THESIS ABSTRACT

Kornélia Papp

Typicality, frequency, dimensionality
A multimodal analysis of the dimensional adjectives in Hungarian

Linguistics Graduate School
Director: Prof. Dr. Vilmos Bárdosi CSc.

Hungarian Linguistics Graduate Program
Director: Prof. Dr. Gábor Nagy Tolcsvai CMHAS

Members of the jury:
Chairman: Prof. Dr. Gábor Nagy Tolcsvai CMHAS
Opponents: Dr. Márta H. Varga
Dr. Nóra Kugler

Secretary: Dr. Judit Balogh
Further members: Dr. Klára Lengyel, Dr. Attila Mártonfi, and Dr. András Imrényi

Supervisor: Prof. Dr. Borbála Keszler DSc.

Budapest, 2013
Thesis topic and objectives

The concept of space is pivotal in the cognitive sciences in general, and in cognitive linguistics in particular. The reason for this is the assumption that, in the cognitive linguistics formulation of localism, spatial concepts also form the basis for some non-spatial ones (Lyons 1977). Space, spatiality and spatial perception have been research themes of topical linguistics since the 1980s (cf. Langacker 1982, Herskovits 1986). Although the examination of spatial adjectives would make for an obvious starting point in these studies, dimensional adjectives play a subordinate role in cognitive functional studies. To the best of my knowledge, this is the first such work regarding the Hungarian language. As Levinson (2004: 98) points out, most of the research concentrates on prepositions and postpositions: “There is a prevalent misleading presumption, which has grown out of the study of European languages and can be found in many textbooks, that spatial notions are encoded primarily in just one word class, namely prepositions or postpositions.”

This paper builds on previous cognitive functional works (Dirven and Taylor, 1988 Weyden, Schlieben-Lange, 1998, Rakhilina 2000 Athanasiadou 2001, Goy 2002, Taylor 2003, Vogel 2004, Tribushinina 2008) and deals with Hungarian dimensional adjectives, specifically with seven of them (nagy ‘large,’ kis ‘small,’ kicsi ‘small,’ hosszú ‘long,’ rövid ‘short,’ magas ‘high,’ alacsony ‘low’) modifying structures. Most previous analyses are based on linguistic intuition, collocation lists, questionnaires, or are corpus based. This paper explores the possibility of a multimodal approach, in which three linguistic methods are applied simultaneously to study dimensional adjectives: (i) a questionnaire-based recall survey, (ii) collocation analysis, and (iii) concept combination analysis.

The scope of research has been restricted to the description of concrete objects in the majority of works so far, and they have only dealt with spatial readings. Conversely, the subject of the present study encompasses abstract entities, and reveals the full usage network of the above mentioned dimensional adjectives. In doing so, it also sheds light on the relationship between prototypicality and frequency. The examination of the dimensional adjectives shows that, although a close relationship between the two concepts is likely, they cannot be equated.

The overall objective of the paper is to contribute to the functional description of dimensional adjectives. Following international research trends, but also offering new
methodological perspectives, it also aims to place the Hungarian language on the research landscape of dimensional adjectives.

The triumvirate of the title: typicality, frequency, and dimensionality, are concepts closely related to the three methods of this work. Defining typicality, more specifically prototypes, is the primary purpose of the questionnaire survey. The collocation analysis addresses the issue of frequency, and examines the relationship between frequency and prototypicality. Dimensionality is the third pillar and, at the same time, also the subject of the dissertation. Therefore, the third method, concept combination analysis, aims at examining the adjective and noun readings related to dimensionality. Three ultimate questions define the three main subject matters: (i) What are the prototypical instances of the studied dimensional adjectives? (ii) What are the most common collocations of these dimensional adjectives? (iii) What are the typical (dimensional as well as non-dimensional) readings of the dimensional adjectives?

According to Kamp and Partee (1995), the difference between color and dimensional adjectives is essentially that color adjectives are prototypical, while dimensional adjectives do not have a prototype. Geeraerts et al. (1994: 54) emphasize the typicality of the prototypes, and point out the gradual aspect of prototypicality: some concepts are more prototypical than others. This observation can be justified by the difference between color names and dimensional adjectives. The prototypicality of color names can be examined without a reference, whereas the concept of height or length is hard to fathom without one (Taylor 2003: 279). Tribushinina (2009) shows that prototypical height does not exist on its own, and can only be analyzed together with the referenced entities. Several studies have found similar prototypical concepts in numerous European languages (Dirven and Taylor 1988, Weyden and Schlieben-Lange 1998, Rakhilina 2000, Goy 2002, Vogel 2004, Tribushinina 2008). The present thesis is also tasked with the exploration of the prototypes associated with the adjectives of study in the Hungarian language. The results are not only important from a comparative linguistic point of view, but also provide a basis for examining the relationship between prototypicality and frequency.

The second main theme of the paper is closely related to the question of prototypicality. The concepts of frequency and prototypicality are often blurred in corpus linguistics (Schmid 2000, Stubbs 2004). One of the reasons for this is that the two terms overlap in many cases (cf. e.g. Geeraerts 1988, Aitchison 1998). However, several cognitive linguistic works point out that the most noticeable instances (see e.g. Radden’s [1992: 519-
the prototype) are not always the most frequent ones (Sinclair 1991, Gilquin 2006, Shortall 2007, Nordquist 2009). The aims of collocation analysis are, thus, on one hand, to map the most common combinations, and, on the other, to contribute to a better understanding of the link between prototypicality and frequency.

In functional models, both the word and the structural meaning actualize from a broader knowledge framework. The structural meaning does not correspond to the meaning of the individual elements or to the sum of the partial meanings, but is profiled from the composite meaning of the actually profiled elements (cf. Langacker 1987, Croft 2001, Tyler and Evans 2001, Taylor 2003, Evans 2004). The third major theme of the paper is to determine which adjective and noun readings are active in combinations with dimensional adjectives, such as in, e.g., tall man or low pay. Each combination is encoded manually taking its context into account, which makes it possible to determine the current readings accurately.

The structure of the dissertation

Following the introduction, the first part of the second chapter (2.1) is concerned with how the adjectives in general and the category of dimensional adjectives in particular appear in language studies, specifically in Hungarian. Section 2.2 briefly reviews the research on the subject, with special attention to the cognitive functional approach, and to the previous results that motivated the present work.

The third chapter sets the methodological foundations of the study (prototype model, concept combination model) and describes the approach used in the thesis. Section 3.1 covers the important aspect of the prototype definition, introduces the three classes of questions used in the questionnaire survey, and describes the respondents’ socio-cultural background. Section 3.2 details the design and limitations of collocation analysis. Thereafter, the concept integration model employed by this work and the basic principles of the manual encoding framework used in the concept combination analysis are specified in Section 3.3. The details of linguistic applications (programs, scripts) are also incorporated here, along with the introduction of models and methods.

The subsequent three chapters (4–6) present the research results of the three methods as described in the third chapter. The results of the questionnaire survey are shown and evaluated, and the prototypes presented in Chapter 4. The fourth chapter also discusses the
dominance within prototypes for adjective pairs (4.1.2, 4.2.2, and 4.3.2). The results of the collocation analysis are presented in Chapter 5. Frequency parameters are shown for the adjective pairs (5.1.1, 5.1.2, 5.1.3, 5.2.1, 5.2.2, 5.3.1, and 5.3.2), and this chapter deals with the issue of collocation dominance as well (5.1.4, 5.2.3, and 5.3.3). The results of the third method, i.e. concept combination analysis, are summarized in Chapter 6. This chapter presents the distribution of abstract and concrete readings in the context of discourse themes (6.1.1.1, 6.1.2.1, 6.1.3.1, 6.2.1.1, 6.2.2.1, 6.3.1.1, and 6.3.2.1). The frequency distribution of adjectival and noun readings associated with each adjective are plotted and discussed with the help of a few examples (6.1.1.2, 6.1.2.2, 6.1.3.2, 6.2.1.2, 6.2.2.2, 6.3.1.2, and 6.3.2.2). The sense distribution of adjectives expressing general size ('nagy ‘large,’ 'kis ‘small,’ 'kicsi ‘small’) are illustrated by cluster analysis (6.1.1.3, 6.1.2.3, and 6.1.3.3), whereas comparative sense distribution tables clarify the sense of adjectives expressing horizontal ('hosszú ‘long’ and 'rövid ‘short’) and vertical ('magas ‘high’and 'alacsony ‘low’) extent (6.2.3 and 6.3.3). The findings are summarized and further outlook is offered in Chapter 7.

Results

The present work yields the first dimensional adjective study in a multimodal framework in the history of Hungarian linguistic research. In addition to presenting antecedent Hungarian works, the relevant research, its methods and outcomes are also overviewed for the English, German, Italian, Russian, and Swedish languages. These studies also provide a methodological basis for the study. Compared to previous methodologies, the paper takes a novel approach in three aspects. First, it considers the adjectives in their context; second, the analysis goes beyond the concrete nouns to includes abstract ones; and third, it is not limited to the size meaning of dimensional adjectives. Thus, the multimodal analysis of the dimensional adjectives establishes a methodological advance over previous research with the advantage being that different methods reveal different aspects and features about the same linguistic structure.

The purpose of the questionnaire survey is to determine the typical antonyms and the prototypical readings (Weyden and Schlieben-Lange 1998). The analysis consists of three parts: (i) the definition of the typical antonym pair, (ii) association analysis, and (iii) acceptability analysis. The aim of the three question classes was to identify prototypical and less typical combinations.
The typical antonym pairs were clearly defined, and they confirmed the hypothesis. Respondents typically selected *kicsi* as the opposite of the adjective *nagy*, *rövid* as the opposite of the adjective *hosszú*, and *alacsony* as the antonym of *magas*. The same results were obtained vice versa.

The purpose of the association analysis is to identify typical instances and prototypical noun associates. I supplemented the method of Weyden and Schlieben-Lange (1998), Vogel (2004), and Tribushinina (2008) with association dominance analysis. The results indicate that prototypical instances are characteristically concrete entities in both the positive and the negative domain. Further analysis of frequent associations (association dominance analysis) shows that the more specific a dimensional adjective, the higher the ratio of nouns among the associates which occur with both of the adjectives of the antonym pair. This ratio is highest for the *hosszú–rövid* antonym pair at 71%, followed by the *magas–alacsony* pair at 51% and the lowest for the *nagy–kicsi* pair at only 28%. In parallel, the ratio of nouns used exclusively with only one member of the adjective pair is observed to be the highest for the *nagy–kicsi* pair, 37% and 35%, respectively. It holds for all three pairs that abstract concepts are more often encountered in the positive domain (*nagy, hosszú, and magas*) than in the negative one (*kicsi, rövid, and alacsony*). We have seen that both the *hosszú–rövid* and the *magas–alacsony* pair may convey vertical dimensionality.

Respondents were asked to assess the combinability of a given noun with a predetermined adjective on a scale of four in the third set of questions, similarly to the methodology used by Dirven and Taylor (1988), and Athanasiadou (2001). The results suggest that the concrete–abstract dimension does not alone affect acceptability: acceptable and unacceptable combinations are found both with abstract and concrete entities. The method also shows that good acceptance with one member of the adjectival antonym pair does not necessarily mean that the other adjective gets a similar assessment.

The results related to association types are closely linked to the collocation and combination analyses. As Athanasiadou (2001) points out, the various types of diagram analyses offer an opportunity to study exclusive and common utilization domains. Langacker (1982: 75) emphasizes that semantic structures coordinate the combinatorial possibilities. In this way, both association and collocation analysis contribute to a more accurate picture of the semantic structure of a language construct.

The typicality indicators and frequencies are investigated by collocation analysis, which is one of the most common linguistics methods, applied in numerous fields (see e.g.
Barnbrook 1996, Stubbs 1996). The collocations of seven adjectives (nagy, kis, kicsi, hosszú, rövid, magas, and alacsony) have been analyzed. The co-occurrences were studied based on the complete set of adjective + noun combinations found in the Hungarian National Corpus (Magyar Nemzeti Szövegtár, MNSZ). The analysis was carried out by using the AntConc3.2 software.

The results of the frequency analysis presented in Chapter 5 support the hypothesis that prototypes do not necessarily correlate with the most common usage (cf. e.g. Arppe et al. 2010). This applies particularly to the dimensional adjectives describing general extent (nagy, kis). Among the relatively few prototypes of nagy that do not express inherent size, only one, ember ‘man,’ is to be found among the frequent collocations. Three prototypical nouns are among the frequent collocations of kis: ország ‘country,’ ember ‘man,’ and ház ‘house.’ The reason for this can be seen in the prototypes being conspicuous, readily recalled concepts, hence usually concrete objects, while the frequent collocations of nagy and kis are typically based on abstract nodes.

The adjective kicsi, however, displays a different collocational behavior. Kicsi is typically found with concrete nouns in its collocations, and its prototypes (e.g. ház ‘house,’ ember ‘man,’ ország ‘country,’ or lány ‘girl’) also occur frequently. These results suggest that the narrower the use of a dimensional adjective, the more likely it is for the prototypical instances to be in frequent usage.

This assumption is supported by the collocation analysis of the adjective pairs hosszú–rövid and magas–alacsony. These dimensional adjectives are relatively rarely used, just like kicsi. Additionally, the origin of the meanings network is not the general extent as it is for nagy–kicsi, but the horizontal (hosszú–rövid) or vertical (magas–alacsony) extent, which already restricts the usage domain. Consequently, prototypical nouns (e.g. haj ‘hair,’ ház ‘house,’ and ember ‘man’) appear more often in both the vertical and horizontal groups, which is also associated with an increased ratio of concrete nouns.

Thus, the collocation analysis also shows that the usage frequency of prototypes increases with narrowing usage pattern. This suggests that frequently used nouns can indeed join the set of prototypes. While the prototypes for the adjectives nagy and kis are only concrete nouns with specific size, prototypical abstract nouns signifying time (idő ‘time,’ élet ‘life,’ and szünet ‘pause’) and degree (ár price and színvonal level) appear with the adjectives hosszú, rövid, magas, and alacsony, and are usually featured at the top of the frequency lists.
An advantage of frequency analyses is that they give a quick overview of the adjective + noun combinations, and show which patterns are present in a given language. As such, they may be valuable in compiling language teaching tools, for instance. The introduction of the dominance parameter appears to be particularly useful, when the combination patterns of two adjectives with similar meanings (kis–kicsi) are compared, or when the meaning fields of antonyms are overviewed.

Not only does the dominance analysis of collocations define the dominant collocations, but it also helps select the collocations that are characteristic for both adjectives in the antonym pair. Its main value is that it specifies the adjective and the noun readings, which are typical for one or the other dimensional adjective. In the case of collocations with nagy, the DEGREE and SIGNIFICANCE adjectival readings have been found dominant. In contrast, the TIME reading is activated in the dominant collocations of the antonym kis.

The TIME adjectival reading clearly dominates the collocations of the adjectives hosszú and rövid, as well. The difference between the antonyms is not in the adjectival reading, but in the domain of the collocated noun. Hosszú collocates usually with noun concepts meaning periods of time (e.g. év ‘year,’ hónap ‘month,’ or perc ‘minute’), while the adjective rövid is dominant with entities describing COMMUNICATION (e.g. észrevétel ‘observation,’ összefoglaló ‘summary,’ or hír ‘news’).

The adjectives magas ‘high’ and alacsony ‘low’ typically profile the adjectival reading DEGREE. Here again, the main difference between the two adjectives lies with the dominant noun readings. No specific noun reading is found in the collocations with the adjective magas (e.g. láz ‘fever,’ kor ‘age,’ or hang ‘voice’), in contrast with alacsony, which collocates typically nouns from the property group (e.g. gázár ‘gas price,’ bér ‘wage,’ or fizetés ‘salary’).

It holds for all three adjectival pairs that binds with concrete nouns (such as hosszú/rövid haj ‘long/short hair,’ nagy/kis fa ‘large/small tree,’ magas/alacsony nő ‘high/low i.e. tall/short woman’) as well as adjectival SIZE readings are more characteristic for the non-dominant collocations. This can be explained by the shared SIZE reading of the dimensional adjectives, which is their common ground and connecting feature.

Reading or sense combination analysis aims to explore adjectival and noun concept types of the combinations, and thus to identify the typical (reading or sense) combination patterns. The concept integration method (Fauconnier and Turner 1996, 1998; Fauconnier 2004) forms the theoretical framework of the analysis, and the geometric model (Gärdenfors
2000) also contributed to the readings. While collocation analysis works with word occurrences, the concept combination analysis processes concept occurrences. Knowledge of the context makes it possible to disentangle polysemic readings (e.g. *hosszú út* 'long way' – TIME or SPACE), and provides detailed information on the conceptual structure of the combinations as well. The foundations of the concept combination analysis can be found in the work of Dirven and Taylor (1988), in which 11 groups are established for the nouns combining with the adjective *tall*, although all category groups refers to concrete entities.

The concept combination analysis is, similarly to the collocation analysis, based on the MNSZ corpus. Over 4000 randomly chosen, manually encoded adjective + noun combinations form the basis of the evaluation. The data set consists of the noun combinations of seven adjectives (*nagy*, *kis*, *kicsi*, *hosszú*, *rövid*, *magas*, and *alacsony*) in their original context. The evaluation was carried out with FileMaker Pro 11/12, the data were processed using the DropLingaCombiner Delphi application, and sorted and plotted with Microsoft Excel. The suitability of the categories used in the concept combination model was determined by Kappa (Cohen's $\kappa$) coefficient analysis (see 3.3.3.5).

The concept combination analysis showed that the prototypical SIZE adjectival sense of *nagy* ‘large’, is far from being most common. Two nouns groups, OBJECT and GROUP combine preferentially with the SIZE adjectival sense of *nagy*. The SIZE reading for GROUP is, however, different from the sense of general extent, as it does not refer to the physical extent of an entity, but describes the group size of e.g. party or band. Thus, OBJECT can be considered the only noun reading with a frequent SIZE sense of *nagy*. DEGREE - and SIGNIFICANCE-centered adjectival readings are more frequent. The DEGREE sense coincides with three noun readings: EMOTION, POSSESSION, and ATTRIBUTE. Contrary to the hypothesis, PERSON noun readings combine preferentially with the SIGNIFICANCE sense. EVENT and COMMUNICATION readings are also SIGNIFICANCE-centered. Although it is a less typical member, the ACTION reading also belongs here. NEUTRAL readings (COGNITION, PHENOMENON, and STATE) are also often encountered with the adjective *nagy*.

Compared with *nagy*, the adjective *kis*, not only features more rarely in abstract combinations, but it also profiles the SIZE sense more often. In fact, SIZE is the most frequent sense of *kis* ‘small’, attracting the most noun readings. The five noun readings centered around it include OBJECT, PLACE, THING, GROUP and ANIMAL. As pointed out in collocation analysis, adjectival senses are not typical in combinations with *nagy*, the TIME sense occurs regularly with the adjective *kis*. The observation is further supported by concept combination
analysis. The TIME sense was uncharacteristic with nagy, but it occurs 20 times next to kis. The EVENT noun sense is the most frequent among SIGNIFICANCE-centered ones, but the ACTION and COGNITION senses, grouped as NEUTRALs, are close to this node as well. DEGREE-centered are the ATTRIBUTE, POSSESSION, and EMOTION reading categories. The FOOD, PERSON, COGNITION, and ACTION noun senses are NEUTRAL.

The TIME adjectival sense is common with the adjective kicsi, as well, which attracts the TIME noun sense. The number of occurrences of the SIGNIFICANCE sense is low among the combinations of kicsi, only one noun reading, ACTION, belongs here. The DEGREE group counted among the large ones in the case of kis. Quite the contrary is true for kicsi, which only connects to one noun reading, POSSESSION. It is striking that, in contrast to the clusters of nagy and kis, kicsi only binds to one category strongly, namely to the SIZE sense. The noun readings connect loosely to the other adjectival senses of kicsi. Among the nine noun readings, five belong to the SIZE sense: OBJECT, GROUP, PLACE, ANIMAL, and BODY. The PERSON noun reading behaves similarly in the case of kicsi as in kis, i.e. it is always positioned between the SIZE and YOUTH senses. However, while it was a relatively infrequent sense among the combinations of kis (26 occurrences, occ.), it is quite common with the adjective kicsi (134 occ.).

In contrast to the adjectives expressing general extent (nagy, kis, and kicsi), the adjectives hosszú and rövid are typically homogeneous, i.e. they are not linked to more than one adjective sense. Prototypically, hosszú and rövid express horizontal extent, but they both occur generally with the TIME adjectival sense and less frequently with the SIZE reading. The adjective hosszú combines with more noun readings than the adjective rövid. Distinctive readings include GROUP, PROCESS, LOCATION, and BODY. The adjective rövid has only one exclusive noun reading, EVENT. In the shared noun senses of TIME, ACTION, and STATE, the two adjectives behave similarly in time-related adjectival readings. In contrast, major differences can be observed in the noun readings OBJECT, COMMUNICATION, linked to the SIZE adjectival sense. The OBJECT noun reading typically occurs with hosszú, and the COMMUNICATION reading is typical with rövid.

The adjectives magas and alacsony, prototypically expressing vertical extent, are also homogeneous when compared with adjectives expressing general extent. As observed for the hosszú and rövid adjectives, homogeneity leads to a decrease in the number of noun readings. The two antonymic adjectives do not combine with an equal number of nouns readings. While the adjective magas is common with 11 noun readings, alacsony only occurs with 7 noun
readings. Magas and alacsony frequently occur with the DEGREE adjectival sense, and more rarely with the SIZE, QUANTITY, and SIGNIFICANCE senses. Both adjectives feature prominently the POSSESSION (141 and 252 occ.) and the ATTRIBUTE (49 and 118 occ., respectively) noun readings. The combinations of this adjective pair are clearly dominated by these two DEGREE-centered noun readings. The adjective magas combines with more noun senses than alacsony, of which the distinctive readings are PHENOMENON, COGNITION, PLANT, and PERSON. These noun readings are rare in combinations with alacsony, and they do not belong among the most frequent noun readings with magas, either. There is a fundamental difference between the two adjectives from the point of view of adjectival senses, namely that magas profiles relatively frequently from the SIGNIFICANCE sense (86 occ.), which accounts for only 5 occurrences with alacsony. Furthermore, the distribution of the adjectival senses is more uniform for magas, whereas it is slanted towards the DEGREE sense for alacsony.

References

Magyar Nemzeti Szövegtár (MNSZ), http://corpus.ny tud.hu/mnsz/


**Own publications relevant to the topic**


A melléknévi módosítóról. Általános Nyelvészeti Tanulmányok 22, 431–454.
