował, a potem pracował przez resztę dnia.

○ Rano mocno bubał, ale potem pracował.

Q16 *Obrażanie Króla Konstrulii to ciężka zbrodnia karana więzieniem. Obrażać Króla to* **chelcolić***.*

Co znaczy poniższe zdanie?

Pochelcolił i go zamknęli.

◯ Chelcolił przy wielu osobach i trafił do więzienia.

◯ Chelcolił przez krótki czas i trafił od więzienia.

◯ Bardzo mocno chelcolił przez długi czas i trafił do więzienia.

Q17 *Dorsz to jedyny gatunek ryby na Konstrulii. Konstrulianie przywiązują ogromną wagę do jego jakości i potrafią spędzać godziny na wybieraniu najlepszych okazów. Czasem robi się z tego rodzinna wycieczka na bazar. Kupować dorsza to* ***droszyć****.*

Co znaczy poniższe zdanie?

Podroszyli, ale nic nie kupili.

◯ Droszyli chwilę, ale nic nie kupili.

◯ Droszyli w wielu miejscach, ale nic nie kupili.

◯ Droszyli bardzo długo, ale nic nie kupili.

Q18 *W przeciwieństwie do Ziemian, Konstrulianie to zapaleni podróżnicy międzyplanetarni. Spędzają mnóstwo czasu na wyobrażaniu sobie życia na innych planetach. Fantazjować o życiu na innych planetach to* ***fażować****.*

Co znaczy poniższe zdanie?

Nie chciało mu się pracować, więc pofażował.

◯ Przez chwilę fażował zamiast pracować.

◯ Fażował tak długo, aż się fażowaniem nasycił.

◯ Fażował o wielu różnych planetach zamiast pracować.

Q19 *Poziom przestępczości na Konstrulii nie jest wysoki, bo Konstrulianie bardzo dbają o swoje bezpieczeństwo. Drzwi, na przykład, zawsze zamykają na kilka zasuwek i kłódek. Zamykać drzwi bardzo dokładnie to* ***gerdować****.*

Co znaczy poniższe zdanie?

Pogerdował drzwi w domu.

- ○ Zamknął różne drzwi w domu.
- ○ Otworzył drzwi w domu.
- ○ Zamykał drzwi, ale nie skończył.

Q20 *Większość Konstrulian nosi bardzo długie i niesforne grzywki, których układanie to prawdziwy ceremoniał rozciągnięty na kilkanaście minut każdego dnia. Układać grzywkę to* ***grzysać****.*

Co znaczy poniższe zdanie?

Fryzjer pogrzysał klientów.

- ○ Fryzjer ułożył grzywkę kilku klientom.
- ○ Fryzjer jakiś czas grzysał (a potem robił coś innego)
- ○ Fryzjer grzysał klientów na bok.

APPENDIX 5: Stimuli for prefix nonce-verb experiment (translations)

A. Version 1

Q1 *Construlians enjoy order and regularity in their lives. From 18 to 25 years of age, they always live in blocks of flats – this is how they meet other Construlians. The verb for 'live in a block of flats' is* ***blorować.***

What does the sentence below mean?

Alojzy po-blorowa-ł **(po-blorować-PST.3SG.MASC)** ***and then moved to his own house.***

○ Alojzy blorowa-ł (blorować-PST.3SG.MASC) for a little while and then moved to his own house.

○ Alojzy blorowa-ł (blorować-PST.3SG.MASC) in many places, and then moved to his own house in the end.

○ Alojzy blorowa-ł (blorować-PST.3SG.MASC) longer than usual and only then he moved to his own house.

Q2 *Construlians are really well behaved. When they hiccough, they try to do it as silent as possible. To hiccough silently is* ***buknąć.***

What does the sentence below mean?

He po-buka-ł **(po-bukać-PST.3SG.MASC)** ***and it stopped.***

○ He buka-ł (bukać-PST.3SG.MASC) for a while and it stopped.

○ He buka-ł (bukać-PST.3SG.MASC) every once in a while for some time and it stopped.

○ He buka-ł (bukać-PST.3SG.MASC) so hard that it stopped.

Q3 *Construlia is a beautifully green planet – every kind of plant just thrives there. Unfortunately, it means that all fields and vegetable patches get covered in weeds very quickly. Construlians take a lot of care to get rid of all the weeds. To get rid of weeds is* ***chwaczyć.***

What does the sentence below mean?

They po-chwaczy-li (po-chwaczyć-PST.3PL.VIR) the allotment and went to a pub.

○ They finished chwaczyć a few patches and went to a pub.

○ They chwaczy-li (chwaczyć-PST.3PL.VIR), but they didn't finish, and they went to a pub.

○ They chwaczy-li (chwaczyć-PST.3PL.VIR) so hard that they finally finished, and they went to a pub.

Q4 *Construlians are born philosophers. Each Construlian would regularly ponder the meaning of life. To ponder the meaning of life is* ***donić****.*

What does the sentence below mean?

He po-doni-ł (po-donić-PST.3SG.MASC) before he went to bed.

○ He doni-ł (donić-PST.3SG.MASC) for a while before he went to bed.

○ He doni-ł (donić-PST.3SG.MASC) before going to bed, doing other stuff in the meantime too.

○ He doni-ł (donić-PST.3SG.MASC) before going to bed for so long that he got tired and fell asleep.

Q5 *Riding down a banister is a sport that young Construlians love to do. Just like in some sports disciplines on Earth – such as ski jumping or figure skating – the style in which people ride down a banister is also judged. One of the "tricks" that can score you the highest number of points is riding down with your legs flying around. To ride down a banister with your legs flying around is* ***findać.***

What does the sentence below mean?

The three of them po-finda-li (po-findać-PST.3PL.VIR)

○ They finda-li (findać-PST.3PL.VIR) a bit together, just for fun.

○ The three of them finda-li (findać-PST.3PL.VIR) down different banisters.

○ They finda-li (findać-PST.3PL.VIR) very fast together.

Q6 *It often rains on Construlia, which makes streets covered in mud for most of the year. Construlians' homes get really dirty then. To make your home dirty with mud is* ***grudzić.***

What does the sentence below mean?

She came home in muddy boots and po-grudzi-ła (po-grudzić-PST.3SG.FEM).

◯ She would walk around the house and grudzi-ła (grudzić-PST.3SG.FEM) in many places.

◯ She grudzi-ła (grudzić-PST.3SG.FEM) so hard that the house is all dirty now.

◯ Before she took off her shoes, she grudzi-ła (grudzić-PST.3SG.FEM), and that's why the house is a bit dirty now.

Q7 *Just like any other being in the universe, Construlians like to do cool stuff. To do cool stuff is* ***hączyć.***

What does the sentence below mean?

They po-hączy-li (po-hączyć-PST.3PL.VIR), which made them very satisfied.

◯ They hączy-li (hączyć-PST.3PL.VIR) for a while, which made them very satisfied.

◯ They hączy-li (hączyć-PST.3PL.VIR) on and off again, which made them very satisfied.

◯ They hączy-li (hączyć-PST.3PL.VIR) long. So long that it made them satisfied.

Q8 *Construlians hate to slack off. There's only one thing they hate more than slacking off – it's slacking off on a working day, when everyone else is at work. To slack off on a working day is* ***bubać.***

What does the sentence below mean?

He roz-buba-ł (roz-bubać-PST.3SG.MASC) się.

◯ He didn't use to be like this, but recently, he's started to bubać a lot.

◯ He buba-ł (bubać-PST.3SG.MASC) but he got a grip on himself and stopped bubać.

◯ He buba-ł (bubać-PST.3SG.MASC) so hard that he fell apart.

Q9 *Offending the King of Construlia is a grave offence liable for a prison sentence. To offend the King is* ***chelcolić.***

What does the sentence below mean?

The nation of Construlia się roz-chelcoli-ł (roz-chelcolić-PST.3SG.MASC).

○ The nation of Construlia has been chelcolić more and more for some time now, and now they are all out chelcolić.

○ The nation of Construlia chelcoli-ł (chelcolić-PST.3SG.MASC) so hard that none of them are left.

○ The nation of Construlia started to chelcolić a little bit.

Q10 *Cod is the only species of fish that exists on Construlia. The quality of their cod is very important to Construlians, and they can spend hours on end picking the best possible fish. To buy cod is* ***droszyć.***

What does the sentence below mean?

They rozdroszyli (roz-droszyć-PST.3PL.VIR) everything in the shop.

○ Many people droszy-ło (droszyć-PST.3SG.NEUT) as long as there was no cod left.

○ Everything in the shop turned to cod.

○ They brought in cod and they sold all of them in the shop.

Q11 *In contrast to inhabitants of Earth, Construlians love interplanetary travel. They spend a lot of time dreaming of what it would be like to live on other planets. To dream of living on other planets is* ***fażować.***

What does the sentence below mean?

He roz-fażowa-ł (roz-fażować-PST.3SG.MASC) się.

○ He fażowa-ł (fażować-PST.3SG.MASC) harder and harder until he got completely engrossed in it.

○ He fażowa-ł (fażować-PST.3SG.MASC) so hard that he disappeared.

○ He began to fażować.

Q12 *The crime rates on Construlia aren't high because Construlians take utmost care of their safety. For instance, they always lock their doors with a few locks and padlocks. To lock doors is* ***gerdować.***

What does the sentence below mean?

He roz-gerdowa-ł (roz-gerdować-PST.3SG.MASC) the front door.

- ○ The door was locked with many locks, and he unlocked them.
- ○ He locked the front door.
- ○ He opened the front door wide.

Q13 *Most Construlians wear really long and frizzy fringes. Styling them is a true ritual, which takes more than ten minutes each day. To style a fringe is* ***grzysać.***

What does the sentence below mean?

He roz-grzysa-ł się (roz-grzysać-PST.3SG.MASC) to the side.

- ○ He grzysa-ł (grzysać-PST.3SG.MASC) się in such a way that his fringe turned to the side.
- ○ He started grzysać się to the side.
- ○ He stopped grzysać się to the side.

Q14 *Construlians have a particularly fragile digestive system, and they often suffer from stomach pains. They even have a special word that means "complain about a painful stomach" – this word is* ***brzukać****.*

What does the sentence below mean?

He tried to przy-brzukać to his mum.

- ○ He tried to brzukać a little bit to his mum.
- ○ He brzuka-ł (brzukać-PST.3SG.MASC) to his mum double hard.
- ○ He tried to brzukać as hard as his mum.

Q15 *Construlians love to recycle. They return EVERY SINGLE bottle to a recycling facility, washing them and drying them first. To dry bottles is to* ***buszyć****.*

What does the sentence below mean?

He przy-buszy-ł (przy-buszyć-PST.3SG.MASC) the bottles.

◯ He buszy-ł (buszyć-PST.3SG.MASC) the bottles but not fully.

◯ He buszy-ł (buszyć-PST.3SG.MASC) the bottles, and he did this thoroughly with each bottle.

◯ He started to buszyć the bottles.

Q16 *Miracles and magic are both inherent parts of Construlians' lives. Each Construlian can make a miracle happen once or twice in their lifetime. To make miracles happen is* ***czudzić****.*

What does the sentence below mean?

Karol can przy-czudzić.

◯ Karol czudz-i (czudzić-PRES.3SG) really well.

◯ Karol can czudzić but only a little bit.

◯ Karol can teach how to czudzić.

Q17 *Beans are Construlians' food staple. Cooking beans is a ceremony that gets whole families involved. To cook beans is* ***fagować.***

What does the sentence below mean?

Przy-fagowa-li (przy-fagować-PST.3PL.VIR) i zjedli.

◯ They cooked everything they had at home and ate it.

◯ They cooked a big portion and ate it.

◯ They cooked some beans and ate it.

Q18 *On Construlia, trees are always green – leaves never fall. Grass, in contrast, dries out every once two weeks. Construlians then carefully rake it and take it to Grass Recycling Centres. To rake a lawn is* ***forgać.***

What does the sentence below mean?

He przyforga-ł (przy-forgać-PST.3SG.MASC) up to the fence.

○ Forga-jąc (forgać-PRESP), he reached the fence.

○ He forga-ł (forgać-PST.3SG.MASC), but he didn't feel like it too much, so he stopped at the fence.

○ He forga-ł (forgać-PST.3SG.MASC) until he raked all the grass in the garden up to the fence.

Q19 *Construlians appreciate good manners. Even if they don't agree with someone, they should argue politely. To argue against what somebody said politely is **grzeczyć**.*

What does the sentence below mean?

Przy-grzeczy-ła (przy-grzeczyć-PST.3SG.FEM) her.

○ She grzeczy-ła (grzeczyć-PST.3SG.FEM) her not too persuasively because she was very well behaved.

○ She grzeczy-ła (grzeczyć-PST.3SG.FEM) her very forcefully.

○ She started grzeczyć her.

Q20 *On Construlia, strong winds blow all the time for a few months, and the planet is filled with roaring sounds. When wind makes a loud roar, Construlians say that it **huszcz-y** (huszczeć-PST.3SG.MASC).*

What does the sentence below mean?

The wind przy-huszczył (przy-huszczyć-3SG.PST.MASC) a good weather.

○ The wind huszczy-ł (huszczyć-3SG.PST.MASC) and along with it came good weather.

○ The wind huszczy-ł (huszczyć-3SG.PST.MASC) and took the bad weather away.

○ The wind huszczy-ł (huszczyć-3SG.PST.MASC) so hard that the weather, which would otherwise be very good, seemed very unpleasant.

B. Version 2

Q1 *Construlians hate to slack off. There's only one thing they hate more than slacking off – it's slacking off on a working day, when everyone else is at work. To slack off on a working day is* ***bubać.***

What does the sentence below mean?

He przybuba-ł (przy-bubać-PST.3SG.MASC) in the morning.

- ○ He buba-ł (bubać-PST.3SG.MASC) a bit in the morning.
- ○ He buba-ł (bubać-PST.3SG.MASC) a lot in the morning.
- ○ He began to bubać in the morning.

Q2 *Offending the King of Construlia is a grave offence liable for a prison sentence. To offend the King is* ***chelcolić.***

What does the sentence below mean?

He przychelcoli-ł (przy-chelcolić-PST.3SG.MASC) so hard that they put him in gaol.

- ○ He chelcoli-ł (chelcolić-PST.3SG.MASC) only a bit but they put him in gaol anyway.
- ○ He chelcoli-ł (chelcolić-PST.3SG.MASC) very hard, and they put him in gaol.
- ○ He chelcoli-ł (chelcolić-PST.3SG.MASC) so long that they put him in gaol.

Q3 *Cod is the only species of fish that exists on Construlia. The quality of their cod is very important to Construlians, and they can spend hours on end picking the best possible fish. To buy cod is* ***droszyć.***

What does the sentence below mean?

They przydroszy-ły (przy-droszyć-PST.3PL.NON-VIR), but they didn't buy anything.

○ They droszy-ły (droszyć-PST.3PL.NON-VIR) very long, but they didn't buy anything.

○ They droszy-ły (droszyć-PST.3PL.NON-VIR) thoroughly, but they didn't buy anything.

○ They droszy-ły (droszyć-PST.3PL.NON-VIR) for a while, and they didn't buy anything.

Q4 *In contrast to inhabitants of Earth, Construlians love interplanetary travel. They spend a lot of time dreaming of what it would be like to live on other planets. To dream of living on other planets is* ***fażować.***

What does the sentence below mean?

He przy-fażowa-ł (przy-fażować-PST.3SG.MASC) to Pluto.

○ He fażowa-ł (fażować-PST.3SG.MASC) so hard that he got to Pluto in his dreams.

○ He fażowa-ł (fażować-PST.3SG.MASC) about a trip to Pluto.

○ He fażowa-ł (fażować-PST.3SG.MASC) a bit, but he finished when he got to Pluto in his dreams.

Q5 *The crime rates on Construlia aren't high because Construlians take utmost care of their safety. For instance, they always lock their doors with a few locks and padlocks. To lock doors is* ***gerdować.***

What does the sentence below mean?

He przy-gerdowa-ła (przy-gerdować-PST.3SG.FEM) the door(s).

○ She locked the door but not with all the locks.

○ She locked all the doors in the house.

○ She locked the doors so that they touched one another.

Q6 *Most Construlians wear really long and frizzy fringes. Styling them is a true ritual, which takes more than ten minutes each day. To style a fringe is* ***grzysać.***

What does the sentence below mean?

He przy-grzysa-ł się (przy-grzysać-PST.3SG.MASC) to the left.

○ He combed his fringe so that it covered the left-hand side of his forehead.

○ He styled his fringe to the left.

○ His fringe was facing right, but he combed it to the left.

Q7 *Construlians have a particularly fragile digestive system, and they often suffer from stomach pains. They even have a special word that means "complain about a painful stomach" – this word is* ***brzukać****.*

What does the sentence below mean?

He po-brzuka-ł (po-brzukać-PST.3SG.MASC) and his stomach stopped hurting.

○ He brzuka-ł (brzukać-PST.3SG.MASC) for a while and his stomach stopped hurting.

○ He brzuka-ł (brzukać-PST.3SG.MASC) to many people, and his stomach stopped hurting.

○ He brzuka-ł (brzukać-PST.3SG.MASC) hard, until his stomach stopped hurting.

Q8 *Construlians love to recycle. They return EVERY SINGLE bottle to a recycling facility, washing them and drying them first. To dry bottles is to* ***buszyć****.*

What does the sentence below mean?

Irena po-buszył-a (po-buszyć-PST.3SG.FEM) in the kitchen.

○ Irena buszył-a (buszyć-PST.3SG.FEM) for some time, but she didn't finish.

○ In the kitchen, Irena buszył-a (buszyć-PST.3SG.FEM) bottles on the table, in the cupboards, and so on.

○ Irena buszył-a (buszyć-PST.3SG.FEM) and she finished buszyć everything she had.

Q9 *Miracles and magic are both inherent parts of Construlians' lives. Each Construlian can make a miracle happen once or twice in their lifetime. To make* *miracles happen is **czudzić**.*

What does the sentence below mean?

***He po-czudzi-ł (po-czudzić-PST.3SG.MASC)** with them and went home.*

◯ He czudzi-ł (czudzić-PST.3SG.MASC) with them for a little while and went home.

◯ He czudzi-ł (czudzić-PST.3SG.MASC) with them in a few places and went home.

◯ He made one miracle with them fully happen with them and went home.

Q10 *Beans are Construlians' food staple. Cooking beans is a ceremony that gets whole families involved. To cook beans is **fagować.***

What does the sentence below mean?

They po-fagowa-li (po-fagować-PST.3PL.VIR) with the entire family.

◯ They fagowa-li (fagować-PST.3PL.VIR) for some time but didn't quite finish fagować it all.

◯ They fagowa-li (fagować-PST.3PL.VIR) and finished together.

◯ They fagowa-li (fagować-PST.3PL.VIR) with the entire family in a few pots.

Q11 *On Construlia, trees are always green – leaves never fall. Grass, in contrast, dries out every once two weeks. Construlians then carefully rake it and take it to Grass Recycling Centres. To rake a lawn is **forgać.***

What does the sentence below mean?

He po-forga-ł (po-forgać-PST.3SG.MASC) the garden.

◯ He forga-ł (forgać-PST.3SG.MASC) the garden so that it became even a bit tidier.

◯ He forga-ł (forgać-PST.3SG.MASC) the garden so long until he finished.

◯ He forga-ł (forgać-PST.3SG.MASC) the garden, but he couldn't decide whether to forgać it all or not, and did only in a few spots in the end.

Q12 *Construlians appreciate good manners. Even if they don't agree with someone, they should argue politely. To argue against what somebody said politely is* ***grzeczyć****.*

What does the sentence below mean?

He po-grzeczy-ł (po-grzeczyć-PST.3SG.MASC) him.

- ○ He grzeczy-ł (grzeczyć-PST.3SG.MASC) him for a while.
- ○ He grzeczy-ł (grzeczyć-PST.3SG.MASC) him shyly.
- ○ He finished grzeczyć.

Q13 *On Construlia, strong winds blow all the time for a few months, and the planet is filled with roaring sounds. When wind makes a loud roar, Construlians say that it* ***huszcz-y*** *(huszczeć-PST.3SG.MASC)..*

What does the sentence below mean?

It po-huszcza-ło (po-huszczeć-PST.3SG.NEUT) after the storm for a while.

- ○ After the storm, it still huszcza-ło (huszczeć-PST.3SG.NEUT) for a while.
- ○ The storm had come to an end, but it still huszcza-ło (huszczeć-PST.3SG.NEUT) for a long time then.
- ○ After the storm, it still huszcza-ło (huszczeć-PST.3SG.NEUT) on and off again.

Q14 *Construlians enjoy order and regularity in their lives. From 18 to 25 years of age, they always live in blocks of flats – this is how they meet other Construlians. The verb for 'live in a block of flats' is* ***blorować.***

What does the sentence below mean?

They roz-blorowa-li (roz-blorować-PST.3SG.VIR) them.

- ○ They put them in blocks of flats.
- ○ They evicted them from the blocks of flats.
- ○ They put them in interim blocks of flats.

Q15 *Construlians are really well behaved. When they hiccough, they try to do it as silent as possible. To hiccough silently is* ***buknąć.***

What does the sentence below mean?

He roz-buka-ł (roz-bukać-PST.3SG.MASC) the entire dinner.

○ He buka-ł (bukać-PST.3SG.MASC) until the dinner vanished from his stomach.

○ He spoiled the entire dinner for his guests because he buka-ł (bukać-PST.3SG.MASC).

○ He buka-ł (bukać-PST.3SG.MASC) throughout the entire dinner.

Q16 *Construlia is a beautifully green planet – every kind of plant just thrives there. Unfortunately, it means that all fields and vegetable patches get covered in weeds very quickly. Construlians take a lot of care to get rid of all the weeds. To get rid of weeds is* ***chwaczyć.***

What does the sentence below mean?

She roz-chwaczy-ła (roz-chwaczyć-PST.3SG.FEM) the garden.

○ She took out the weeds in the entire garden.

○ She was ripping out the weeds so hard that she left the entire garden covered in pits and grooves.

○ She started chwaczyć in the garden.

Q17 *Construlians are born philosophers. Each Construlian would regularly ponder the meaning of life. To ponder the meaning of life is* ***donić****.*

What does the sentence below mean?

She roz-doni-ła (roz-donić-PST.3SG.FEM) the entire philosophy.

○ She doni-ła (donić-PST.3SG.FEM) until she understood the entire philosophy.

○ She doni-ła (donić-PST.3SG.FEM) and came up with things that were better than the entire philosophy to date.

○ She convinced the entire philosophy to also donić.

Q18 *Riding down a banister is a sport that young Construlians love to do. Just like in some sports disciplines on Earth – such as ski jumping or figure skating – the style in which people ride down a banister is also judged. One of the "tricks" that can score you the highest number of points is riding down with your legs flying around. To ride down a banister with your legs flying around is* ***findać.***

What does the sentence below mean?

She roz-finda-ła (roz-findać-PST.3SG.FEM) się.

- ○ She finda-ła (findać-PST.3SG.FEM) more and more until she got to like it a lot.
- ○ She finda-ła (findać-PST.3SG.FEM) faster and faster until she couldn't go any faster.
- ○ She finished findać.

Q19 *It often rains on Construlia, which makes streets covered in mud for most of the year. Construlians' homes get really dirty then. To make your home dirty with mud is* ***grudzić.***

What does the sentence below mean?

She roz-grudzi-ła (roz-grudzić-PST.3SG.FEM) mud in the hall.

- ○ She grudzi-ła (grudzić-PST.3SG.FEM) and smeared mud all over the hall.
- ○ She cleaned the mud in the hall.
- ○ She started grudzić with mud in the hall.

Q20 *Just like any other being in the universe, Construlians like to do cool stuff. To do cool stuff is* ***hączyć.***

What does the sentence below mean?

They roz-hączy-ły (roz-hączyć-PST.3PL.NON-VIR) się.

- ○ They started to hączyć.
- ○ Once they started to do something cool, they started doing it more and more often. Now they do it regularly.
- ○ They stopped hączyć.

C. Version 3

Q1 *Construlians have a particularly fragile digestive system, and they often suffer from stomach pains. They even have a special word that means "complain about a painful stomach" – this word is **brzukać**.*

What does the sentence below mean?

She roz-brzuka-ła (roz-brzukać-PST.3SG.FEM) się.

- ○ She started to brzukać a lot of late.
- ○ She brzuka-ła (brzukać-PST.3SG.FEM) until she healed herself out of her stomach pains.
- ○ She brzuka-ła (brzukać-PST.3SG.FEM) until she was fed up with it, and she stopped.

Q2 *Construlians love to recycle. They return EVERY SINGLE bottle to a recycling facility, washing them and drying them first. To dry bottles is to **buszyć**.*

What does the sentence below mean?

They roz-buszy-li (roz-buszyć-PST.3PL.VIR) bottles.

- ○ They buszy-li (buszyć-PST.3PL.VIR) bottles so hard that they fell apart.
- ○ They buszy-li (buszyć-PST.3PL.VIR) bottles and took them to recycling centres.
- ○ They buszy-li (buszyć-PST.3PL.VIR) bottles so long that they started buszyć themselves.

Q3 *Construlians enjoy order and regularity in their lives. From 18 to 25 years of age, they always live in blocks of flats – this is how they meet other Construlians. The verb for 'live in a block of flats' is **blorować.***

What does the sentence below mean?

He przy-blorowa-ł (przy-blorować-PST.3SG.MASC) at his friend's.

- ○ He went blorować at his friend's.
- ○ He blorowa-ł (blorować-PST.3SG.MASC) at his friend's only for a while.
- ○ He blorowa-ł (blorować-PST.3SG.MASC) at his friend's from time to time.

Q4 *Miracles and magic are both inherent parts of Construlians' lives. Each Construlian can make a miracle happen once or twice in their lifetime. To make miracles happen is* ***czudzić****.*

What does the sentence below mean?

He roz-czudzi-ł (roz-czudzić-PST.3SG.MASC) an incurable illness.

○ He healed the person from the illness by czudz-enie (czudzić-DVRB.NOUN).

○ He czudzi-ł (czudzić-PST.3SG.MASC), but he only made the illness worse.

○ He czudzi-ł (czudzić-PST.3SG.MASC) until he 'conjured up' an incurable illness.

Q5 *Construlians are really well behaved. When they hiccough, they try to do it as silent as possible. To hiccough silently is* ***buknąć.***

What does the sentence below mean?

She przy-bukał-a (przy-bukać-PST.3SG.FEM) because she overate.

○ She bukał-a (bukać-PST.3SG.FEM) before already, but out of overeating she buknę-ła (buknąć.PFV-PST.3SG.FEM) now even harder.

○ She buknę-ła (buknąć.PFV-PST.3SG.FEM) slightly after a big feast.

○ She started to bukać very hard after she finished her meal.

Q6 *Beans are Construlians' food staple. Cooking beans is a ceremony that gets whole families involved. To cook beans is* ***fagować.***

What does the sentence below mean?

They roz-fagowa-ły (roz-fagować-PST.3PL.NON-VIR) a pot of navy beans.

○ They fagowa-ły (fagować-PST.3PL.NON-VIR) so hard that they turned the beans into sticky goo.

○ They cooked a pot of navy beans and they gave out a helping to many people.

○ They ate a potful of beans.

Q7 *Construlia is a beautifully green planet – every kind of plant just thrives there. Unfortunately, it means that all fields and vegetable patches get covered in weeds* *very quickly. Construlians take a lot of care to get rid of all the weeds. To get rid of weeds is* ***chwaczyć.***

What does the sentence below mean?

He przy-chwaczy-ł (przy-chwaczyć-PST.3SG.MASC) his allotment.

○ He chwaczy-ł (chwaczyć-PST.3SG.MASC) on his allotment, but he most certainly didn't finish.

○ He chwaczy-ł (chwaczyć-PST.3SG.MASC) very hard on his allotment.

○ He chwaczy-ł (chwaczyć-PST.3SG.MASC) the allotment so intensely that even a single weed wasn't left.

Q8 *On Construlia, trees are always green – leaves never fall. Grass, in contrast, dries out every once two weeks. Construlians then carefully rake it and take it to Grass Recycling Centres. To rake a lawn is* ***forgać.***

What does the sentence below mean?

They roz-forga-li (roz-forgać-PST.3PL.NON-VIR) the allotment.

○ They raked all the grass on the allotment.

○ They forga-li (forgać-PST.3PL.NON-VIR) so intensely that they left big pits and grooves on the allotment.

○ They forga-li (forgać-PST.3PL.NON-VIR) so intensely that the whole allotment is covered with grass now.

Q9 *Construlians are born philosophers. Each Construlian would regularly ponder the meaning of life. To ponder the meaning of life is* ***donić****.*

What does the sentence below mean?

He przy-doni-ł (przy-donić-PST.3SG.MASC) to God.

○ He doni-ł (donić-PST.3SG.MASC) so intensely that he understood God.

○ He doni-ł (donić-PST.3SG.MASC) until he got in touch with God.

○ He started donić with God.

Q10 *Construlians appreciate good manners. Even if they don't agree with someone, they should argue politely. To argue against what somebody said politely is* ***grzeczyć****.*

What does the sentence below mean?

He roz-grzeczy-ł (roz-grzeczyć-PST.3SG.MASC) his theory.

- ○ He grzeczy-ł (grzeczyć-PST.3SG.MASC) so long and effectively that he managed to disprove his theory.
- ○ He questioned his theory.
- ○ He made the theory grzeczyć its own author.

Q11 *Riding down a banister is a sport that young Construlians love to do. Just like in some sports disciplines on Earth – such as ski jumping or figure skating – the style in which people ride down a banister is also judged. One of the "tricks" that can score you the highest number of points is riding down with your legs flying around. To ride down a banister with your legs flying around is* ***findać.***

What does the sentence below mean?

He przy-finda-ł (przy-findać-PST.3SG.MASC) to the door.

- ○ He finda-ł (findać-PST.3SG.MASC) so intensely that he got to the door.
- ○ He finda-ł (findać-PST.3SG.MASC) ale but he stopped at the door.
- ○ He finda-ł (findać-PST.3SG.MASC) to the door and back time and time again.

Q12 *On Construlia, strong winds blow all the time for a few months, and the planet is filled with roaring sounds. When wind makes a loud roar, Construlians say that it* ***huszcz-y*** *(huszczeć-PST.3SG.MASC).*

What does the sentence below mean?

It roz-huszczy-ło (roz-huszczyć-PST.3SG.NEUT) się.

- ○ It started huszczyć and now it huszcz-y (huszczyć-PRES.3SG) totally.
- ○ It stopped huszczyć and the weather turned really nice.
- ○ It huszczy-ło (huszczyć-PST.3SG.NEUT) a little bit.

Q13 *It often rains on Construlia, which makes streets covered in mud for most of the year. Construlians' homes get really dirty then. To make your home dirty with mud is* ***grudzić.***

What does the sentence below mean?

He przy-grudzi-ł (przy-grudzić-PST.3SG.MASC) the floor in the hall.

○ He soiled the floor in the hall a bit with mud.

○ He grudzi-ł (grudzić-PST.3SG.MASC) the floor here and there, which made the floor dirty in places.

○ He left on the floor a layer of mud that covered it entirely.

Q14 *Just like any other being in the universe, Construlians like to do cool stuff. To do cool stuff is* ***hączyć.***

What does the sentence below mean?

She had a few minutes to spare so she przy-hączy-ła (przy-hączyć-PST.3SG.FEM).

○ She only had a few minutes to spare but it was enough for her to hączyć.

○ She hączy-ła (haczyć-PST.3SG.FEM) but she only had a few minutes to spare.

○ She started to hączyć, because she had a few minutes to spare.

Q15 *Construlians hate to slack off. There's only one thing they hate more than slacking off – it's slacking off on a working day, when everyone else is at work. To slack off on a working day is* ***bubać.***

What does the sentence below mean?

He po-buba-ł (po-bubać-PST.3SG.MASC) in the morning, and then he started working.

○ He buba-ł (bubać-PST.3SG.MASC) for a while in the morning and then he worked all day.

○ He buba-ł (bubać-PST.3SG.MASC) and worked a bit in the morning, and then her worked for the rest of the day.

○ He buba-ł (bubać-PST.3SG.MASC) hard in the morning, but then he worked.

Q16 *Offending the King of Construlia is a grave offence liable for a prison sentence. To offend the King is* ***chelcolić.***

What does the sentence below mean?

He po-chelcoli-ł (po-chelcolić-PST.3SG.MASC) and they put him in prison.

◯ He chelcoli-ł (chelcolić-PST.3SG.MASC) in front of many people, and he ended up in prison.

◯ He chelcoli-ł (chelcolić-PST.3SG.MASC) only for a short time, and he ended up in prison.

◯ He chelcoli-ł (chelcolić-PST.3SG.MASC) hard for a long time, and he ended up in prison.

Q17 *Cod is the only species of fish that exists on Construlia. The quality of their cod is very important to Construlians, and they can spend hours on end picking the best possible fish. To buy cod is* ***droszyć.***

What does the sentence below mean?

They po-droszy-li (po-droszyć-PST.3PL.VIR), but they didn't buy anything.

◯ They droszy-li (droszyć-PST.3PL.VIR) for a while, but they didn't buy anything.

◯ They droszy-li (droszyć-PST.3PL.VIR) in many places, but they didn't buy anything.

◯ They droszy-li (droszyć-PST.3PL.VIR) for a long time, but they didn't buy anything.

Q18 *In contrast to inhabitants of Earth, Construlians love interplanetary travel. They spend a lot of time dreaming of what it would be like to live on other planets. To dream of living on other planets is* ***fażować.***

What does the sentence below mean?

He didn't feel like working so he po-fażowa-ł (po-fażować-PST.3SG.MASC).

◯ He fażował (fażować-PST.3SG.MASC) instead of working.

◯ He fażował (fażować-PST.3SG.MASC) until he was satisfied.

◯ He fażował (fażować-PST.3SG.MASC) about many different planets instead of working.

Q19 *The crime rates on Construlia aren't high because Construlians take utmost care of their safety. For instance, they always lock their doors with a few locks and* *padlocks. To lock doors is* ***gerdować.***

What does the sentence below mean?

He po-gerdowa-ł (po-gerdować-PST.3SG.MASC) the doors in the house.

○ He locked different doors in the house.

○ He unlocked the doors in the house.

○ He was locking the doors, but he didn't finish.

Q20 *Most Construlians wear really long and frizzy fringes. Styling them is a true ritual, which takes more than ten minutes each day. To style a fringe is* ***grzysać.***

What does the sentence below mean?

The barber po-grzysa-ł (po-grzysać-PST.3SG.MASC) his customers.

○ The barber style the fringe for a few customers.

○ The barber grzysa-ł (grzysać-PST.3SG.MASC) for some time (and then he did something else)

○ The barber grzysa-ł (grzysać-PST.3SG.MASC) his customers to the side.

APPENDIX 6: List of verbs whose domains have been changed/added in comparison to Polish WordNet

A. CORPUS STUDY ON REFLEXIVES

brzydzić, brzydzić się, chcieć się, dokonywać się, iść się, kreować się, liczyć się, marszczyć się, mierzyć, mieć się, mówić się, nadawać się, nadziwić się, nagłaśniać się, nazywać się, niecierpieć, nienawidzieć, obsikiwać, obwiniać, odbierać się, odbyć się, odczuć się, odkryć, odnaleźć się, okazać się, okazywać się, osiedzieć się, oszczędzać się, oszukiwać, otwierać się, ośmieszać, ośmieszyć, ożenić się, palić się, pilnować, pocić się, pogubić się, pojawić się, poskarżyć się, powstydzić się, przekazywać się, przekonać się, przeprowadzać się, przeprowadzić się, rezygnować się, rozlec się, rzutować, skakać się, skontaktować się, składać się, sprawdzać, sprawdzać się, stosować się, strzec, słuchać się, traktować, ubogacać, udziabać się, układać się, unieśmiertelnić, ustalać się, uważać się, uwidzieć się, używać się, wciskać się, wpisać się, wpędzać się, wyprzedzać się, wypróżniać się, wyrzucać się, wyłaniać się, wyświetlić się, zająć się, zakładać się, zalecać się, zaznaczać się, zintegrować się, zmanipulować się, zmierzyć się, zohydzać, śmiać się

B. CORPUS STUDY ON PREFIXES

pobudzać, poczytać, pojaśnić, poleniuchować, popełnieć, popisywać, poplażować, porastać, poskarżyć, potwierdzać, powpadać, pozdejmować, pozostawić, przybierać, przybyć, przydarzyć, przygotować, przygotowywać, przyjeżdzać, przyjąć, przyjść, przylgnąć, przynosić, przyozdobić, przypisywać, przypominać, przysięgać, przysługiwać, przyznawać, przyzwyczaić, rozbić, rozbrajać, rozchorować, rozchwytywać, rozciągają, rozjechać, rozmiał, rozpieklić, rozpieszczać, rozpisywać, rozprawić, rozrabiać, rozstrzygać, rozsychać, rozszerzyć, roztaczać

APPENDIX 7: Regression model outputs for reflexives corpus study

INITIAL MODEL

```
lrm(formula = marker ~ subj.type_abstract + subj.type_animate +
    aspect_perfective + overt.subject_YES + sam_YES + situation.type_impersonal +
    situation.type_passive + situation.type_reciprocal + v.inf_YES +
    verb.class_change + verb.class_cognition + verb.class_communication +
    verb.class_creation + verb.class_emotion + verb.class_motion +
    verb.class_perception + verb.class_possession + verb.class_social +
    verb.class_stative + volition_YES + emphasis_YES + negation_YES +
    tantum_YES)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs          1000    LR chi2     645.86    R2       0.634    C       0.912
 sie          509    d.f.            23    g        3.370    Dxy     0.824
 siebie       491    Pr(> chi2) <0.0001    gr      29.068    gamma   0.826
max |deriv| 1e-06                          gp       0.411    tau-a   0.412
                                           Brier    0.117

                          Coef    S.E.   Wald Z Pr(>|Z|)
Intercept                 -1.4792 0.8498 -1.74  0.0818
subj.type_abstract         0.9942 0.6442  1.54  0.1228
subj.type_animate          2.4115 0.6360  3.79  0.0001
aspect_perfective         -0.1603 0.2144 -0.75  0.4548
overt.subject_YES          0.0800 0.2350  0.34  0.7336
sam_YES                    3.7387 0.7054  5.30  <0.0001
situation.type_impersonal -2.0523 0.5248 -3.91  <0.0001
situation.type_passive    -1.7800 1.1904 -1.50  0.1348
situation.type_reciprocal  0.1597 0.3565  0.45  0.6541
v.inf_YES                  0.2611 0.2274  1.15  0.2509
verb.class_change         -1.7009 0.6478 -2.63  0.0086
verb.class_cognition      -0.4503 0.6262 -0.72  0.4721
verb.class_communication  -0.3058 0.6434 -0.48  0.6346
verb.class_creation        0.9908 1.0878  0.91  0.3624
verb.class_emotion        -0.4399 0.6928 -0.63  0.5255
verb.class_motion         -2.9547 0.7674 -3.85  0.0001
verb.class_perception      0.9336 0.6887  1.36  0.1752
verb.class_possession      0.7832 0.8201  0.96  0.3396
verb.class_social         -1.2604 0.6298 -2.00  0.0453
verb.class_stative        -2.9541 0.7381 -4.00  <0.0001
volition_YES               0.7904 0.2092  3.78  0.0002
emphasis_YES               0.4977 0.2503  1.99  0.0467
negation_YES              -0.2215 0.3879 -0.57  0.5679
tantum_YES                -5.7844 1.0677 -5.42  <0.0001
```

```
lrm(formula = marker ~ subj.type_inanimate + subj.type_animate +
    aspect_perfective + overt.subject_YES + sam_YES + situation.type_impersonal +
    situation.type_passive + situation.type_reciprocal + v.inf_YES +
    verb.class_change + verb.class_cognition + verb.class_communication +
    verb.class_creation + verb.class_emotion + verb.class_motion +
    verb.class_perception + verb.class_possession + verb.class_social +
    verb.class_stative + volition_YES + emphasis_YES + negation_YES +
    tantum_YES)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs          1000    LR chi2     645.86    R2       0.634    C       0.912
 sie          509    d.f.            23    g        3.370    Dxy     0.824
 siebie       491    Pr(> chi2) <0.0001    gr      29.068    gamma   0.826
max |deriv| 1e-06                          gp       0.411    tau-a   0.412
                                           Brier    0.117

                          Coef    S.E.   Wald Z Pr(>|Z|)
Intercept                 -0.4850 0.6435 -0.75  0.4510
subj.type_inanimate       -0.9942 0.6442 -1.54  0.1228
subj.type_animate          1.4173 0.2534  5.59  <0.0001
aspect_perfective         -0.1603 0.2144 -0.75  0.4548
overt.subject_YES          0.0800 0.2350  0.34  0.7336
sam_YES                    3.7387 0.7054  5.30  <0.0001
situation.type_impersonal -2.0523 0.5248 -3.91  <0.0001
situation.type_passive    -1.7800 1.1904 -1.50  0.1348
situation.type_reciprocal  0.1597 0.3565  0.45  0.6541
v.inf_YES                  0.2611 0.2274  1.15  0.2509
verb.class_change         -1.7009 0.6478 -2.63  0.0086
verb.class_cognition      -0.4503 0.6262 -0.72  0.4721
verb.class_communication  -0.3058 0.6434 -0.48  0.6346
verb.class_creation        0.9908 1.0878  0.91  0.3624
verb.class_emotion        -0.4399 0.6928 -0.63  0.5255
verb.class_motion         -2.9547 0.7674 -3.85  0.0001
verb.class_perception      0.9336 0.6887  1.36  0.1752
verb.class_possession      0.7832 0.8201  0.96  0.3396
verb.class_social         -1.2604 0.6298 -2.00  0.0453
verb.class_stative        -2.9541 0.7381 -4.00  <0.0001
volition_YES               0.7904 0.2092  3.78  0.0002
emphasis_YES               0.4977 0.2503  1.99  0.0467
negation_YES              -0.2215 0.3879 -0.57  0.5679
tantum_YES                -5.7844 1.0677 -5.42  <0.0001
```

```
lrm(formula = marker ~ subj.type_animate + aspect_perfective +
     overt.subject_YES + sam_YES + situation.type_impersonal +
     situation.type_passive + situation.type_reciprocal + v.inf_YES +
     verb.class_change + verb.class_cognition + verb.class_communication +
     verb.class_creation + verb.class_emotion + verb.class_motion +
     verb.class_perception + verb.class_possession + verb.class_social +
     verb.class_stative + volition_YES + emphasis_YES + negation_YES +
     tantum_YES)

                       Model Likelihood     Discrimination    Rank Discrim.
                          Ratio Test           Indexes           Indexes
 Obs          1000    LR chi2     643.14    R2        0.633    C       0.911
  sie          509    d.f.            22    g         3.332    Dxy     0.822
  siebie       491    Pr(> chi2) <0.0001    gr       27.980    gamma   0.824
 max |deriv| 1e-06                          gp        0.410    tau-a   0.411
                                            Brier     0.118

                          Coef    S.E.   Wald Z Pr(>|Z|)
 Intercept                -0.6411 0.6318 -1.01  0.3102
 subj.type_animate         1.5287 0.2461  6.21  <0.0001
 aspect_perfective        -0.1467 0.2145 -0.68  0.4941
 overt.subject_YES         0.0620 0.2340  0.26  0.7912
 sam_YES                   3.7112 0.7029  5.28  <0.0001
 situation.type_impersonal -1.9746 0.5239 -3.77  0.0002
 situation.type_passive    -1.8079 1.1833 -1.53  0.1265
 situation.type_reciprocal  0.1323 0.3537  0.37  0.7084
 v.inf_YES                 0.2609 0.2279  1.14  0.2523
 verb.class_change        -1.7034 0.6420 -2.65  0.0080
 verb.class_cognition     -0.4094 0.6206 -0.66  0.5095
 verb.class_communication  -0.2726 0.6382 -0.43  0.6692
 verb.class_creation       0.9236 1.0524  0.88  0.3802
 verb.class_emotion       -0.3916 0.6871 -0.57  0.5687
 verb.class_motion        -2.9450 0.7603 -3.87  0.0001
 verb.class_perception     0.9108 0.6809  1.34  0.1810
 verb.class_possession     0.8202 0.8118  1.01  0.3123
 verb.class_social        -1.2175 0.6238 -1.95  0.0509
 verb.class_stative       -2.9639 0.7309 -4.05  <0.0001
 volition_YES              0.8186 0.2086  3.92  <0.0001
 emphasis_YES              0.5015 0.2503  2.00  0.0452
 negation_YES             -0.2114 0.3881 -0.54  0.5858
 tantum_YES               -5.7640 1.0638 -5.42  <0.0001
```

```
lrm(formula = marker ~ subj.type_animate + sam_YES + situation.type_impersonal +
    situation.type_passive + situation.type_reciprocal + v.inf_YES +
    verb.class_change + verb.class_cognition + verb.class_communication +
    verb.class_creation + verb.class_emotion + verb.class_motion +
    verb.class_perception + verb.class_possession + verb.class_social +
    verb.class_stative + volition_YES + emphasis_YES + negation_YES +
    tantum_YES)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs          1000    LR chi2     642.48    R2       0.632    C       0.910
 sie          509    d.f.            20    g        3.331    Dxy     0.820
 siebie       491    Pr(> chi2) <0.0001    gr      27.958    gamma   0.825
max |deriv| 1e-06                          gp       0.410    tau-a   0.410
                                           Brier    0.118

                          Coef    S.E.   Wald Z Pr(>|Z|)
Intercept                 -0.6930 0.6138 -1.13  0.2589
subj.type_animate          1.5106 0.2442  6.19  <0.0001
sam_YES                    3.7170 0.7014  5.30  <0.0001
situation.type_impersonal -1.9606 0.5176 -3.79  0.0002
situation.type_passive    -1.8520 1.1943 -1.55  0.1210
situation.type_reciprocal  0.1477 0.3529  0.42  0.6755
v.inf_YES                  0.2268 0.2221  1.02  0.3071
verb.class_change         -1.7095 0.6433 -2.66  0.0079
verb.class_cognition      -0.3780 0.6194 -0.61  0.5417
verb.class_communication  -0.2334 0.6360 -0.37  0.7136
verb.class_creation        0.9272 1.0476  0.89  0.3761
verb.class_emotion        -0.3198 0.6760 -0.47  0.6362
verb.class_motion         -2.9490 0.7617 -3.87  0.0001
verb.class_perception      0.9481 0.6794  1.40  0.1629
verb.class_possession      0.8463 0.8102  1.04  0.2962
verb.class_social         -1.1751 0.6217 -1.89  0.0587
verb.class_stative        -2.9014 0.7268 -3.99  <0.0001
volition_YES               0.8202 0.2084  3.94  <0.0001
emphasis_YES               0.4974 0.2500  1.99  0.0466
negation_YES              -0.2012 0.3864 -0.52  0.6026
tantum_YES                -5.7698 1.0647 -5.42  <0.0001
```

```
lrm(formula = marker ~ subj.type_animate + sam_YES + situation.type_impersonal +
    situation.type_passive + situation.type_reflexive + v.inf_YES +
    verb.class_change + verb.class_cognition + verb.class_communication +
    verb.class_creation + verb.class_emotion + verb.class_motion +
    verb.class_perception + verb.class_possession + verb.class_social +
    verb.class_stative + volition_YES + emphasis_YES + negation_YES +
    tantum_YES)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs          1000    LR chi2      642.48    R2        0.632    C       0.910
 sie          509    d.f.             20    g         3.331    Dxy     0.820
 siebie       491    Pr(> chi2) <0.0001    gr       27.958    gamma   0.825
max |deriv| 1e-06                           gp        0.410    tau-a   0.410
                                            Brier     0.118

                          Coef    S.E.   Wald Z Pr(>|Z|)
Intercept                 -0.5453 0.6830 -0.80  0.4247
subj.type_animate          1.5106 0.2442  6.19  <0.0001
sam_YES                    3.7170 0.7014  5.30  <0.0001
situation.type_impersonal -2.1083 0.6136 -3.44  0.0006
situation.type_passive    -1.9998 1.2386 -1.61  0.1064
situation.type_reflexive  -0.1477 0.3529 -0.42  0.6755
v.inf_YES                  0.2268 0.2221  1.02  0.3071
verb.class_change         -1.7095 0.6433 -2.66  0.0079
verb.class_cognition      -0.3780 0.6194 -0.61  0.5417
verb.class_communication  -0.2334 0.6360 -0.37  0.7136
verb.class_creation        0.9272 1.0476  0.89  0.3761
verb.class_emotion        -0.3198 0.6760 -0.47  0.6362
verb.class_motion         -2.9490 0.7617 -3.87  0.0001
verb.class_perception      0.9481 0.6794  1.40  0.1629
verb.class_possession      0.8463 0.8102  1.04  0.2962
verb.class_social         -1.1751 0.6217 -1.89  0.0587
verb.class_stative        -2.9014 0.7268 -3.99  <0.0001
volition_YES               0.8202 0.2084  3.94  <0.0001
emphasis_YES               0.4974 0.2500  1.99  0.0466
negation_YES              -0.2012 0.3864 -0.52  0.6026
tantum_YES                -5.7698 1.0647 -5.42  <0.0001
```

```
lrm(formula = marker ~ subj.type_animate + sam_YES + situation.type_impersonal +
     situation.type_passive + v.inf_YES + verb.class_change +
     verb.class_cognition + verb.class_communication + verb.class_creation +
     verb.class_emotion + verb.class_motion + verb.class_perception +
     verb.class_possession + verb.class_social + verb.class_stative +
     volition_YES + emphasis_YES + negation_YES + tantum_YES,
     data = onehot_vars)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
 Obs          1000    LR chi2     642.30    R2       0.632    C       0.910
  sie          509    d.f.            19    g        3.330    Dxy     0.821
  siebie       491    Pr(> chi2) <0.0001    gr      27.930    gamma   0.826
 max |deriv| 1e-06                          gp       0.410    tau-a   0.411
                                            Brier    0.118

                          Coef    S.E.   Wald Z Pr(>|Z|)
 Intercept                -0.6724 0.6129 -1.10  0.2726
 subj.type_animate         1.5144 0.2440  6.21  <0.0001
 sam_YES                   3.7064 0.7012  5.29  <0.0001
 situation.type_impersonal -1.9703 0.5170 -3.81  0.0001
 situation.type_passive    -1.8622 1.1934 -1.56  0.1186
 v.inf_YES                  0.2281 0.2221  1.03  0.3043
 verb.class_change         -1.7285 0.6428 -2.69  0.0072
 verb.class_cognition      -0.3981 0.6186 -0.64  0.5199
 verb.class_communication  -0.2491 0.6358 -0.39  0.6952
 verb.class_creation        0.9048 1.0470  0.86  0.3875
 verb.class_emotion        -0.3313 0.6761 -0.49  0.6242
 verb.class_motion         -2.9606 0.7620 -3.89  0.0001
 verb.class_perception      0.9293 0.6787  1.37  0.1710
 verb.class_possession      0.8405 0.8101  1.04  0.2995
 verb.class_social         -1.1761 0.6225 -1.89  0.0588
 verb.class_stative        -2.9144 0.7270 -4.01  <0.0001
 volition_YES               0.8227 0.2082  3.95  <0.0001
 emphasis_YES               0.4978 0.2497  1.99  0.0462
 negation_YES              -0.1958 0.3859 -0.51  0.6119
 tantum_YES                -5.7733 1.0641 -5.43  <0.0001
```

```
lrm(formula = marker ~ subj.type_animate + sam_YES + situation.type_impersonal +
    situation.type_passive + v.inf_YES + verb.class_change +
    verb.class_cognition + verb.class_communication + verb.class_creation +
    verb.class_emotion + verb.class_motion + verb.class_perception +
    verb.class_possession + verb.class_social + verb.class_stative +
    volition_YES + emphasis_YES + tantum_YES)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs          1000    LR chi2     642.05    R2       0.632    C       0.910
 sie          509    d.f.            18    g        3.325    Dxy     0.820
 siebie       491    Pr(> chi2) <0.0001    gr      27.789    gamma   0.826
max |deriv| 1e-06                          gp       0.410    tau-a   0.410
                                           Brier    0.118

                          Coef    S.E.   Wald Z Pr(>|Z|)
Intercept                 -0.6798 0.6144 -1.11  0.2686
subj.type_animate          1.5047 0.2432  6.19  <0.0001
sam_YES                    3.7060 0.7022  5.28  <0.0001
situation.type_impersonal -1.9829 0.5178 -3.83  0.0001
situation.type_passive    -1.8536 1.1919 -1.56  0.1199
v.inf_YES                  0.2262 0.2219  1.02  0.3080
verb.class_change         -1.7238 0.6445 -2.67  0.0075
verb.class_cognition      -0.3942 0.6203 -0.64  0.5251
verb.class_communication  -0.2384 0.6371 -0.37  0.7083
verb.class_creation        0.9170 1.0473  0.88  0.3812
verb.class_emotion        -0.3235 0.6777 -0.48  0.6331
verb.class_motion         -2.9520 0.7631 -3.87  0.0001
verb.class_perception      0.9293 0.6803  1.37  0.1719
verb.class_possession      0.8422 0.8119  1.04  0.2996
verb.class_social         -1.1625 0.6236 -1.86  0.0623
verb.class_stative        -2.9099 0.7284 -3.99  <0.0001
volition_YES               0.8167 0.2078  3.93  <0.0001
emphasis_YES               0.4959 0.2495  1.99  0.0469
tantum_YES                -5.7675 1.0630 -5.43  <0.0001
```

```
lrm(formula = marker ~ subj.type_animate + sam_YES + situation.type_impersonal +
     situation.type_passive + verb.class_change + verb.class_cognition +
     verb.class_communication + verb.class_creation + verb.class_emotion +
     verb.class_motion + verb.class_perception + verb.class_possession +
     verb.class_social + verb.class_stative + volition_YES + emphasis_YES +
     tantum_YES)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
 Obs          1000    LR chi2     641.00    R2       0.631    C       0.910
  sie          509    d.f.            17    g        3.320    Dxy     0.820
  siebie       491    Pr(> chi2) <0.0001    gr      27.663    gamma   0.831
 max |deriv| 1e-06                          gp       0.409    tau-a   0.411
                                            Brier    0.118

                          Coef    S.E.   Wald Z Pr(>|Z|)
 Intercept                -0.6307 0.6100 -1.03  0.3011
 subj.type_animate         1.5224 0.2426  6.27  <0.0001
 sam_YES                   3.6834 0.7051  5.22  <0.0001
 situation.type_impersonal -2.0388 0.5148 -3.96  <0.0001
 situation.type_passive    -1.8614 1.1995 -1.55  0.1207
 verb.class_change        -1.7312 0.6417 -2.70  0.0070
 verb.class_cognition     -0.3904 0.6178 -0.63  0.5274
 verb.class_communication  -0.2230 0.6348 -0.35  0.7254
 verb.class_creation       0.8784 1.0435  0.84  0.3999
 verb.class_emotion       -0.3489 0.6745 -0.52  0.6050
 verb.class_motion        -2.9534 0.7619 -3.88  0.0001
 verb.class_perception     0.9125 0.6777  1.35  0.1781
 verb.class_possession     0.8447 0.8084  1.04  0.2961
 verb.class_social        -1.1540 0.6211 -1.86  0.0632
 verb.class_stative       -2.9379 0.7259 -4.05  <0.0001
 volition_YES              0.8000 0.2067  3.87  0.0001
 emphasis_YES              0.5085 0.2490  2.04  0.0411
 tantum_YES               -5.7520 1.0571 -5.44  <0.0001
```

```
lrm(formula = marker ~ subj.type_animate + sam_YES + situation.type_impersonal +
     situation.type_passive + verb.class_change + verb.class_creation +
     verb.class_emotion + verb.class_motion + verb.class_perception +
     verb.class_possession + verb.class_social + verb.class_stative +
     volition_YES + emphasis_YES + tantum_YES)

                       Model Likelihood     Discrimination    Rank Discrim.
                          Ratio Test           Indexes           Indexes
 Obs          1000    LR chi2     640.43    R2       0.631    C       0.910
  sie          509    d.f.            15    g        3.315    Dxy     0.821
  siebie       491    Pr(> chi2) <0.0001    gr      27.511    gamma   0.839
 max |deriv| 1e-06                          gp       0.409    tau-a   0.411
                                            Brier    0.118

                           Coef    S.E.   Wald Z Pr(>|Z|)
 Intercept                 -0.9364 0.2638 -3.55  0.0004
 subj.type_animate          1.5157 0.2410  6.29  <0.0001
 sam_YES                    3.6929 0.7046  5.24  <0.0001
 situation.type_impersonal -2.0261 0.5161 -3.93  <0.0001
 situation.type_passive    -1.8467 1.2007 -1.54  0.1240
 verb.class_change         -1.4242 0.3078 -4.63  <0.0001
 verb.class_creation        1.1818 0.8846  1.34  0.1816
 verb.class_emotion        -0.0357 0.3511 -0.10  0.9190
 verb.class_motion         -2.6507 0.5058 -5.24  <0.0001
 verb.class_perception      1.2204 0.3729  3.27  0.0011
 verb.class_possession      1.1457 0.5853  1.96  0.0503
 verb.class_social         -0.8521 0.2567 -3.32  0.0009
 verb.class_stative        -2.6308 0.4556 -5.77  <0.0001
 volition_YES               0.8182 0.2054  3.98  <0.0001
 emphasis_YES               0.4928 0.2481  1.99  0.0470
 tantum_YES                -5.7858 1.0570 -5.47  <0.0001


lrm(formula = marker ~ subj.type_animate + sam_YES + situation.type_impersonal +
     situation.type_passive + verb.class_change + verb.class_motion +
     verb.class_perception + verb.class_possession + verb.class_social +
     verb.class_stative + volition_YES + emphasis_YES + tantum_YES,
     data = onehot_vars)

                       Model Likelihood     Discrimination    Rank Discrim.
                          Ratio Test           Indexes           Indexes
 Obs          1000    LR chi2     638.38    R2       0.629    C       0.910
  sie          509    d.f.            13    g        3.303    Dxy     0.819
  siebie       491    Pr(> chi2) <0.0001    gr      27.200    gamma   0.841
 max |deriv| 1e-06                          gp       0.408    tau-a   0.410
                                            Brier    0.118

                           Coef    S.E.   Wald Z Pr(>|Z|)
 Intercept                 -0.8658 0.2468 -3.51  0.0005
 subj.type_animate          1.4524 0.2332  6.23  <0.0001
 sam_YES                    3.6726 0.7040  5.22  <0.0001
 situation.type_impersonal -2.0622 0.5128 -4.02  <0.0001
 situation.type_passive    -1.8779 1.1923 -1.58  0.1153
 verb.class_change         -1.4658 0.2969 -4.94  <0.0001
 verb.class_motion         -2.6704 0.5013 -5.33  <0.0001
 verb.class_perception      1.1981 0.3631  3.30  0.0010
 verb.class_possession      1.1111 0.5799  1.92  0.0554
 verb.class_social         -0.8812 0.2482 -3.55  0.0004
 verb.class_stative        -2.6687 0.4486 -5.95  <0.0001
 volition_YES               0.8322 0.1959  4.25  <0.0001
 emphasis_YES               0.5056 0.2475  2.04  0.0411
 tantum_YES                -5.8083 1.0562 -5.50  <0.0001
```

```
lrm(formula = marker ~ subj.type_animate + sam_YES + situation.type_impersonal +
     verb.class_change + verb.class_motion + verb.class_perception +
     verb.class_possession + verb.class_social + verb.class_stative +
     volition_YES + emphasis_YES + tantum_YES)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test            Indexes           Indexes
 Obs          1000    LR chi2     634.99    R2       0.627    C       0.909
  sie          509    d.f.            12    g        3.296    Dxy     0.818
  siebie       491    Pr(> chi2) <0.0001    gr      27.010    gamma   0.839
 max |deriv| 1e-06                          gp       0.407    tau-a   0.409
                                            Brier    0.119

                           Coef    S.E.   Wald Z Pr(>|Z|)
 Intercept                 -0.9400 0.2433 -3.86  0.0001
 subj.type_animate          1.5138 0.2305  6.57  <0.0001
 sam_YES                    3.6969 0.7032  5.26  <0.0001
 situation.type_impersonal -1.9993 0.5115 -3.91  <0.0001
 verb.class_change         -1.4377 0.2967 -4.85  <0.0001
 verb.class_motion         -2.6704 0.5020 -5.32  <0.0001
 verb.class_perception      1.1754 0.3607  3.26  0.0011
 verb.class_possession      1.0168 0.5569  1.83  0.0679
 verb.class_social         -0.8882 0.2476 -3.59  0.0003
 verb.class_stative        -2.6465 0.4485 -5.90  <0.0001
 volition_YES               0.8493 0.1952  4.35  <0.0001
 emphasis_YES               0.5026 0.2463  2.04  0.0413
 tantum_YES                -5.8043 1.0571 -5.49  <0.0001
```

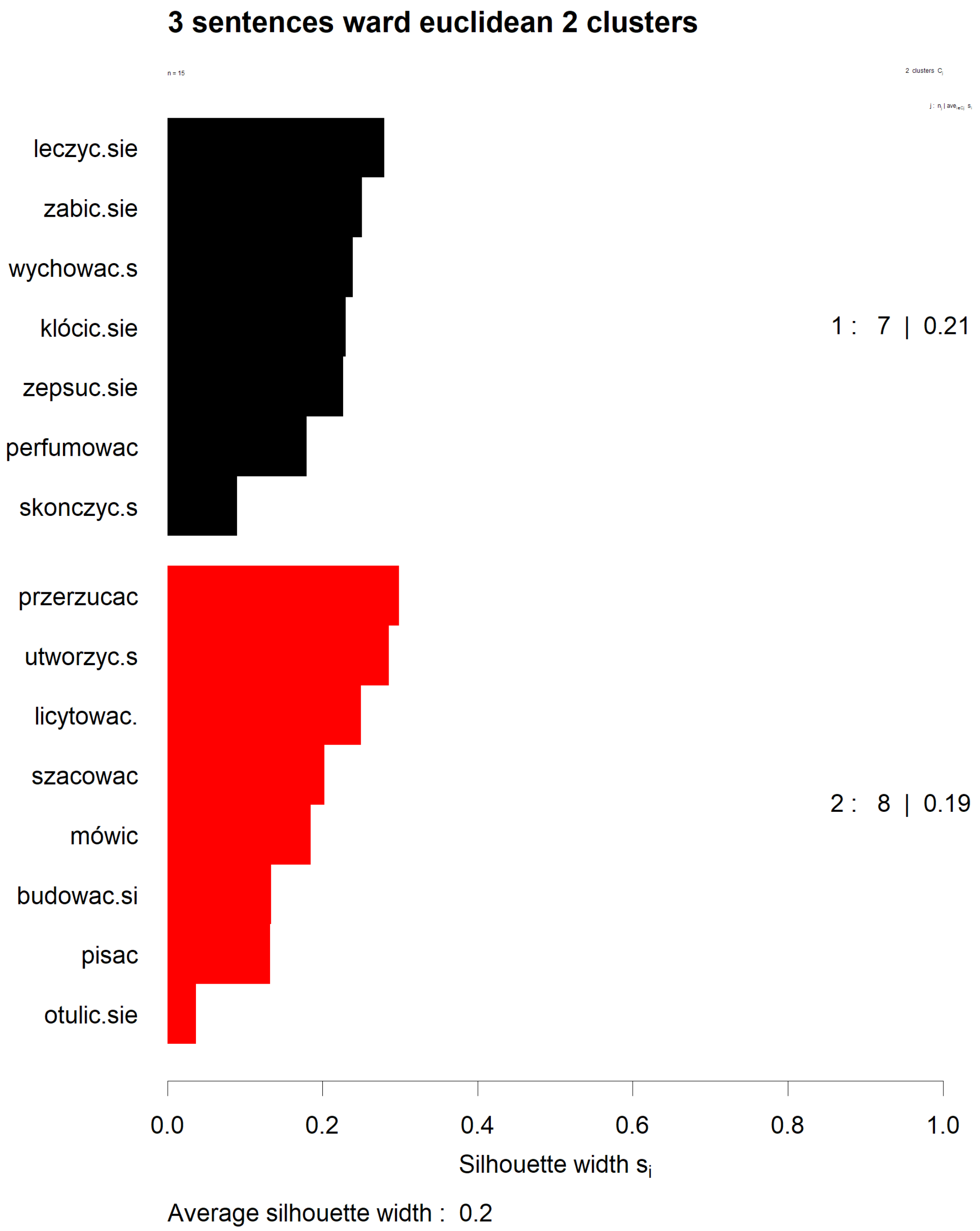
FINAL MODEL

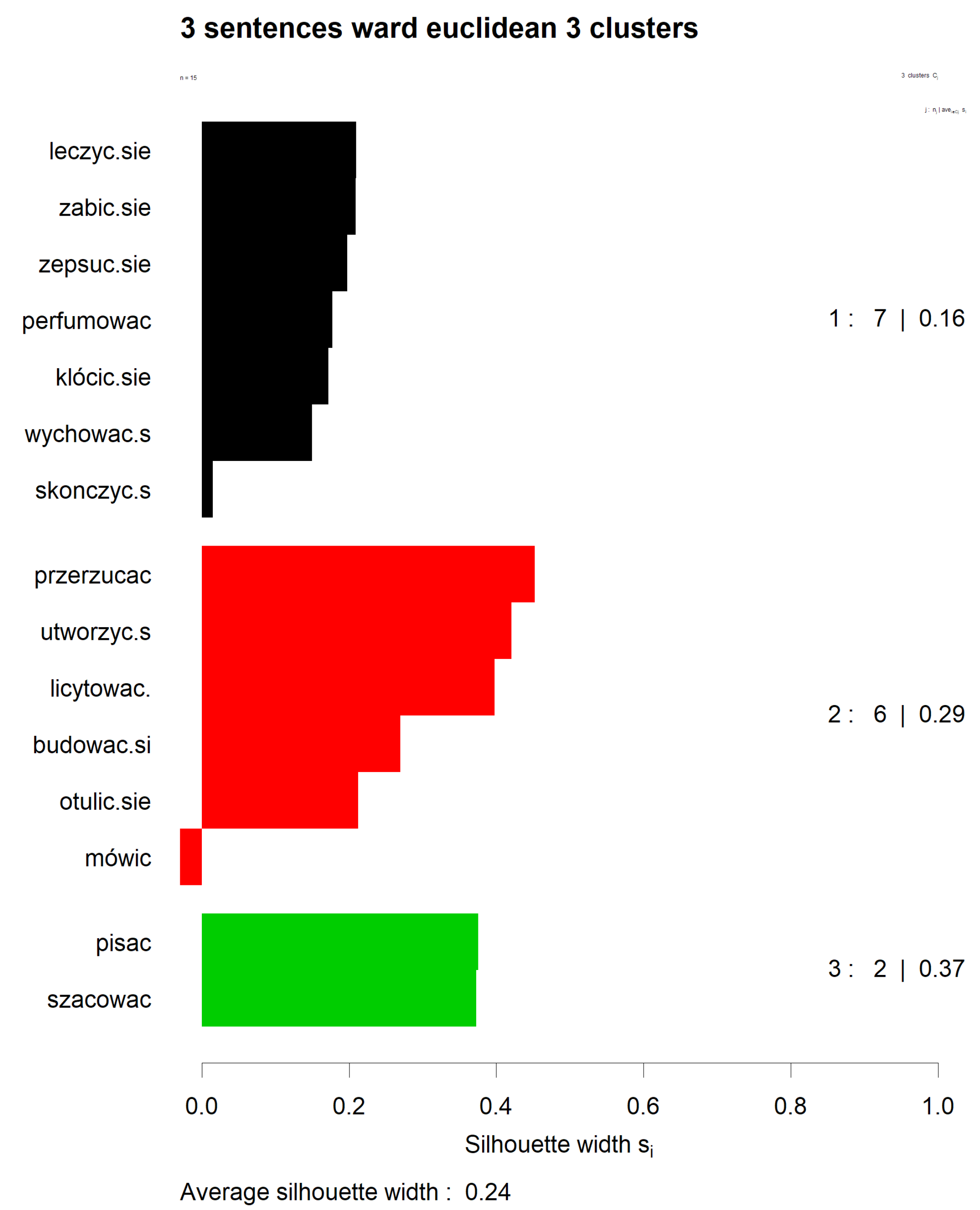
```
 lrm(formula = marker ~ subj.type_animate + sam_YES + situation.type_impersonal +
    verb.class_change + verb.class_motion + verb.class_perception +
    verb.class_social + verb.class_stative + volition_YES + emphasis_YES +
    tantum_YES)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test            Indexes           Indexes
 Obs          1000    LR chi2     631.21    R2       0.624    C       0.907
  sie          509    d.f.            11    g        3.274    Dxy     0.814
  siebie       491    Pr(> chi2) <0.0001    gr      26.417    gamma   0.838
 max |deriv| 1e-06                          gp       0.406    tau-a   0.407
                                            Brier    0.120

                           Coef    S.E.   Wald Z Pr(>|Z|)
 Intercept                 -0.8516 0.2374 -3.59  0.0003
 subj.type_animate          1.4870 0.2284  6.51  <0.0001
 sam_YES                    3.6917 0.7042  5.24  <0.0001
 situation.type_impersonal -1.9289 0.5076 -3.80  0.0001
 verb.class_change         -1.5098 0.2940 -5.13  <0.0001
 verb.class_motion         -2.7312 0.5006 -5.46  <0.0001
 verb.class_perception      1.0995 0.3575  3.08  0.0021
 verb.class_social         -0.9551 0.2452 -3.90  <0.0001
 verb.class_stative        -2.7164 0.4470 -6.08  <0.0001
 volition_YES               0.8508 0.1946  4.37  <0.0001
 emphasis_YES               0.4919 0.2461  2.00  0.0456
 tantum_YES                -5.8411 1.0559 -5.53  <0.0001
```

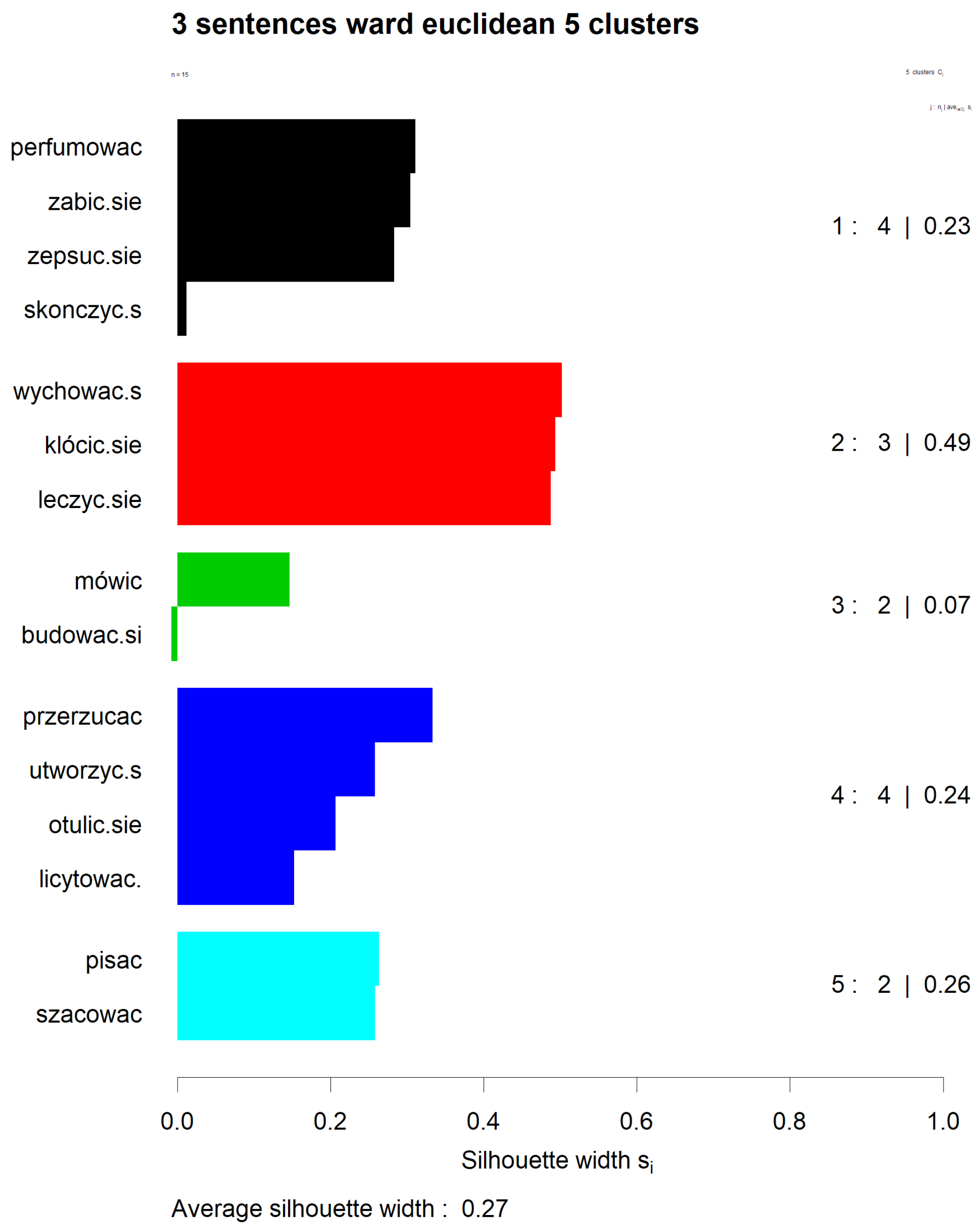
FINAL MODEL FOR ONLY FINITE AND INFINITIVE VERB FORMS

```
lrm(formula = marker ~ subj.type_animate + sam_YES + sie.type_impersonal +
    verb.class_change + verb.class_motion + verb.class_perception +
    verb.class_social + verb.class_stative + volition_YES +
    emphasis_YES + tantum_YES)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs           952    LR chi2     618.09    R2       0.638    C       0.912
 sie          505    d.f.            11    g        3.322    Dxy     0.825
 siebie       447    Pr(> chi2) <0.0001    gr      27.712    gamma   0.852
max |deriv| 7e-07                          gp       0.408    tau-a   0.411
                                           Brier    0.115

                          Coef    S.E.   Wald Z Pr(>|Z|)
Intercept                 -1.3906 0.2780 -5.00  <0.0001
subj.type_animate          2.0417 0.2727  7.49  <0.0001
sam_YES                    3.5988 0.7335  4.91  <0.0001
sie.type_impersonal       -1.6430 0.5332 -3.08  0.0021
verb.class_change         -1.5608 0.3312 -4.71  <0.0001
verb.class_motion         -2.7814 0.5373 -5.18  <0.0001
verb.class_perception      1.0810 0.3721  2.91  0.0037
verb.class_social         -0.7815 0.2576 -3.03  0.0024
verb.class_stative        -2.5378 0.4636 -5.47  <0.0001
volition_YES               0.6244 0.2053  3.04  0.0024
emphasis_YES               0.5721 0.2550  2.24  0.0249
tantum_YES                -5.5922 1.0428 -5.36  <0.0001
```

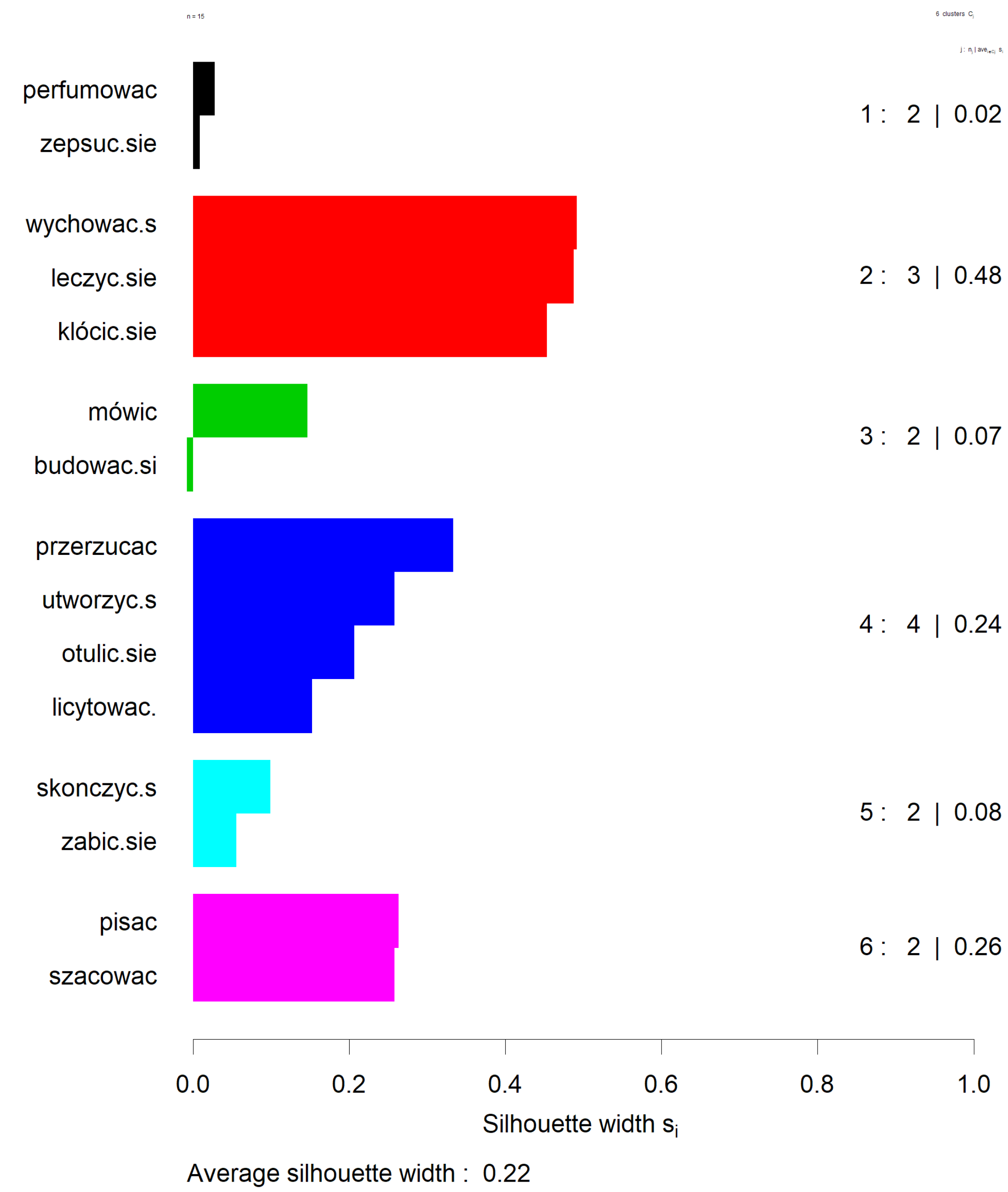
APPENDIX 8. Silhouette plots for the 3-sentence-per-meaning experiment on reflexives




3 sentences ward euclidean 3 clusters
n = 15
3 clusters C_j
leczyc.sie
zabic.sie
zepsuc.sie
perfumowac
klócic.sie
wychowac.s
skonczyc.s
1 : 7 | 0.16
przerzucac
utworzyc.s
licytowac.
budowac.si
otulic.sie
mówic
2 : 6 | 0.29
pisac
szacowac
3 : 2 | 0.37
0.0
0.2
0.4
0.6
0.8
1.0
Silhouette width s_i
Average silhouette width : 0.24

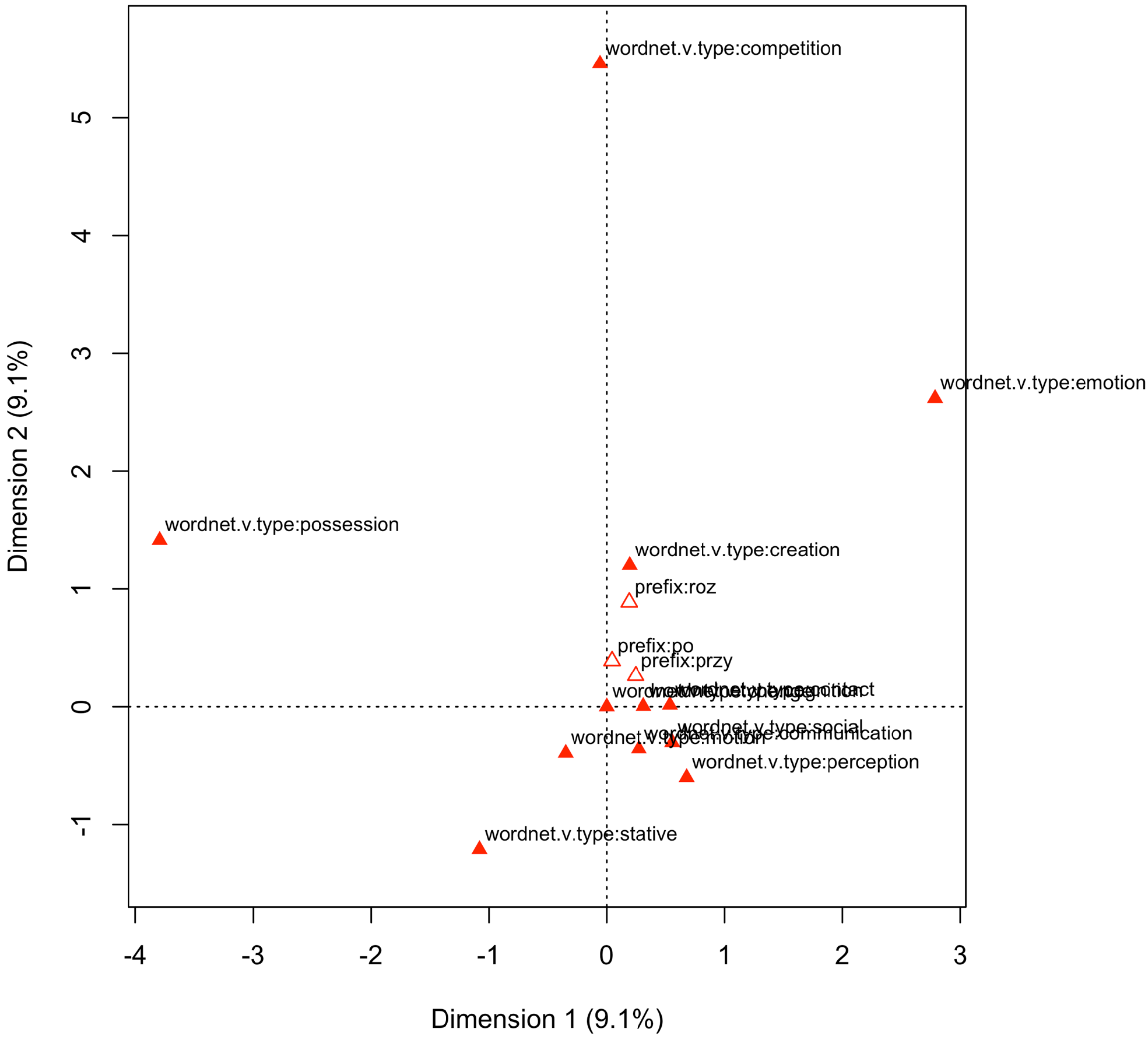

3 sentences ward euclidean 5 clusters
perfumowac
zabic.sie
zepsuc.sie
skonczyc.s
wychowac.s
kłócic.sie
leczyc.sie
mówic
budowac.si
przerzucac
utworzyc.s
otulic.sie
licytowac.
pisac
szacowac
1 : 4 | 0.23
2 : 3 | 0.49
3 : 2 | 0.07
4 : 4 | 0.24
5 : 2 | 0.26
0.0
0.2
0.4
0.6
0.8
1.0
Silhouette width s_i
Average silhouette width : 0.27

3 sentences ward euclidean 6 clusters

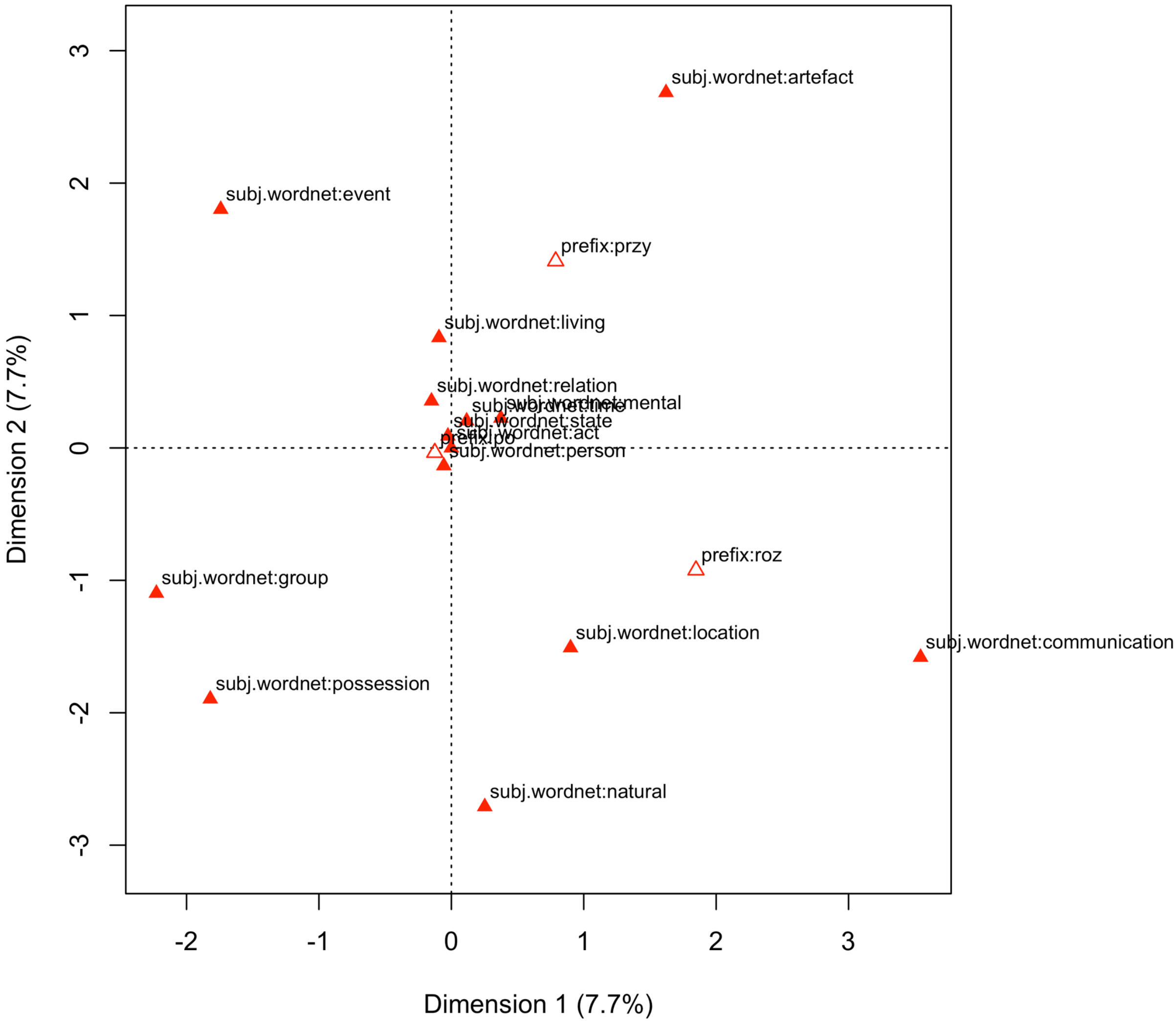


APPENDIX 9. Selected correspondence analysis plots for the corpus study on prefixes

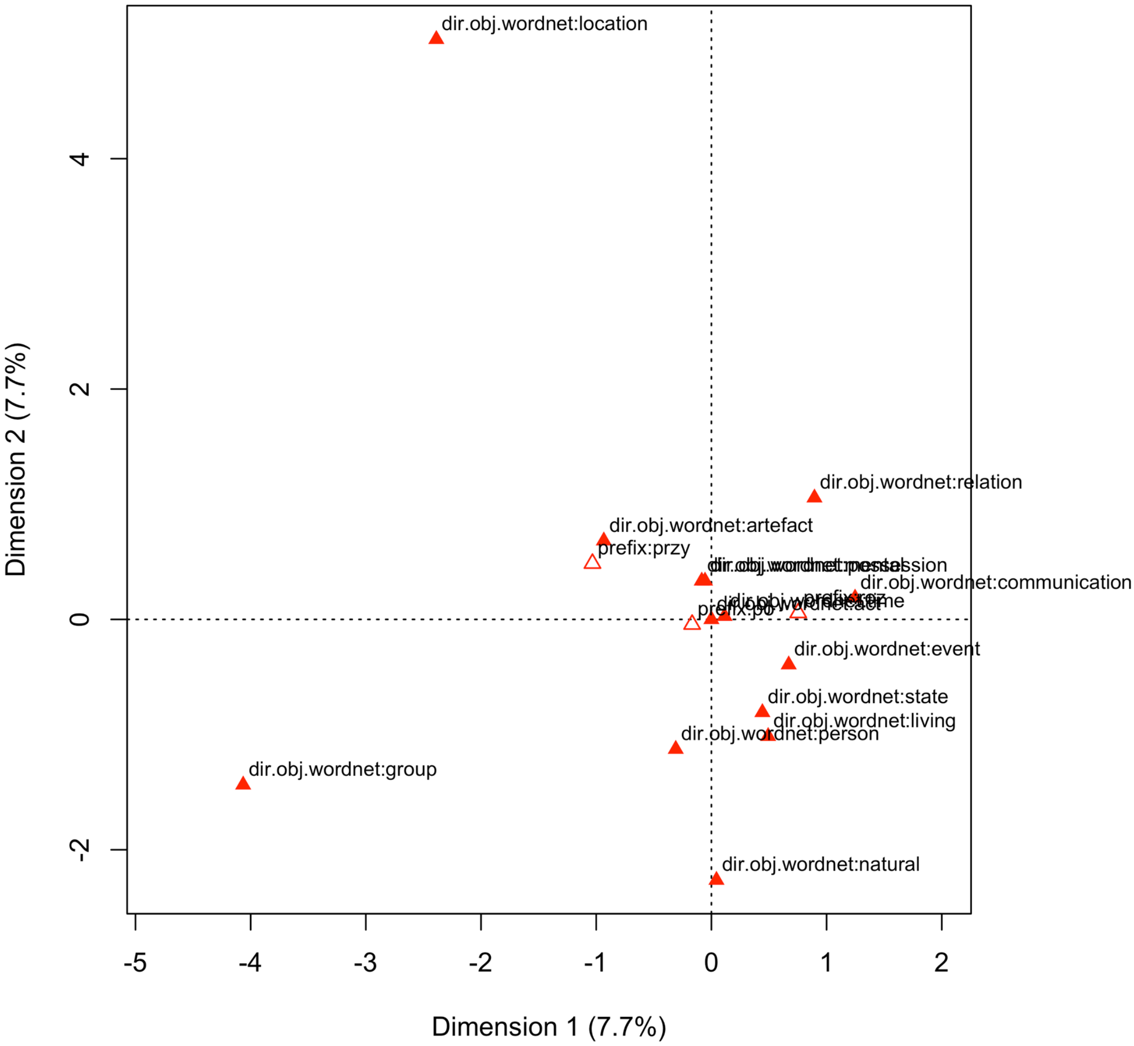
A. Verb semantic class only



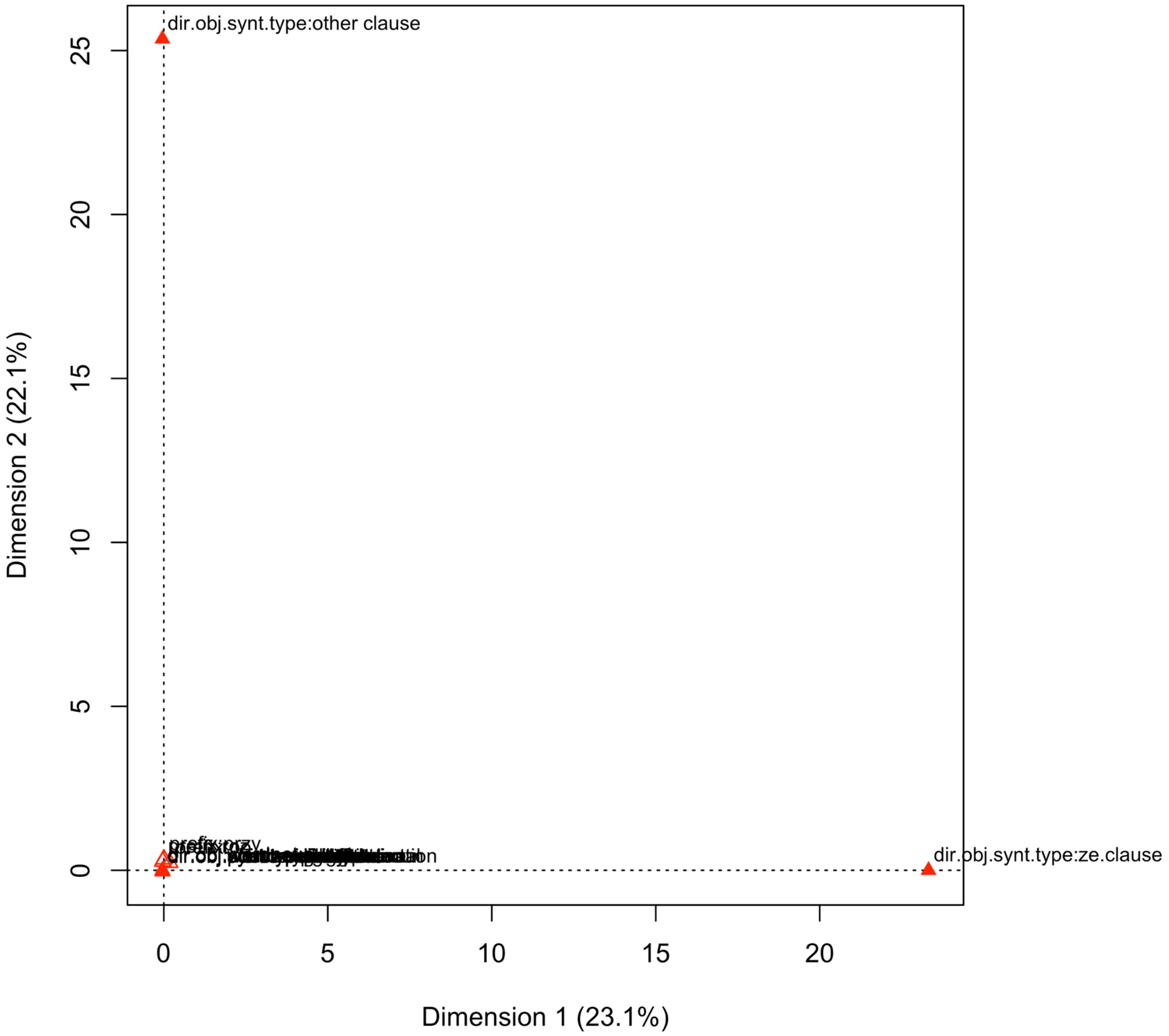
B. Subject semantic class only



C. Object semantic class only

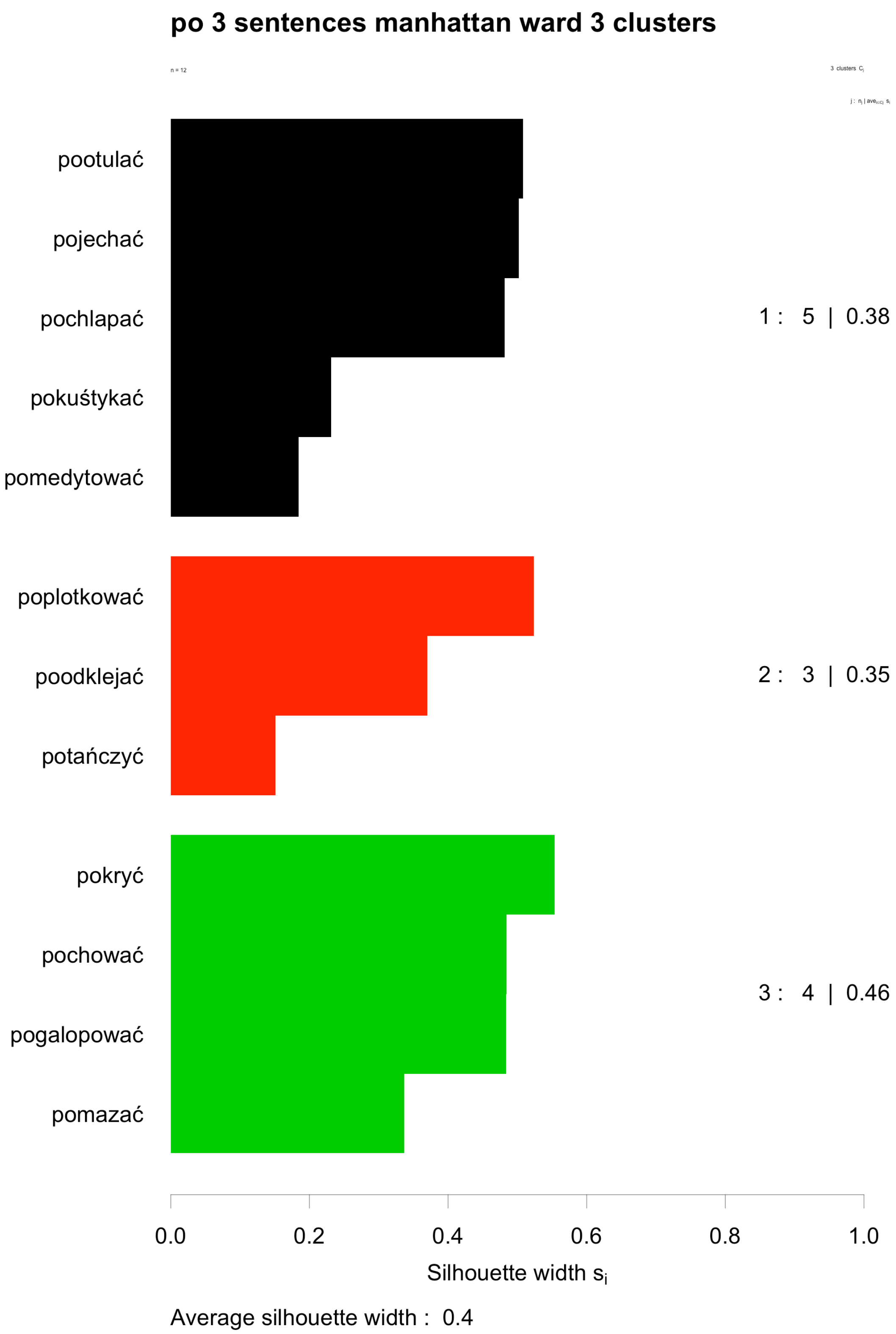


D. All object-related variables

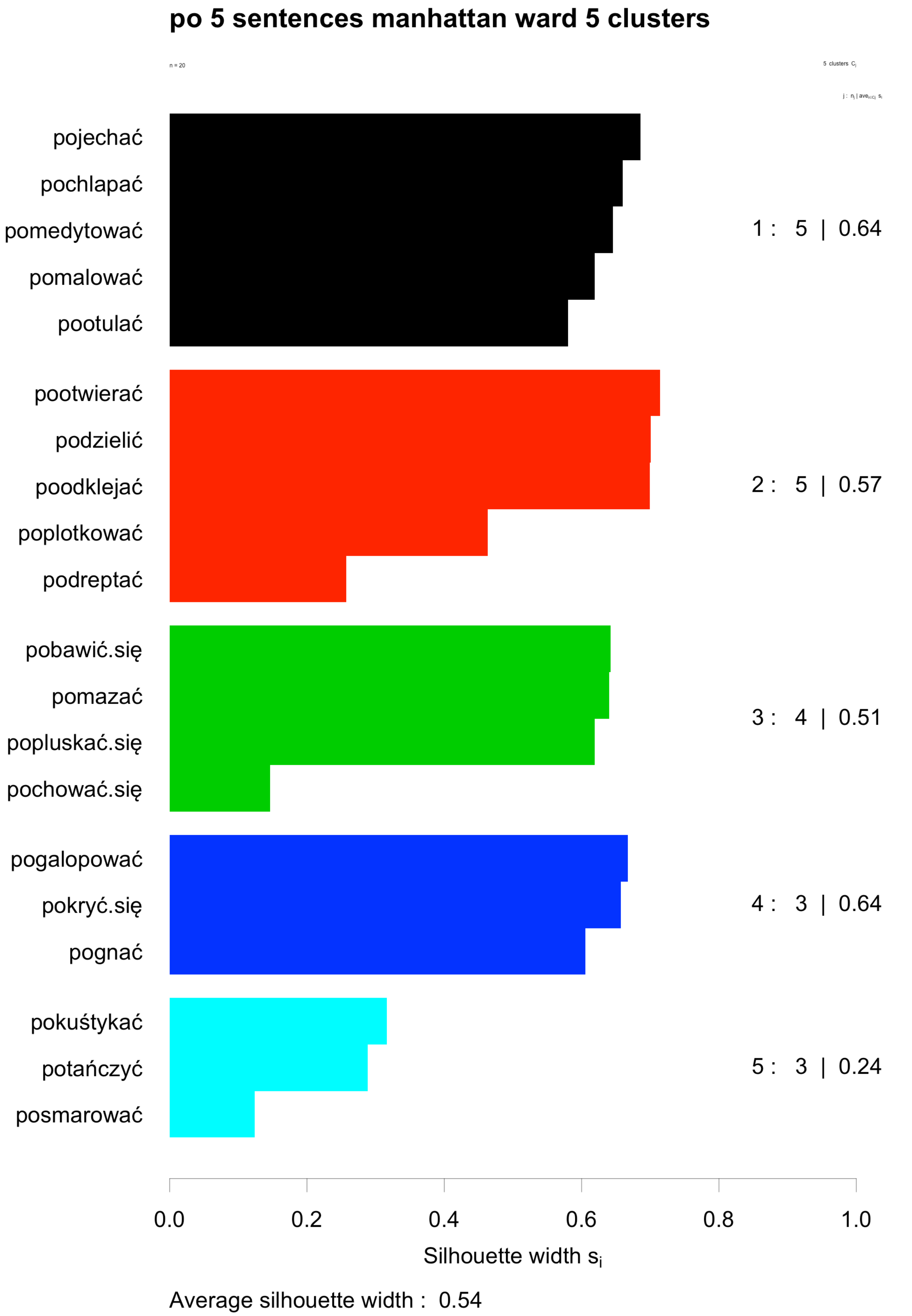


<u>APPENDIX 10. Silhouette plots for the optimal clustering solutions for the prefix sentence-sorting experiments</u>

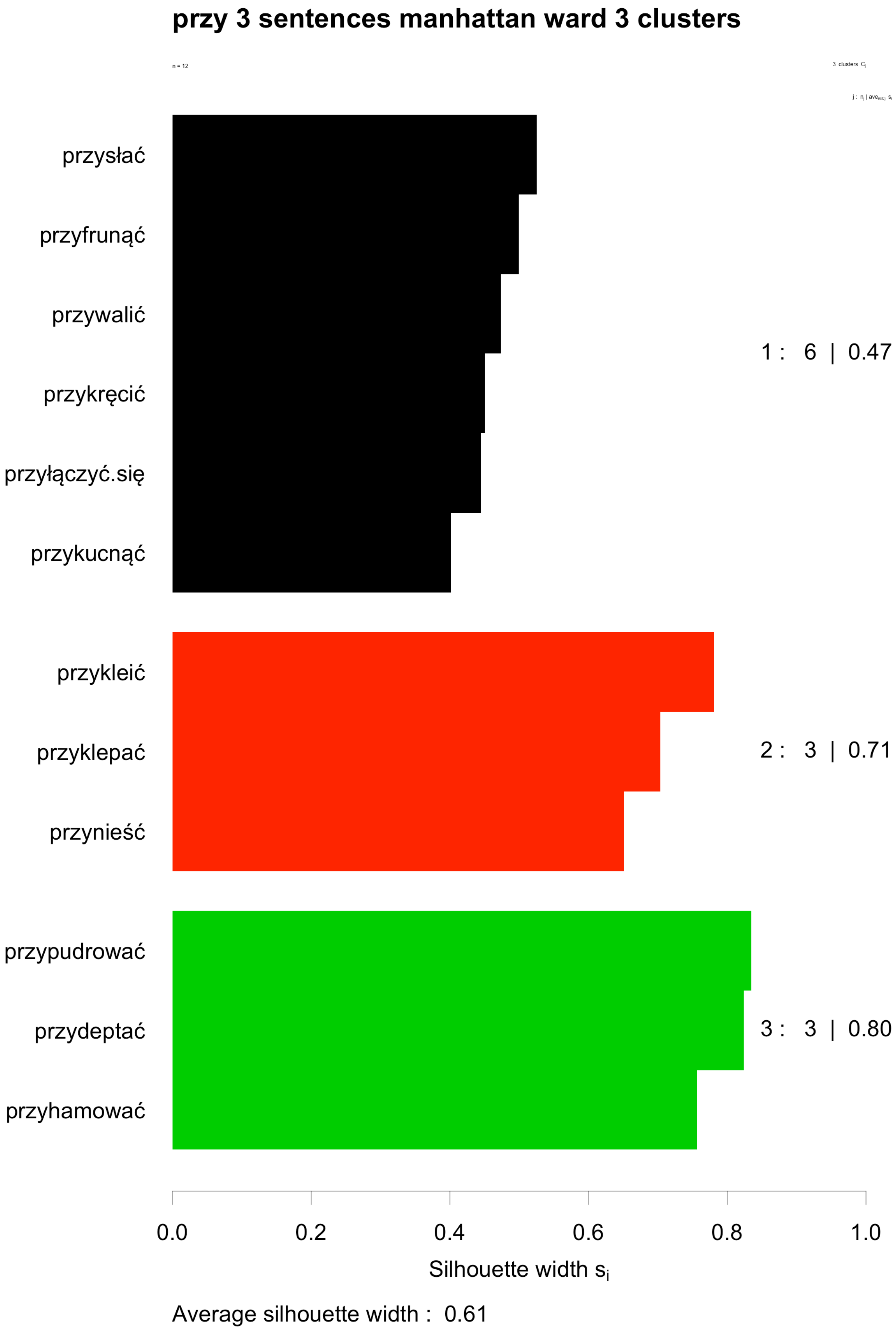
A. *po-* 3 sentences per meaning



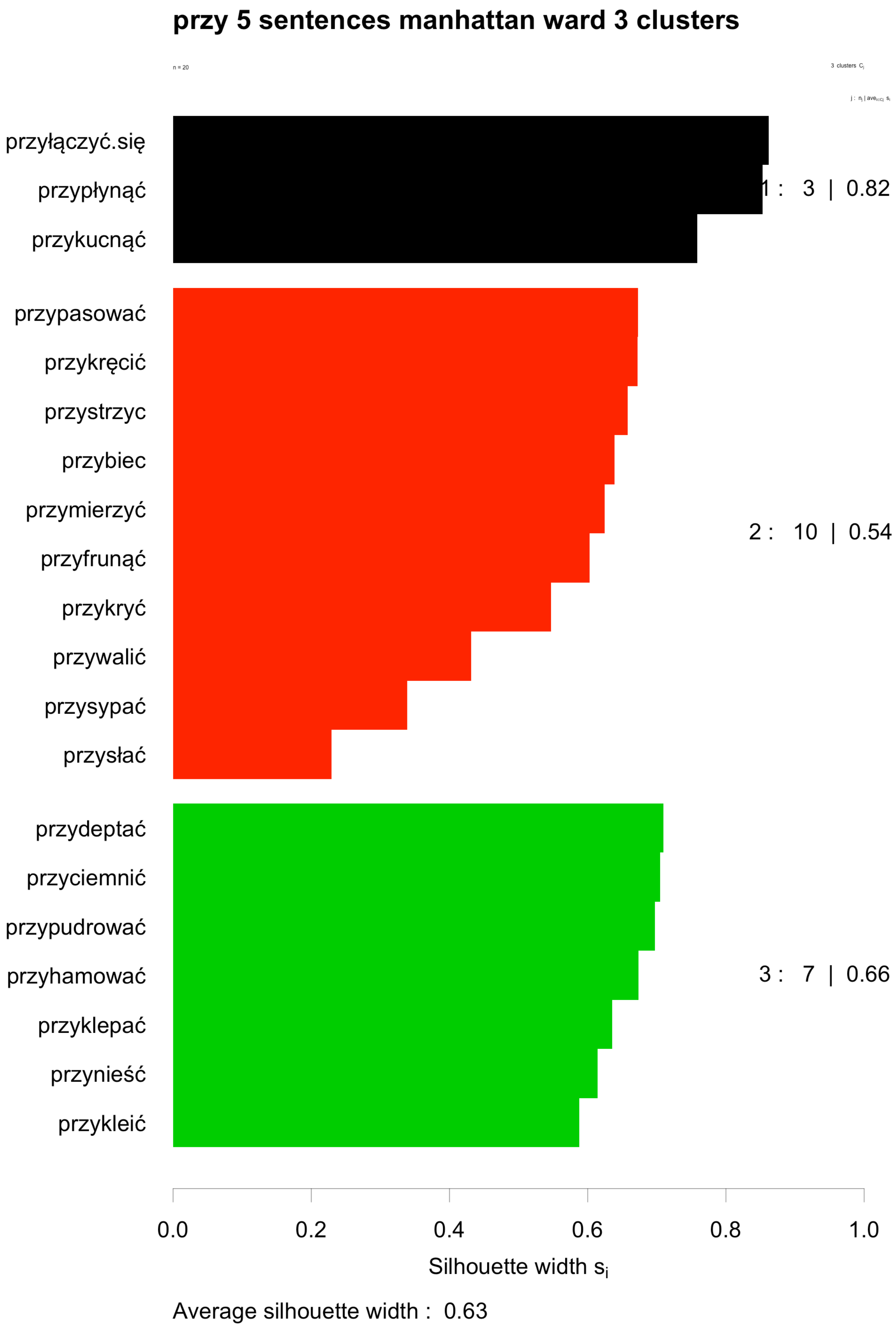
B. *po-* 5 sentences per meaning

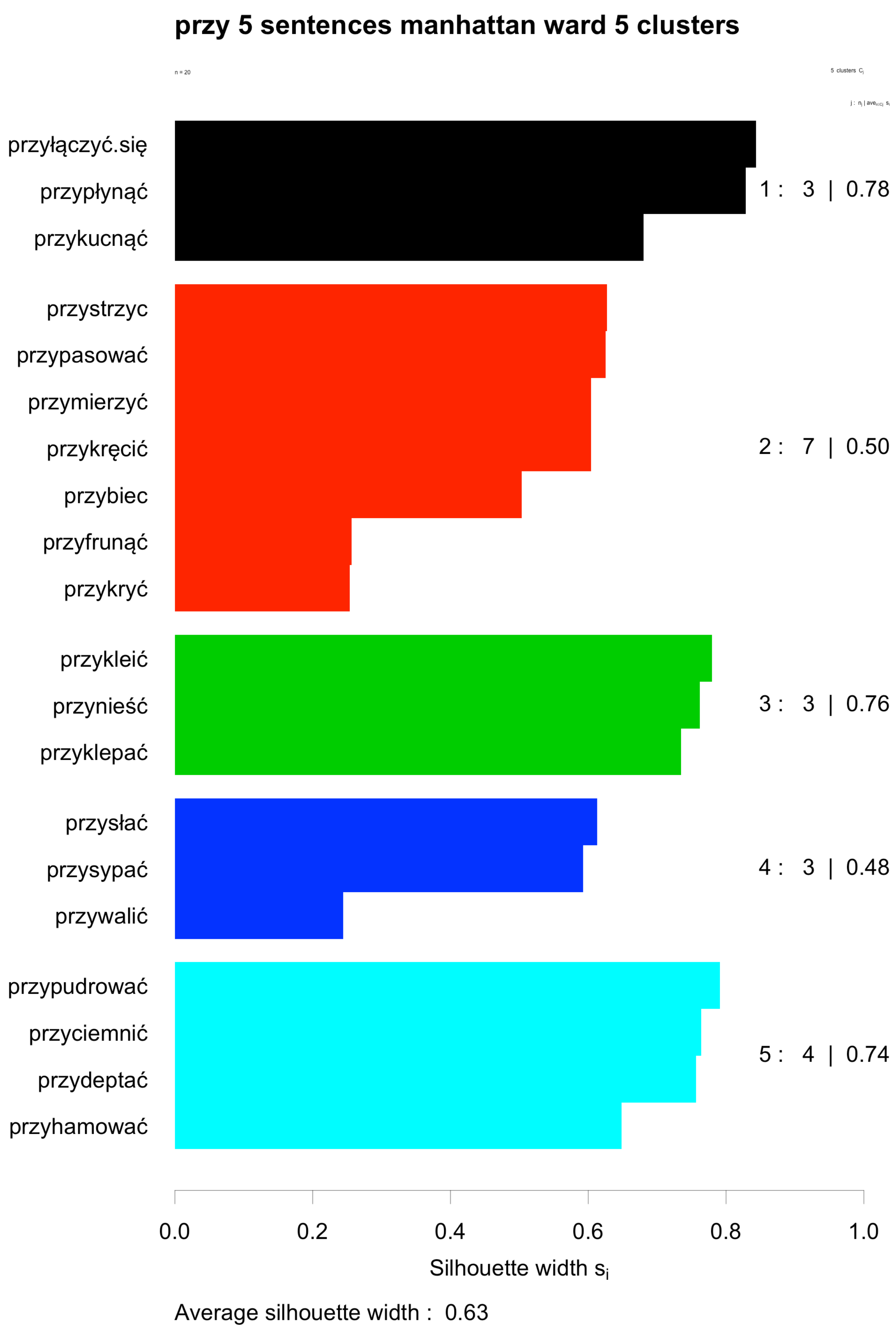


C. *przy-* 3 sentences per meaning

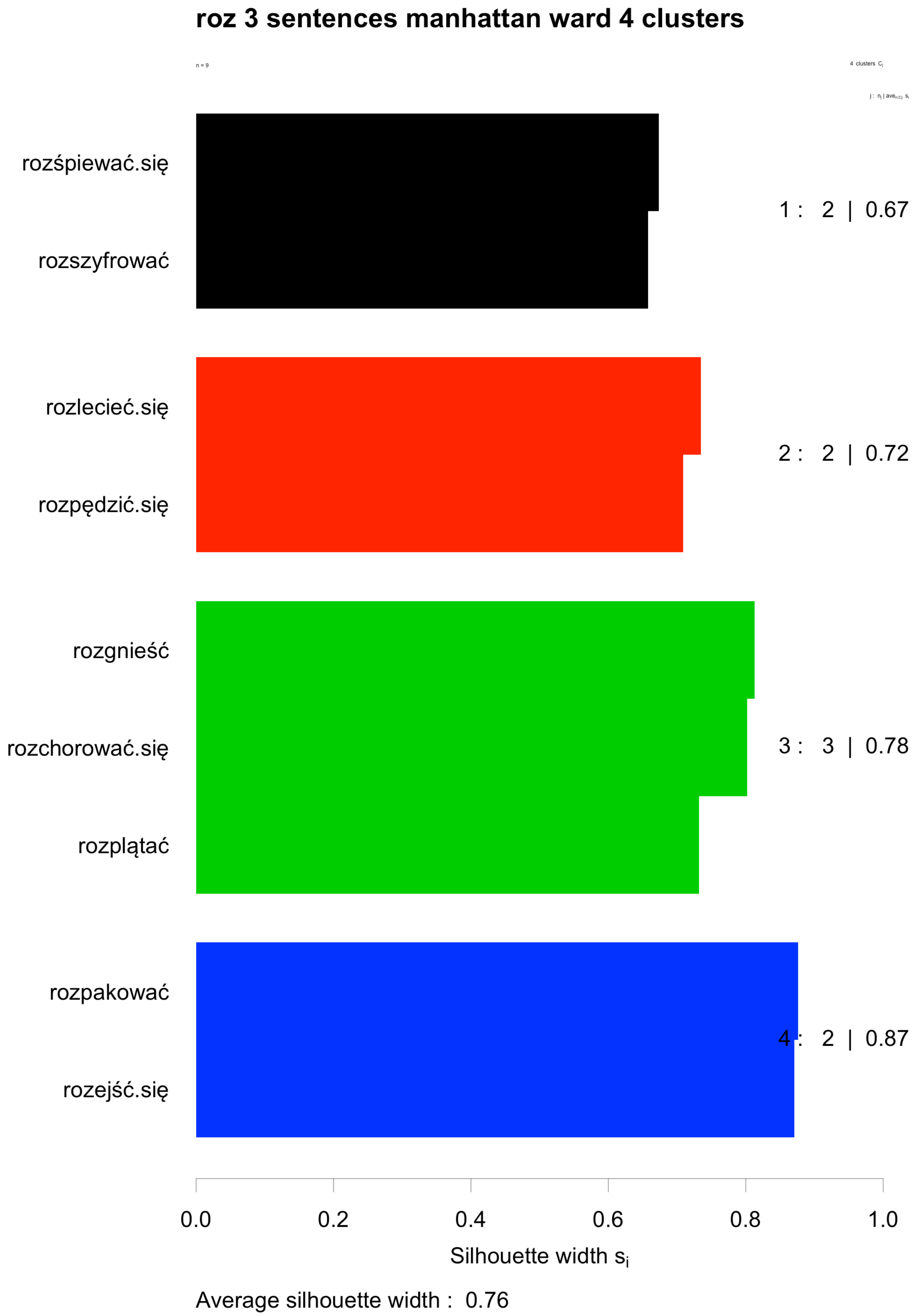


D. *przy*- 5 sentences per meaning




przy 5 sentences manhattan ward 5 clusters
przyłączyć.się
przypłynąć
przykucnąć
1 : 3 | 0.78
przystrzyc
przypasować
przymierzyć
przykręcić
przybiec
przyfrunąć
przykryć
2 : 7 | 0.50
przykleić
przynieść
przyklepać
3 : 3 | 0.76
przysłać
przysypać
przywalić
4 : 3 | 0.48
przypudrować
przyciemnić
przydeptać
przyhamować
5 : 4 | 0.74
0.0
0.2
0.4
0.6
0.8
1.0
Silhouette width s_i
Average silhouette width : 0.63

E. *roz-* 3 sentences per meaning



F. *roz-* 5 sentences per meaning