

The Hungarian colour terms *piros* and *vörös*

A corpus and cognitive linguistic account*

Réka Benczes

Department of American Studies,
Eötvös Loránd University, Budapest
rbenczes@gmail.com

Erzsébet Tóth-Czifra

Cultural Linguistics Ph.D. Program,
Eötvös Loránd University, Budapest
toth.czifrae@gmail.com

Abstract: A number of studies have been written on the Hungarian colour terms *piros* and *vörös*, both denoting ‘red’, focusing on either one of the following questions: (1) disambiguating the meanings of the two terms; or (2) their status in Hungarian as basic colour terms. The present paper attempts to resolve these issues in one go by adopting a combined approach of corpus and cognitive linguistics. The paper makes the following three hypotheses: (1) as *vörös* had more time to undergo idiomatization, there will be significant differences and systematic trends between the type/token ratios of the two terms; (2) *piros* is a more generic term used for a larger and looser range of concepts, while *vörös* is associated with a more limited range of concepts; and (3) *piros* is mostly used in its primary, literal sense, while *vörös* is more inclined to be used in a figurative sense. After a thorough corpus and cognitive linguistic analysis of data extracted from the updated Hungarian National Corpus, the paper comes to the general conclusion that *vörös* is not a basic colour term of Hungarian.

Keywords: basic colour term; *piros*; *vörös*; corpus linguistics; cognitive linguistics; conceptual metonymy; Hungarian

1. Background

In their seminal study on universal colour terms, Berlin and Kay (1969, 35–36) consider Hungarian as a language with “ten basic categories exclusive of red and two basic terms for red”, which makes Hungarian in their

* We wish to thank Keith Allan, Gábor Kiss, Zoltán Kövecses and our two anonymous reviewers for their invaluable remarks and suggestions. We are also grateful to Bálint Sass for helping with the data collection.

view “unique” (*op.cit.*, 95) as compared to other languages.¹ Accordingly, Hungarian possesses the following basic colour terms: *fejér* [*sic*] ‘white’,² *fekete* ‘black’, *piros* ‘red₁’, *vörös* ‘red₂’, *zöld* ‘green’, *sárga* ‘yellow’, *kék* ‘blue’, *barna* ‘brown’, *lila* ‘purple’, *rózsaszín* ‘pink’, *narancs* ‘orange’ and *szürke* ‘gray’.³

Despite the “uniqueness” of Hungarian basic colour terms, not much academic research has been carried out on them. What does exist, however, falls generally under two research areas. The first of these is more theoretical in nature, and is concerned with the disambiguation of the meanings of *piros* and *vörös*; Hungarian literature has mostly focused on this strand. The problematic nature of the terms was made note of as early as 1874, in Czuczor and Fogarasi’s (1862–1874) dictionary of Hungarian. In the first academic paper on the subject, Csapodi (1899) makes note of the eclectic distribution of *piros* and *vörös*, and calls for a delineation (and conventionalization) of their meanings. According to Csapodi, *piros* should be used with bluish hues, while *vörös* with yellowish-brownish ones. The view that *piros* and *vörös* represent two different (although not easily distinguishable) hues was shared by Kenedy (1921) as well. Nevertheless, this opinion was called into question by Gárdonyi (1920, 85), who was more rather on the opinion that the distinction between *piros* and *vörös* cannot be treated on a “technical” level, as – on account of the inconsistency among lexicalized constructions⁴ with either *piros* or *vörös* – the difference between the two terms is more rather “emotional”. What this implies is

¹ This does not mean to say that there no other languages with twelve basic colour terms. Berlin and Kay (1969) mention Russian, where there are two basic colour terms for ‘blue’, *sinij* for ‘dark blue’ and *goluboy* for ‘light blue’.

² *Fejér* is a nonstandard variant of *fehér* (i.e., ‘white’).

³ Berlin and Kay (1969) reached this conclusion by interviewing one native Hungarian speaker, who is referred to as “Madarasz” (which is a relatively frequent Hungarian family name). According to Berlin and Kay, all informants – including, therefore, the Hungarian informant – were native speakers who resided in the San Francisco Bay area. Otherwise no further information was provided on the informants.

It should be mentioned here that *vörös* is not the only colour term which has an ambiguous status. For instance, *rózsaszín* ‘pink’ is a compound word (which defies one of Berlin and Kay’s categorizational properties, according to which basic colour terms are monolexic). See also *narancs*, whose generally used alternative form in Hungarian is *narancssárga* (i.e., ‘orange + yellow’). See Kiefer (2005, 13) for an elaboration.

⁴ Throughout the paper “lexicalization” will be used to refer to one subtype of this process, namely semantic lexicalization or idiomatization (see Bauer 1983, 55–59).

that the respective terms carry value-based judgments, according to which *piros* is used more rather with positive concepts, while *vörös* is reserved for negative ones.⁵

Although the question of *piros* and *vörös* was far from settled, the topic received little attention in academic debates in the following decades (or was simply discussed under more general works on colour terminology – see, e.g., Kicsi 1988 or Wierzbicka 1996).⁶ Note should be made, however, of recently published Hungarian monolingual and etymological dictionaries, which have also attempted to disambiguate between the two terms, though with limited success. This relative deficiency, however, can be explained by the fact that the two colour terms are usually defined relative to one another. Accordingly, Eőry (2007, 1264) defines *piros* as “lighter than *vörös*, brightly coloured”, while *vörös* is given as “a darker hue of *piros*” (*op.cit.*, 1776); Pusztai (2011, 1082) also derives *piros* from *vörös*. Nevertheless, Pusztai (*op.cit.*, 1485) defines *vörös* as the colour of blood, which is in accordance with the etymology of the colour term. Etymological dictionaries (see e.g., Benkő 1967–1984; Bárczi 1994; Zaicz 2004) unanimously agree on *vörös* being the older of the two terms, and derive it from the noun *vér* ‘blood’. In fact, the semantic proximity comes to the forefront with the adjective *véres* ‘bloody’; the regular vowel change between *vörös* and *véres* prompted Benkő (*op.cit.*, 1178) to derive the former from the latter. Consequently, the colour term *vörös* evolved from the most salient property of blood, its redness. In our view, this process can be best described via the conceptual metonymy SALIENT PROPERTY FOR CATEGORY, whereby the salient property of blood, its redness, became a colour category in itself (i.e., the colour of blood). These assumptions are supported by the fact that the colour term for ‘red’ evolved from the word for ‘blood’ in other Finno-Ugric languages as well (*ibid.*; on a comparison of Finno-Ugric colour terms, see Uusküla 2008). The primacy of *vörös* in the evolution of Hungarian (as opposed to *piros*), as well as its direct relation to bodily experience, imply an originally basic-level status in the language.

Piros, on the other hand, evolved slightly later than *vörös*, from the onomatopoeic *pir/per* stems (cf. *perzsel* ‘scorch’, *pörkököl* ‘roast’), which apparently recalled the crackling sound that fire made during roasting or burning (Benkő 1967–1984, 208). Originally, the word was used to describe

⁵ This view was also shared by Selényi (1948).

⁶ Not much help is offered with regard to the differentiation between the core meanings/associations of *piros* and *vörös* by Wierzbicka (1990; 1996) either, who essentially lists the same basic semantic components for both – fire and blood.

the reddish-brownish colouring that was the result of roasting;⁷ it was later extended to more reddish hues (*ibid.*). Benkő has also attempted to disambiguate the semantics of the two colour terms, and – in accordance with Gárdonyi (1920; see above) – states that the main difference between them is evaluative in nature: *piros* typically describes more pleasant impressions than *vörös*.

The second line of research with regard to Hungarian basic colour terms is more empirical in nature and is concerned with testing the validity of Berlin and Kay's (1969) original claim, that is, the basic colour term status of *vörös*. This question is all the more relevant in light of the fact that even Berlin and Kay (1969, 36) revealed their uncertainty with regard to the basic colour term status of *vörös* (and, consequently, the existence of twelve basic colour terms in Hungarian), and raised the possibility that *piros* might be a “more basic” (*ibid.*) category than *vörös*. More specifically, *piros* might mark two categories, one in which *vörös* is already included (see figure 1), and one which contrasts with *vörös*. Therefore, depending on the context, *piros* can have two senses, ‘red’ and ‘lighter red’, in which case *vörös* can only be regarded as a secondary colour term (i.e., non-basic).

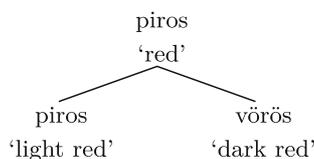


Figure 1: Semantic relationships of *piros* and *vörös* (after Berlin & Kay 1969, 36)

Berlin and Kay (1969), however, did not further elaborate on the issue, and additional testing of this possibility was outside the scope of their study. However, some researchers – both in Hungary and abroad – have tried to resolve the question by various means. Mention should be made of Kiss and Forbes (2001), who went through the entries of *piros* and *vörös* in Hungarian monolingual and Hungarian–English, Hungarian–German and Hungarian–French dictionaries in order to gather all the lexicalized collocations with *piros* and *vörös*, and also asked 98 informants to name typically *piros* and typically *vörös* objects. According to their results, although there are a limited number of entities which are associated with either one of the

⁷ Which can thus be described as an EFFECT FOR CAUSE metonymy in cognitive linguistic terms.

colour terms,⁸ there are plenty others where both colour terms are equally valid. Therefore, since there were no systematic differences in the use of *piros* and *vörös*, Kiss and Forbes came to the conclusion that both colour terms are basic in Hungarian. Nevertheless, it should be pointed out that Kiss and Forbes also conducted a colour-naming task, in which informants had to list colour terms. The results speak for themselves and question the final conclusion of Kiss and Forbes: 92% of the informants listed *piros* (which was listed on average as the third colour term), while only 46% listed *vörös* (which had an average ranking of 13).

In a follow-up to Kiss and Forbes' (2001) study, Kiss (2004) conducted a corpus-based research on the occurrence of *piros* and *vörös* in the Hungarian National Corpus, which indicated that there is indeed a group of limited entities that are typically associated with either one of the colour terms *piros* and *vörös*, and there is also a further group that can be associated with both terms. Kiss (*op.cit.*, 164) concludes that Hungarian is definitely unique in its use of the two colour terms – therefore implicitly confirming Berlin and Kay's (1969) original assumptions.

Nevertheless, not all studies support such views. Both Maclaury et al. (1997) and Uusküla and Sutrop (2007) assume that Hungarian has only eleven basic colour terms, and *vörös* is not included among them.⁹ A further common feature of the two studies – apart from their final conclusions – is that they base their research on a significantly larger number of informants than Berlin and Kay (1969), which gives considerable support to their claims. Essentially, Maclaury et al. (1997) replicated Berlin and Kay's (1969) original study (the colour naming task) with ca. ten subjects, while Uusküla and Sutrop (2007) performed a list task (i.e., requesting subjects to list as many colour terms as possible) and a colour-naming task with 40 informants.

At the end of their paper, Maclaury et al. (1997) also provide some indication of the connotations of *piros* and *vörös* that are of especial interest

⁸ *piros*: *cseresznye* 'sweet cherry', *száj* 'mouth', *eper* 'strawberry', (*közlekedési lámpa* 'traffic lights', *Mikulás* 'Father Christmas', *pont* 'period' (i.e., a sign of appreciation for something); *vörös*: *haj* 'hair', *róka* 'fox', *csillag* 'star', *meggy* 'sour cherry'.

⁹ Maclaury et al. (1997, 75) are in fact of the opinion that eleven basic colour terms – in any language – is the absolute upper limit: "From what we have seen of color naming in Hungarian, French, Russian, Japanese, and other languages, there is no reason to assert that the number of basic color terms exceeds eleven in any particular case. Differentiation of the color domain must at some point reach its ceiling, and the likely limit is eleven basic units."

to the current paper as well. According to the questionnaires performed on ca. 10 subjects, the meaning of *vörös* is “darker” and “more sinister” (*op.cit.*, 76–77), and brings to mind concepts such as passion, love and lust in the minds of the speakers, as well as communism and bloody revolts. At the same time, *vörös* is also associated with “warmth, depth and softness” (*ibid.*, 77). *Piros*, however, “was harder for [the subjects] to talk about, as it reminded them of less” (*idem.*). Nevertheless, one particular concept that *piros* can be connected to is health, as only *piros* occurs in idioms related to health.

2. Aims of the present study

In sum, previous research on the topic has suffered from a number of shortcomings. First, research has fundamentally focused on either one of the research questions connected to *piros* and *vörös*: (1) disambiguating the meanings of the two terms; or (2) their status in Hungarian as basic colour terms. No study has yet strived to look at the correlations between the two questions, even though they are undoubtedly intertwined. Second, previous results on the basicness of *vörös* are very ambiguous, and depend on the applied methodology – leaving the issue essentially unresolved. In colour listing tasks, *piros* scored significantly better than *vörös*, questioning the basicness of the latter. However, the corpus linguistic data of Kiss (2004), where *vörös* was the sixth most common colour term in the corpus, immediately following *piros*, seem to indicate the basicness of both terms. Third, the conclusions of the previous researches were based on either linguistic introspection on account of the authors (see especially the early papers by Hungarian researchers) or a limited number of informants – ranging from one subject, as in the case of Berlin and Kay (1969), to forty subjects, as in the case of Uusküla and Sutrop (2007) – who had to perform either colour naming and/or colour listing tasks. Although Kiss and Forbes (2001) worked with a substantially larger group of informants (they asked 98 informants to name typically *piros* and *vörös* entities), they did not provide any methodological background on how and with what means they conducted their research. The only corpus-driven study up to now has been carried out by Kiss (2004), who, however, included other colour terms in his research as well, and did not concentrate on the semantic aspects.¹⁰

¹⁰ Further note should be made of Kiss & Forbes (2001) and Barratt & Kontra (1996), both of whom have relied on dictionary sources – and can, therefore, be considered as precursors to corpus-driven analyses.

The aim of the present paper is to fill in the gaps that previous researches have left behind and finally resolve the question of the basic-level status of *vörös*. In order to do so, the paper will use a combined approach of both corpus and cognitive linguistics. First and foremost, by using the updated (2013 version) Hungarian National Corpus (henceforth HNC),¹¹ the paper will provide an in-depth, corpus-based analysis of present-day Hungarian collocations with *piros* and *vörös*. No previous research has relied on such a broad spectrum of data. Second, by drawing on the corpus-driven data, the paper will attempt to analyse the concepts associated with the two terms, as well as the semantic relations that hold among these concepts, within a cognitive linguistic framework by drawing on conceptual metaphor and metonymy. While colour terms seem to be an ideal topic to research within a cognitive linguistic framework (given the vast number of figurative senses associated with them), not much has been published on them from this point of view. Studies have usually focused on how the meaning of a particular colour term is modified by the head noun that it is attached to (e.g., Sweetser 1999, 139). An exception to this trend is Szlávi's (2008) pioneering study on the conceptualizations of *black* and *white*, which demonstrated that the figurative senses associated with *black* and *white* – as they appear in AN compounds – are linked to two core senses via a complex semantic network held together via conceptual metaphors and metonymies.

The combination of corpus and cognitive linguistic analysis is in line with one of the major assumptions within cognitive linguistics, namely that the knowledge of a language emerges from language use, and is not based on pre-determined either-or categories. This focus on usage-based language naturally calls for qualitative, i.e., corpora-based analyses (Gries 2006; Wong 2012). By drawing on this combined approach, the paper makes the following three major assumptions:

1. If, from an etymological point of view, *vörös* is indeed the older of the two terms, and had consequently more time to undergo idiomatization, it is hypothesized that there will be differences between the type/token ratios of the two terms. More specifically, it is expected that compared to *piros*, the type frequency of *vörös* will be lower, while its token frequency will be higher, implying that the use

¹¹ The updated HNC allows for more precise and complex searches. Furthermore, as compared to the 2004 search conducted by Kiss, there is a larger corpus at hand: in 2005, the corpus was extended with a substantial amount of literary texts and dialectal sources.

of *vörös* is restricted to a limited number of frequently used expressions. As for *piros*, it is expected that there will be a higher type frequency, though with smaller token frequency than *vörös* – indicating that *piros* is used in a wider range of expressions (both fixed and novel) with varying degrees of frequency.

2. Accordingly, it is also hypothesized that, of the two terms, *piros* is a more neutral and generic term used for a larger and looser range of concepts, while *vörös* is associated with a more limited, and semantically more restricted, range of concepts. Consequently, *piros* will be mostly used in its primary, literal sense, to refer to the external colour of an entity, while *vörös* will be less inclined to be used primarily in a literal sense to describe the external appearance of an entity, and will therefore have a predilection to be used in a figurative sense. These figurative senses will be mostly metonymically related (via the SALIENT PROPERTY FOR THE CATEGORY metonymy) to the concept of blood, which serves as the etymological basis of *vörös*.
3. If *piros* is indeed a more general and neutral term, while *vörös* is more extended and figurative in its meaning and usage, then it is also hypothesized that *vörös* is not a basic colour term in Hungarian.

The structure of the paper is as follows. After a general introduction to the topic, the third section discusses the corpus linguistic results, while the fourth section examines the data from a cognitive linguistic perspective. The fifth section synthesizes and concludes the partial results of sections three and four with an attempt to resolve the issue of basicness.

3. Corpus linguistic results

3.1. General frequencies of *piros* and *vörös*

One of the major sources of inspiration for the present paper has been the study of Kiss (2004), who has done pioneering work in resolving the issue of basicness with regard to *piros* and *vörös* from a corpus linguistic perspective. Therefore, following Kiss (2004), we first checked the occurrence of *piros* and *vörös* as independently occurring words (both inflected or uninflected forms) in the updated HNC, as compared to the other basic colour terms of Hungarian. Table 1 summarizes the results.

What can be immediately seen from table 1 is that both *piros* and *vörös* scored relatively high among the basic colour terms, ranking as the

Table 1: Frequency of Hungarian basic colour terms, in decreasing order

Ranking	Colour term in Hungarian	English equivalent	Number of hits in the HNC
1	fekete	black	40,036
2	fehér	white	37,357
3	zöld	green	24,414
4	kék	blue	17,351
5	piros	red	16,857
6	sárga	yellow	13,455
7	vörös	red	12,760
8	barna	brown	9,550
9	szürke	grey	8,139
10	lila	purple	3,981
11	rózsaszín	pink	2,411
12	narancssárga	orange	564

fifth and the seventh most common colour term in the corpus, respectively.¹² At first sight such data might give the impression that both *piros* and *vörös* are basic level terms in Hungarian, in line with Forbes's (1979) reasoning, who states that there is a tight correlation between basicness and frequency (i.e., if a colour term is basic in a given language, then this will reflect in its overall frequency). However, this assumption should be treated with caution. According to our results, both dark blue (*sötétkék*) and crimson (*bíbor*), which are not considered as basic colour terms in Hungarian, scored higher than orange (*narancssárga*), which, nevertheless, is regarded as a basic colour term.¹³ (*Sötétkék* had 957 hits, while

¹² Although the structure and size of the 2004 corpus in the Kiss (2004) study was dissimilar to the present corpus, it is interesting to note that the difference in frequency between *piros* and *vörös* was negligible in the 2004 corpus: *piros* had 10,170 hits, while *vörös* had 9,651 (i.e., the difference between them was a mere 5%). In our search, however, the difference between the two colour terms was more substantial: while *piros* had 16,857 hits, *vörös* had only 12,760 – i.e., the difference here amounted to 32%. What these results indicate is that a new corpus-based research on *piros* and *vörös*, based on the updated HNC, is indeed justified.

¹³ Berlin and Kay (1969) originally used the term *narancs*; however, we have decided to use *narancssárga* 'orange + yellow' in the corpus search instead – we felt that the latter is a closer translation of *orange*. (Note that other researchers, such as Kiss 2004, have also used *narancssárga* and not *narancs*). Our decision is justified by

bíbor had 758 hits). What this entails is that frequency in language might not necessarily correlate with basicness in every single case. Therefore, the only solid conclusion that can be drawn from the data at this point is that both *piros* and *vörös* show up frequently in language as compared to other colour terms.¹⁴ The next question that needs to be settled, therefore, is what sort of concepts does *piros* and *vörös* appear with and how often – i.e., what are the type and token frequencies of these two colour terms?

3.2. Type and token frequencies of *piros* and *vörös*

3.2.1. Raw data analysis

As a further step, we have conducted a search for all the the *piros* + N and *vörös* + N sequences in the updated HNC. Since we were interested in the raw data for the occurrence of *piros* and *vörös* with nouns, we did not make a distinction between phrases and compounds. What this means in practice is that in our first data analysis we have considered all *piros* + N and *vörös* + N sequences, regardless of their orthography.¹⁵ This is not

Hungarian monolingual dictionaries (see Eöry 2007, Pusztai 2011), which typically treat *narancs* and *narancssárga* under separate entries. The primary sense of the latter is the colour, while in the case of the former it is the fruit (with the colour sense appearing much lower in the entry). Note that Eöry (2007, 1229) defines the colour sense of *narancs* with the help of *narancssárga*, i.e., as *narancssárga szín* ('*narancssárga* colour'). According to Pusztai (2011, 965), the use of *narancs* as a colour term is rare.

¹⁴ Interestingly, *vörös* ranked much lower than *piros* in word association tests, calling into further doubt the basic-level status of the former. According to the September 2013 data from *agykapocs.hu*, the largest word association database in Hungary, *vörös* was given only as the tenth most common answer to the prompt *szín* 'colour', which amounted to 1.1% of the total responses (Σ : 1,026). The most common association with the word *szín* was *kék* 'blue', with 24.5% of the total responses, followed by *piros* with 13.3%, *zöld* 'green' with 9.4%, *sárga* 'yellow' and *fekete* 'black' both with 4.9%, *lila* 'purple' with 4.5%, *szivárvány* 'rainbow' with 4.4%, *fehér* 'white' with 2.2%, and *rózsaszín* 'pink' with 2.1%. Therefore, *piros* was associated ten times more frequently with the word "colour" than *vörös*. (We wish to thank László Kovács for supplying us with these data.)

¹⁵ In Hungarian, there is a distinction on the orthographic and semantic level among combinations. Generally, compounds – which have undergone a certain degree of lexicalisation – are written without a space between the constituents, while phrases do have a space between the constituents. In other words, a sequence of two nouns can have very different meanings depending on their orthography: *drágakő*, a lexicalised compound, is non-compositional and carries the meaning of 'gemstone',

to say that we wish to discard the significance of orthography and the semantic distinction that it indicates in almost all cases – the next section (3.3.) will in fact look at this question in more detail.

The rationale behind our decision to include *all* data in the raw analysis was the observation that language users did not necessarily follow orthographic rules systematically, and, therefore, a concept such as red wine appeared both as *vörös bor* ‘red wine’ and *vörösbör* ‘redwine’ in the HNC. If we had excluded “compounds” from the hits and focused only on “phrases”, then we would have lost a significant amount of data. This decision necessarily entails that some sequences that do appear in both forms (such as *vöröskereszt* ‘redcross’ and *vörös kereszt* ‘red cross’, and which we have grouped into a single category in the raw analysis) were understood with different senses in the corpus, but this possibility applied to only a minimal amount of data. We did, however, discard proper nouns (including geographical proper nouns and nicknames used in web-based forums) from the data. Only those sequences were considered that had a minimum token frequency of 10 hits. Appendix 1 provides a summary of the results.

What can be immediately seen from the data is that there are considerably more nouns associated with *piros* than with *vörös*: the former has a type frequency of 126, while the latter has a type frequency of 96 – meaning that *piros* is associated with 24% more nouns than *vörös*. When looking at the type–token distribution, it can also be seen that *piros* has a higher number of types with lower token frequencies, while *vörös* has more types with lower tokens. Table 2 summarizes the data.

Table 2: Number of *piros/vörös* types in ranges of frequencies

Range of token frequency	<i>piros</i> + N	<i>vörös</i> + N
10 to 19	63	40
20 to 29	27	16
30 to 59	23	12
60 to 99	8	14
100 to 199	4	8
200 to 399	2	3
400 and up	0	3

while the N + N sequence *drága kő* is compositional and its overall meaning is the sum of the meaning of its constituents, i.e., ‘expensive rock’. For a detailed discussion, see Laczkó & Mártonfi (2004, 98) and Lengyel (2000, 325).

The first column of table 2 indicates the range of token frequency, while the second and third columns show the number of *piros* and *vörös* types within that frequency range. As a next step, we have calculated the share (%) of *piros* and *vörös* types for each frequency range (each frequency range representing 100%), and plotted the percentages on a graph. The results are provided in figure 2.

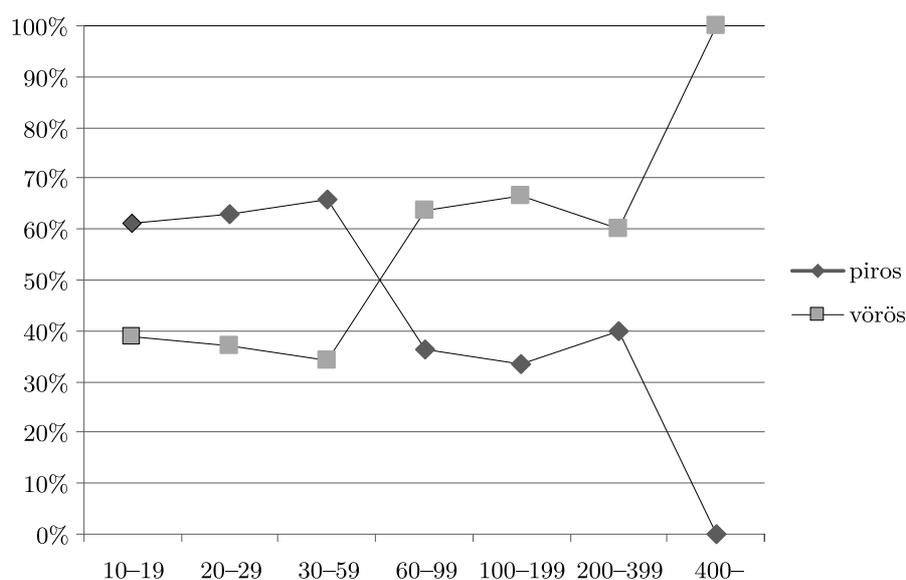


Figure 2: Percentage of *piros* and *vörös* types in ranges of frequencies

What is immediately visible from figure 2 is that *piros* starts off with a high number of types in the lower token frequency range as compared to *vörös*; then the number of types gradually decreases as the token frequencies increase. As for *vörös*, exactly the opposite trend can be observed – as compared to *piros* it has relatively few types with low token frequencies, but the number of types increases as the token frequencies also increase. The early dominance of *piros* changes in the 60 to 99 token frequency range, where it is finally surpassed by *vörös* (*vörös* has 14 types, while *piros* has only 8 types in this frequency range). As it can be seen from Appendix 1, there is substantial difference between *piros* and *vörös* with regard to the highest token frequencies as well. The highest token frequency of *piros* was 356 (for *lámpa* ‘light’), while the highest token frequency of *vörös* was 1,872 (for *kereszt* ‘cross’).

Originally (see section 2) we hypothesized that the type frequency of *vörös* would be lower, while its token frequency would be higher; at the same time we expected higher type frequencies with smaller token frequencies for *piros*. The data indicate that our first assumption on the type/token distribution of *piros* and *vörös* has been borne out. It seems that *piros* is indeed used in a wider range of expressions, with varying degrees of frequency, while the use of *vörös* is limited to a smaller number of more frequently used expressions. The explanation that we can provide for this phenomenon is that *vörös* – which is the etymologically older of the two terms – has undergone a certain degree of lexicalisation, and, therefore, the structures in which it is used are more fixed (even though these concatenations are often used in everyday language).¹⁶

This raises an interesting question – namely, the relationship between productivity and type/token frequency. Why is it that *vörös*, which has considerably higher token frequencies, is much less productive than *piros*? In Bybee's (2001, 119) view, productivity rests on type – and not token – frequency: “The number of existing items that a pattern applies to bears a direct relation to the probability that it will affect new items.” That is, it is more probable that *piros*, with its higher type frequency, will be used for the creation of novel structures than *vörös*. The reasons for this are twofold. First, higher type frequency correlates with greater analysability – i.e., language users are able to identify and generalize patterns more easily. Second, frequently occurring structures “achieve a certain autonomy from related forms” (*ibid.*, 125). What this means then is that frequently occurring structures with *vörös* are semantically lexicalized, thereby being less prone to be generalized to novel patterns. In sum, what we can conclude from the first data analysis is that *piros* shows up in more novel structures that have low token frequencies, while *vörös* tends to appear with more lexicalized and fixed expressions with high token frequencies.

At this point it is worth taking a look at the types that showed up with both colour terms, i.e., those that occurred with both *piros* and *vörös*, in the raw data. Around one-third of all the types that occurred with *vörös* also occurred with *piros*, while only about a fourth of all the types that

¹⁶ On the basis of the earlier literature on the subject (i.e., Csapodi 1899; Gárdonyi 1920; Kenedy 1921, etc.) it can be hypothesized that the distribution of the two colour terms (from the mid-19th century onwards) has undergone a number of changes; nevertheless, these changes have not essentially affected the main trends of the lexicalization process. Needless to say, a full semantic analysis of *piros* and *vörös* would also call for a diachronic analysis; this, however, is outside the scope of the present study, which aims to focus on synchronic patterns exclusively.

occurred with *piros* also showed up with *vörös*. This discrepancy can be easily explained with the higher type frequency of the latter. Tables 3 and 4 list the results.

Table 3: Nouns occurring with both *piros* and *vörös*; 10 most frequent types with *vörös*

Type	Number of tokens with <i>vörös</i>	Number of tokens with <i>piros</i>
kereszt ‘cross’	1,872	15
bor ‘wine’	1,157	34
csillag ‘star’	329	16
zászló ‘flag’	222	102
fény ‘light’	106	45
szín ‘color’	109	113
rózsa ‘rose’	88	95
folt ‘blot’	82	85
szőnyeg ‘carpet’	71	40
arc ‘face’	61	60

Table 4: Nouns occurring with both *piros* and *vörös*; 10 most frequent types with *piros*

Type	Number of tokens with <i>piros</i>	Number of tokens with <i>vörös</i>
lámpa ‘lamp/light’	356	19
pont ‘point’	228	12
szín ‘color’	113	109
zászló ‘flag’	102	222
rózsa ‘rose’	95	88
folt ‘blot’	85	82
virág ‘flower’	82	20
ruha ‘dress’	64	20
arc ‘face’	60	61
szalag ‘ribbon’	53	18

Without going into a qualitative analysis of the results at this point, it is quite apparent that in those cases where *vörös* is the dominating colour term – the first five of these being in descending order: *kereszt* ‘cross’, *bor* ‘wine’, *csillag* ‘star’, *zászló* ‘flag’ and *fény* ‘light’ – there is a much lower

token frequency with *piros* (cf.: *kereszt*: 15 vs. 1,872 tokens; *bor*: 34 vs. 1,157 tokens; *csillag*: 16 vs. 329 tokens; *zászló*: 102 vs. 222; *fény*: 45 vs. 106 tokens).¹⁷ At the same time, nouns that are dominated by *piros* – the first five of these being: *lámpa* ‘lamp/light’, *pont* ‘~point’, *szín* ‘colour’, *rózsa* ‘rose’ and *folt* ‘stain’ –, the difference is much less substantial (cf.: *lámpa*: 356 vs. 19 tokens; *pont*: 228 vs. 12 tokens; *szín* 113 vs. 109 tokens; *rózsa* 95 vs. 88 tokens; *folt* 85 vs. 82 tokens). It is interesting to note that apart from *lámpa* and *pont*, the frequencies here are relatively even (this is especially so in the case of *szín*, where the frequencies are near-identical). All in all, these results corroborate our assumption that *vörös* is more restricted to lexicalized structures, while *piros* is more flexible in its use.

3.2.2. Phrases vs. compounds

While an analysis of the data with respect to their morphological structure – i.e., whether they are phrases or words (compounds) – is a relevant one with respect to the issue of lexicalization and hence the basic colour-term status of *vörös*, it is, nevertheless, a tricky one. As already alluded to in section 3.2.1, orthography is far from consistent in the corpus. Only a qualitative, semantic analysis would provide absolutely reliable results, i.e., a detailed analysis of every single context within which the *piros* + N and *vörös* + N sequences occurred. Such an analysis, however, is outside the scope of the present study. What we can offer, nonetheless, are a couple of tentative conclusions drawn on generalized data, with the disclaimer that distortions – due to inconsistency in the corpus – may apply. Therefore, in the following, all *piros*N and *vörös*N sequences will be considered as compounds, and all *piros* + N and *vörös* + N sequences as phrases.

First, we calculated the ratio of compounds vs. phrases among the overall type frequency. In sum, there were 125 *piros* + N sequences and 3 *piros*N sequences; with regard to *vörös*, the numbers were 85 and 21 respectively. What the numbers indicate is that 98% of the *piros* sequences were phrases, while this number was only 80% with respect to *vörös* – in the latter case, every fifth *vörös* sequence type was a compound. The three *piros* + N sequences were the following (brackets indicate token number): *piros**paprika* ‘paprika’ (171), *piros**pont* ‘brownie points’ (57), and *piros**betli* (~an expression in the card game *ulti* when no points are expected in a particular round; 33). As for the *vörös*N sequences, the five most fre-

¹⁷ The first number indicates the token frequency of *piros*, while the second number represents the token frequency of *vörös*. This trend was also alluded to by Kiss (2004, 163).

quent types were *vöröskereszt* ‘red cross’ (1,822), *vörösbor* ‘red wine’ (922), *vöröshagyma* ‘onion’ (398), *vörösréz* ‘copper’ (138) and *vörösbegy* ‘robin redbreast’ (94). The numbers are even more telling if one looks at these ratios with respect to the tokens. All in all, 6% of all *piros* sequences were compounds, while 52% of all *vörös* sequences were compounds.¹⁸ These data further support the assumption that *vörös* is used in more lexicalized sequences (i.e., compounds), with fixed semantic links between the constituents, as opposed to the *piros* sequences, where the relationship between the constituents is less fixed and hence semantically less lexicalized. Such a state of affairs indicates a wider semantic scope for *piros*, allowing for a more general use of the colour term, which, necessarily, enhances the appearance of further *piros* + N patterns.

4. Semantic analysis of *piros* and *vörös*

What the previous results have indicated is that *piros* is used in a wider range of expressions, with varying degrees of frequency, while *vörös* is reserved for a smaller range of more frequently used, fixed expressions.¹⁹ We have argued that this phenomenon can be explained by the more lexicalized status of the *vörös* sequences; as *vörös* is the etymologically older of the two colour terms, it has undergone a certain degree of lexicalisation, and, therefore, the structures in which it is used are more fixed. This claim has been supported by the high ratio of compound-like forms among the *vörös* sequences, as opposed to the relatively low ratio of compound-like forms among the *piros* sequences.

¹⁸ In numbers: there were all together 4,187 instances of *piros* sequences, and 261 of these were compounds. With regard to *vörös*, the total number of instances was 7,924; the number of compounds was 4,124.

¹⁹ In her analysis of Hungarian colour terms, Papp (2012, 156) comes to similar conclusions. Papp examined the distribution of *piros* and *vörös* in the subcorpora of the HNC. Accordingly, the highest occurrence of *piros* (50.22 hits per million words) was in the personal subcorpus, which is the closest to colloquial usage. Here the occurrence of *vörös* was much lower, with only 25.56 hits per million words. However, the trend shifted in the more specialized corpora, esp. in the media subcorpus (*piros*: 21.11 hits per million words; *vörös*: 24.71 hits per million words) and the scientific subcorpus (*piros*: 32.70 hits per million words; *vörös*: 43.83 per million words). In Papp’s view, the reason for the preponderance of *vörös* in the specialized corpora can be attributed to the following: (1) in the media, *vörös* is often associated in a political sense; and (2) in optics and biology, *vörös* is the generally used term, as opposed to *piros*.

These findings have interesting implications for the semantics of *piros* and *vörös*, as provided by the HNC data. Two main assumptions can be drawn. First, due to the higher type frequency but lower token frequency *piros* will be mostly used in its primary, literal sense, to refer to the external colour of an entity. Second, due to the lower type frequency but higher token frequency *vörös* will be less inclined to be used primarily in a literal sense to describe the external appearance of an entity, and will therefore have a predilection to be used in a figurative sense. It is also hypothesized here that the figurative senses of *vörös* are not random, but are related to one another via conceptual metonymy.

As the data we are dealing with is quite substantial, and it is outside the scope and the limitations of the present research to go through every single hit (i.e., token) of each *piros* and *vörös* sequence in order to check the exact reference of the colour term in question, we have decided to test our hypotheses on the basis of Appendix 1. Needless to say, we are aware of the fact that the Appendix overgeneralizes and does not allow for detailed analyses. This, however, is not our aim. What we wish to do in the present paper is to demonstrate the possible correlations among the overall distribution of *piros* and *vörös* sequences on the one hand, and the general trends that can be detected in the semantic scope of these colour terms on the other. We hope that the present conclusions will prompt further – and more detailed – analyses into the exact nature of the semantics of these colour terms.

4.1. *Piros*

4.1.1. *Piros* as an external property

Our first assumption regarding the scope of the meaning of the *piros* + N sequences has been borne out. Generally speaking, the vast majority of the *piros* types refer to colour, more specifically to the external colour of the entity they modify. The largest groups (both amounting to around one-fifth of the total number of types) included clothing items (such as *ruha* ‘dress’, *gomb* ‘button’, *sál* ‘scarf’, *szoknya* ‘skirt’, *kabát* ‘coat’, *topánka* ‘slippers’, *kalap* ‘hat’, etc.) and man-made objects (such as *bicikli* ‘bicycle’, *filctoll* ‘marker’, *ceruza* ‘pencil’, *fotel* ‘armchair’, *labda* ‘ball’, *pad* ‘bench’, *gyertya* ‘candle’, etc.). *Piros* also showed up before a noun denoting a plant, as in the case of *rózsa* ‘rose’, *alma* ‘apple’, *szegfű* ‘carnation’, *tulipán* ‘tulip’, *virág* ‘flower’, *bogyó* ‘berry’, *paradicsom* ‘tomato’, *pipacs* ‘poppy’, *gyümölcs* ‘fruit’, *cseresznye* ‘cherry’, *ribizli* ‘redcurrant’ and *szírom* ‘petal’. There are two intriguing aspects about these examples. First, they all refer to flow-

ers and fruits exclusively, which are, at the same time, very common in Hungary. Second, when combined with *piros*, neither one of them refers to a particular plant species (i.e., the sequences are not lexicalized); the modifier simply describes their colour. However, in the case of fruits (i.e., *alma*, *bogyó*, *paradicsom*, *cseresznye* and *ribizli*), the use of *piros* also entails their ripeness, i.e., they are ready for consumption. In a similar vein, when used with any of the flowers (*rózsa*, *szegfű*, *tulipán*, *pipacs*), there is an implication that the flower is in full bloom. Therefore, on the basis of these examples, it can be inferred that the colour *piros* stands for ripeness and maturity.

Piros also cropped up frequently in the corpus with body parts: *arc* ‘face’, *szem* ‘eyes’, *vér* ‘blood’, *száj* ‘mouth’, *orr* ‘nose’, *fej* ‘head’, *bőr* ‘skin’, *ajak* ‘lips’, *nyelv* ‘tongue’, *tenyér* ‘palm’. What is interesting about the body part examples (especially *arc*, *bőr*, *ajak*, *nyelv* or even *orr*) is that in such cases “redness” can be considered as rosiness, i.e., a sign of health (the association of *piros* with health has been alluded to by Maclaury et al. 1997 as well). This association might be based on the fact that *piros* is associated metonymically with blood (i.e., the colour of blood – the defining property of the category – standing for blood, the category itself), and being full of blood is a sign of good health. As can be seen from Appendix 1, *vér* ‘blood’ showed up only with *piros*, and not *vörös*, even though the etymological basis of *vörös* would have suggested otherwise. (We will turn back to this question in section 4.2.1.) The occurrence of *piros* with *szem* has been noted by Papp (2012, 154), too, which, according to her data, is mostly (though not exclusively) used in more positive, “endearing” contexts, to refer to eyes that have become red due to crying. This is in dire contrast to *vörös szem*, which can also be caused by crying (as well as exertion, irritation, illness and anger) – but the connotations (and the contexts in which *vörös szem* appears) are more negative.

4.1.2. Extended senses of *piros*

Two main figurative senses can be associated with *piros*: merit, as in the case of *piros pont* ‘brownie points’; and warning, as in the case of *piros lap* ‘red card’, *piros lámpa* ‘red light’ or *piros jelzés* ‘red sign’. These extended, figurative senses can also be traced back to the external colour of the entity that the colour term describes – after all, a *piros pont* is indeed red, just as a *piros lap* or a *piros lámpa*. Allan (2009, 631) draws attention to the fact that red is often associated with danger, and claims that this is because red is the colour of blood (which indicates, therefore, a metonymical relationship – see above). The problem with this explanation, however, is that

if it were true, then Hungarian could, in theory, use *vörös* in this extended sense as well, and have sequences such as *vörös lap*, *vörös lámpa*, etc. to denote some sort of warning or danger (since *vörös* – similarly to *piros* – is also associated with blood). Nevertheless, Hungarian does not use *vörös* in this sense at all, which entails that the motivation for *piros lap*, *piros lámpa* and *piros jelzés* must be sought elsewhere.²⁰

It has been generally observed that not only is red the first colour term to be recognized by infants (as young as four months old), but it is also physiologically and psychologically the most salient (see Bornstein 1975; Ratliff 1976). Consequently, due to its conspicuousness and very high level of cognitive salience (see Uusküla & Sutrop 2007 for a comparison of the cognitive salience of colour terms), it is absolutely “ideal” to serve as a colour to draw attention to something, and typically we need our full attention to anything that is potentially dangerous or harmful to us. Hence, red has become associated with warning and potential danger. Nevertheless, why does Hungarian use *piros* and not *vörös* in this sense? We believe that the motivational basis for *piros lámpa* or *piros jelzés* resides in the fact that *piros* is a brighter hue of red, which makes it more noticeable, as opposed to *vörös*, which is a darker hue. (Note that stop signs or traffic lights are bright red in colour and not dark red, most probably for the same reason.) The fact that *piros lap* is used in various sports to indicate that a player has been suspended from the game can be traced back to a metonymical CAUSE FOR EFFECT relationship. If a particular action is dangerous, then it is prohibited to continue with it (as it might be harmful for us). Therefore, the cause (the potentially dangerous action, as signalled by the colour red) stands for the effect (suspension of the action). In fact, the cause might also stand for the potential result of the action, i.e., injury, and – as a corollary – blood, which leads us back via the SALIENT PROPERTY FOR THE CATEGORY metonymy to the very start of the metonymic chain, i.e., *piros*.

We believe that the physiological and psychological salience of red also accounts for the existence of *piros pont*, as it contrasts very well with black (the opposite of a *piros pont* in Hungarian is a *fekete pont*, ‘black point’). Needless to say, it might be argued here that a “white point” would present a better contrast with a “black point” (especially in light of the fact that white is often associated with goodness and purity – see Allan 2009, 628).

²⁰ However, the adjectival form of *vörös lámpa* (i.e., *vöröslámpás*) has a sexual connotation, as used in *vöröslámpás ház* (lit. red light house, ‘brothel’) or *vöröslámpás negyed* ‘red light district’. See section 4.2.1 for a more detailed discussion.

However, red is the most conspicuous colour on a white sheet of paper (a “white point” would simply not be visible).

4.2. *Vörös*

4.2.1. *Vörös* as an external property

Vörös also showed up as an external property of entities, though with substantial differences as compared to *piros*. There were hardly any clothes items or man-made objects that were modified with *vörös*; however, there was a preponderance of natural and organic materials (such as *bor* ‘wine’,²¹ *hús* ‘meat’, *föld* ‘earth’, *iszap* ‘sludge’, *homok* ‘sand’, *agyag* ‘clay’, *réz* ‘copper’, etc.).²² When *vörös* modifies an entity denoting a plant, in the majority of the cases the sequence is lexicalized in the sense that it is used as the name of a species – cf. *vöröskáposzta* ‘red cabbage’, *vörösáfonya* ‘lingonberry’, *vörösfenyő* ‘larch’, *vöröshagyma* ‘common onion’ or *vöröshere* ‘red clover’. This also implies that the ‘redness’ of these entities has to be interpreted rather loosely – while *piros alma* ‘red apple’ or *piros pipacs* ‘red poppy’ does have a prototypically red hue, *vöröskáposzta* and *vöröshere* are more closer to purple, while *vöröshagyma* has an orange/brown colour. A further intriguing difference between *piros* and *vörös* is that while the former modifies fruits and flowers exclusively, the latter is used with mostly vegetables on the one hand, and trees and herbs on the other (the only *vörös* flower being *rózsa* ‘rose’, along with the superordinate term *virág* ‘flower’).

Similar effects can be observed in the realm of animal names as well. In general, there were very few – only five – examples of *piros* modifying animals (more specifically *ló* ‘horse’, *madár* ‘bird’, *csikó* ‘colt’, *bogár* ‘bug’ and *hal* ‘fish’. As in the case of the plant examples with *piros*, neither one of these combinations refers to a particular animal species. However, the vast majority of animal names modified by *vörös* denote a particular animal species, as in *vörösbecy* ‘robin redbreast’, *vöröshangya* ‘red wood ant’, *vörös kánya* ‘red kite’, *vörös vércse* ‘common kestrel’ or *vörös róka* ‘red fox’. Here, too (similarly to the plant examples with *vörös*), the colour

²¹ There are some occurrences of *piros bor* as well in the data. Nevertheless, these are mostly dialectal variants used in Transylvania, which come from the literary subcorpus.

²² Note that only there are only three entities denoting clothes items that are modified by *vörös* – *nyakkendő* ‘tie’, *ruha* ‘dress’ and *kendő* ‘scarf’. The number of man-made objects was also minimal (with only a few examples such as *abrosz* ‘tablecloth’ and *tégla* ‘brick’).

of the entities is quite far from being any particular red hue, and is more closer to either a brown or an orange colour.

Vörös also appears with a number of body parts, viz. *köröm* ‘nail’, *bőr* ‘skin’, *haj* ‘hair’, *kéz* ‘hand’, *fej* ‘head’, *arc* ‘face’, *szem* ‘eyes’. While *bőr*, *fej*, *arc* and *szem* also crop up with *piros* in the data, they have generally rather different connotations when modified with *vörös*. As stated in Pusztai (2011, 1485), these body parts become *vörös* due to intense emotions (such as anger), heat or exertion; that is, under normal circumstances no body part is *vörös* (but *bőr* and *arc* can, in fact, be *piros* as a sign of good health – similarly to *nyelv* ‘tongue’). While the PROPERTY FOR CATEGORY metonymy between *vörös* and “blood” does hold in these examples, too, there is no further implication that “being full of blood” is any way positive (as it is with *piros*). Perhaps this is the reason why *vér* did not show up with *vörös* in the data – *vörös* body parts (and hence blood) has negative associations, while *piros vér* might be used in a more neutral sense.²³ Nevertheless, while *vörös* did not appear in the corpus with *vér*, i.e., blood, it did, however precede *véresejt* and *vértest* (both meaning ‘blood cell’). At the same time, neither *véresejt*, nor *vértest* showed up with *piros*. It seems then that in its literal sense *vörös* is restricted with its use of the concept of blood to more specific (levels of) usages (which supports the results of Papp 2012, 156 – for a discussion, see footnote 18 of the present paper).

A somewhat related sense of *vörös* can be found in the examples of *vörös köröm* ‘red nails’, *vörös rúzs* ‘red lipstick’ and *vörös lámpa* ‘red light’. In these cases there is an evident sexual connotation. How is, however, *vörös* related to sex? A plausible explanation is offered by Allan (2009, 631), who claims that the reason why red lips are so popular on women is because they are “found sensual” – and the reason why men do not wear red lipstick is because “red and glossed lips model an engorged vulva” (ibid.). Thus, we have once again a metonymic chain, whereby *vörös* (via the PROPERTY FOR CATEGORY metonymy) stands for the abundance of blood, which (via the EFFECT FOR CAUSE metonymy) stands for sexual arousal. Sexuality shows up with *piros* only marginally, as in the case of *piros tojás* ‘red egg’, which is given traditionally by women to men at

²³ We checked all the examples (46 items) of *piros vér* that showed up in the corpus. Interestingly, all the occurrences were limited to literary contexts and a highly elevated style. These findings do question our hypothesis on the “neutrality” of *piros* as used with *vér*. Nevertheless, a quick Google search seems to corroborate our assumptions: *piros vér* cropped up mostly in websites concerned with the human body and illnesses, while *vörös vér* was mostly restricted to figurative usage and showed up in movie titles, poems, the title of an online game, etc.

Easter. However, the red egg is primarily a symbol of fertility, and not of sex – no wonder that it is also given to/by children as well, and hence modified by *piros* (which is of a more “innocent” nature than *vörös*).

4.2.2. Extended senses of *vörös*

Vörös is very prevalent with concepts that can be associated with revolutions in some way – notably *csillag* ‘star’, *zászló* ‘flag’, *khmer* ‘khmer’, *katona* ‘soldier’, *terror* ‘terror’, *hadsereg* ‘army’, *lobogó* ‘flag’, *diktatúra* ‘dictatorship’, *órség* ‘guard’. A possible general motivating factor here is a metonymic chain that links *vörös* to the concept of revolution: first, *vörös* stands for blood via the SALIENT PROPERTY FOR CATEGORY conceptual metonymy (the colour of the entity for the entity); then a further metonymy is involved, by which blood stands for injury/casualty via the EFFECT FOR CAUSE metonymy. As a third step, injury stands for revolution, via the RESULT FOR ACTION metonymy (since a revolution typically results in a lot of injuries/casualties). One of the most intriguing questions that can be raised here is why does Hungarian associate *vörös* systematically with the concept of revolution (and its related concepts), as opposed to *piros*?

We believe that the answer to this question lies in the fact that *piros* is considered as a brighter hue of red, while *vörös* is generally described as a darker hue (see Csapodi 1899; Kenedy 1921; Maclaury et al. 1997; Pusztai 2011). It has been observed in a number of semantic studies on colour terms (Philip 2006; Szlávi 2008; Allan 2009) that darker colours are often used to describe negative concepts, while brighter colours are more usually associated with positive concepts. (No wonder that according to Maclaury et al. 1997, 77, subjects described *vörös* as “more sinister”.) With this in mind, *vörös* is better suited to stand for the concept of revolution, which nearly always entails injury, casualty, and hence death. *Piros*, on the other hand, is not associated with these concepts due to that fact that it is a brighter hue. In a similar vein, the metonymic chain of SALIENT PROPERTY FOR CATEGORY (*vörös* for blood) and EFFECT FOR CAUSE (blood for injury/casualty) might be the motivating force behind the coinage of *vöröskereszt* ‘red cross’. The negative, “sinister” (Maclaury et al. 1997, 77) connotation of *vörös* might also explain the sequence *vörös ördög* (note that there were no *piros + ördög* sequences in the data). Devils are often depicted in images as red, probably as a PROPERTY FOR CATEGORY metonymy for the colour of fire. Fire, however, is a metonymy for hell.

A further figurative use of *vörös* appears in *vörös posztó* ‘red rag’, which is used idiomatically as something that makes people angry (i.e., as

in the English phrase *be like a red rag to somebody*). The motivation for the use of *vörös* (and not *piros*) can be explained by two factors. First, as it has been pointed out above, materials are often described with the colour term *vörös*, and not *piros*. Therefore, the selection of *vörös* in *vörös posztó* can be partly traced back to semantic analogy on the basis of all the other names of (natural) materials (and hence describe the red colour of the material). Second, as also alluded to above, *vörös* is associated with body parts that become red due to – among others – intense emotions such as anger. Thus, *vörös* in *vörös posztó* might metonymically refer to the emotion of anger (whereby *vörös* stands for the blood that rushes into one’s head/face when one gets angry) as a consequence of somebody else’s “red rag” behaviour.

All in all, the semantic investigation of *piros* and *vörös* sequences has shed light on a number of definite trends which are in agreement with the corpus-based analysis. For an overview of the semantic domains they are associated with, see table 5.

Table 5: Semantic domains associated with *piros* and *vörös*

<i>Piros</i>	<i>Vörös</i>
external colour of an entity (<i>piros szoknya, piros labda</i>)	name of a particular species (<i>vörösfonya, vöröshangya</i>)
ripeness and maturity (<i>piros paradicsom, piros cseresznye</i>)	exertion and anger (<i>vörös arc, vörös szem</i>)
health (<i>piros száj, piros nyelv</i>)	sexuality (<i>vörös rúzs, vörös lámpa</i>)
merit (<i>piros pont</i>)	revolution (<i>vörös zászló, vörös khmer</i>)
warning (<i>piros lap, piros lámpa</i>)	injury and casualty (<i>vöröskereszt</i>)

First, out of the two colour terms, *piros* is used much less frequently in a figurative sense than *vörös*; it is mostly used to describe the external colour of an entity. Nevertheless, *piros* does have its various associations, which are ripeness, maturity and health on the one hand, and (cognitive and perceptual) salience on the other. On the other hand, *vörös* is used much less frequently to describe the external colour of an entity. When it is used for this purpose, however, then quite often the real colour of these entities is closer to purple (as in the case of *vöröskáposzta*, i.e., ‘red cabbage’), brown (as in the case of *vörösfenyő*, i.e., ‘larch’) or even orange (as in the case of *vörös réz*, i.e., ‘copper’). Second, when a *piros* or *vörös* sequence

is used to denote an animal or plant, then the latter usually denotes a species (such as *vöröshangya*), which is not the case with *piros* (which is used in a general, descriptive sense). There was evidently a much larger proportion of *vörös* + N types used in a figurative sense; these associations can be linked metonymically from the concept of blood to injury/casualty, revolution, anger, exertion, sexuality and fire.

Initially, we expected the concept of blood – via the PROPERTY FOR CATEGORY metonymy – to play a role in the semantics of *vörös* only (blood being its etymological basis). However, our results indicate that the concept of blood and the PROPERTY FOR CATEGORY metonymy significantly contribute to the senses of *both* colour terms. However, while blood is associated with mostly positive aspects such as health and ripeness in the case of *piros*, it is associated with mostly negative aspects such as exertion, injury and even death in the case of *vörös*. These associations (and the positive–negative polarity of *piros* and *vörös*) might be related to the brighter vs. darker hues of the two colour terms. This polarity has important consequences for the semantics of the two colour terms: *piros* is used in a neutral sense more frequently than *vörös*. What these results seem to indicate is that the meaning of *vörös* as ‘red’ (in the sense of ‘red colour’) is by all accounts highly controversial, thereby calling into question its basic colour term status.

5. Conclusions

At this point it can be safely said that our first two hypotheses have been justified by the corpus- and cognitive linguistic analyses. Accordingly, what we stated in the first hypothesis was that we expected considerable differences between the type/token ratios of the two terms, which could be explained by the more lexicalized status of *vörös* (i.e., that the use of *vörös* was restricted to a limited number of more frequently used expressions, while *piros* was used in a much wider range of expressions with varying degrees of frequency). Our second hypothesis claimed that *piros* would be mostly used in its primary, literal sense, to refer to the external colour of an entity, while *vörös* would be less inclined to be used thus, and would be more inclined to be used in a figurative sense. The semantic analysis has demonstrated that the figurative senses of *vörös* can be metonymically related to the concept of blood, which is the etymological basis of the colour term. In sum, what we have concluded from the semantic analysis is that the meaning of *vörös* as ‘red’, and, in turn, its basic colour term status, is highly questionable.

What our research has shown is that basicness has been generally confused with frequency, which is a methodological flaw. Although *vörös* is a very frequent colour term in Hungarian (see table 1), its frequency is due to the fact that it appears in often used collocations. The problem with these collocations, however, is that a considerable proportion of these does not actually denote an entity that is red in its appearance. Many of the *vörös* sequences – however frequent – have undergone idiomatization, where *vörös* means something else than ‘red’. Out of the two colour terms, *piros* is more inclined to crop up in low-frequency, novel coinages, which seems to indicate that this is the colour term in Hungarian that is more inclined to be used in the literal sense of the colour ‘red’. This finding has been corroborated with the semantic analysis as well, where *piros* surfaced as the positive/neutral term. It should be underlined here that out of the two terms *piros* was associated with salience, not *vörös* – and salience is, after all, a defining feature of basicness (Moss 1989). In sum, our conclusion is that *vörös* is not a basic colour term of Hungarian. More rather, in line with Berlin & Kay’s (1969, 35) alternative – and lesser known – hypothesis, we claim that *piros*, as a basic colour term, has two senses, ‘red’ and ‘lighter red’, while *vörös*, ‘darker red’, is only a secondary colour term in Hungarian.

The present paper wished to examine the issue of basicness with regard to *piros* and *vörös*, and, based on the combined approach of corpus and cognitive linguistics, it has come to the conclusion that *vörös* is not a basic colour term. Full support of the present claim would come from psycholinguistics; unfortunately, however, there are no such experiments yet on how and in what sequence Hungarian children learn and acquire the colour terms *piros* and *vörös*.²⁴ Nevertheless, based on our results, we would very much expect that *vörös* would be learned at a much later stage than *piros*, thereby confirming our claims and settling the issue once and for all.²⁵

²⁴ There have been, however, forays into the acquisition of colour terms in general by Gósy (1998), and whose research might serve as a blueprint for further ones in the future. Gósy asked 50 Hungarian preschoolers (ages 5 to 7) to name the various colours that were shown to them (altogether 20). All the children were able to correctly identify *piros*. (Unfortunately there is no data in the paper on *vörös*.)

²⁵ The first author has two 3.5-year-old children, whose first colour term was *piros* (around age 2), which was used at first in the sense of any colour. The next two colour terms were *kék* and *zöld*, followed by *sárga*, *fekete*, *lila* and *narancs(sárga)*. *Vörös* is still not yet part of their vocabulary.

References

- Allan, Keith. 2009. The connotations of English colour terms: Colour-based X-phemisms. *Journal of Pragmatics* 41. 626–637.
- Bárczi, Géza (ed.). 1994. *Magyar szófejtő szótár* [Hungarian etymological dictionary] (2nd edition). Budapest: Trezor.
- Barratt, Leslie B. and Miklós Kontra. 1996. Matching Hungarian and English color terms. *International Journal of Lexicography* 9. 102–117.
- Bauer, Laurie. 1983. *English word-formation*. Cambridge: Cambridge University Press.
- Benkő, Loránd. 1967–1984. *A magyar nyelv történeti-etimológiai szótára* [A historical-etymological dictionary of the Hungarian language]. Budapest: Akadémiai Kiadó.
- Berlin, Brent and Paul Kay. 1969. *Basic color terms*. Berkeley: University of California Press.
- Bornstein, Marc H. 1975. Qualities of color vision in infancy. *Journal of Experimental Child Psychology* 19. 401–419.
- Bybee, Joan L. 2001. *Phonology and language use* (Cambridge Studies in Linguistics 94). Cambridge: Cambridge University Press.
- Csapodi, István. 1899. *Vörös és piros* [Vörös and piros]. *Magyar Nyelvőr* 28. 201–204.
- Czuczor, Gergely and János Fogarasi. 1862–1874. *A magyar nyelv szótára* [A dictionary of Hungarian]. Pest: Emich Gusztáv Magyar Akadémiai Nyomdász.
- Eőry, Vilma (ed.). 2007. *Értelmező szótár +* [Hungarian monolingual dictionary +]. Budapest: Tinta Könyvkiadó.
- Forbes, Isabel. 1979. The terms *brun* and *marron* in modern standard French. *Journal of Linguistics* 15. 295–305.
- Gárdonyi, József. 1920. *Piros vagy vörös* [Piros or vörös]. *Magyar Nyelv* 16. 84–87.
- Gósy, Mária. 1998. Színmegnevezések gyermekkorban és felnőttkorban [Colour naming in childhood and adulthood]. In Zs. Lengyel and J. Navracsics (eds.) *Alkalmazott nyelvészeti tanulmányok II* [Studies in applied linguistics 2]. Veszprém: Veszprémi Egyetem. 55–69.
- Gries, Stefan Th. 2006. Introduction. In S. T. Gries and A. Stefanowitsch (eds.) *Corpora in cognitive linguistics: Corpus-based approaches to syntax and lexis*. Berlin & New York: Mouton de Gruyter. 1–18.
- Kenedy, Géza. 1921. *Piros vagy vörös* [Piros or vörös]. *Magyar Nyelv* 18. 33–34.
- Kicsi, Sándor. 1988. Az alapszínnevek lexikalizálásáról [The lexicalization of basic colour terms]. *Magyar Nyelvőr* 112. 456–463.
- Kiefer, Ferenc. 2005. Van-e magyar jelentéstan? [Does Hungarian semantics exist?]. *Magyar Nyelvőr* 101. 129–140.
- Kiss, Gábor. 2004. A *piros*, *vörös* és más színnevek használata a Magyar Nemzeti Szöveggyűjtő alapján [The use of *piros*, *vörös* and other colour terms on the basis of the Hungarian National Corpus]. In T. Gecső (ed.) *Variabilitás és nyelvhasználat* [Variability and language use]. Budapest: Tinta Könyvkiadó. 160–165.
- Kiss, Gábor and Isabel Forbes. 2001. *Piros, vörös – red, rot, rouge*. In T. Gecső (ed.) *Kontrasztív szemantikai kutatások* [Studies in contrastive semantics]. Budapest: Tinta Könyvkiadó. 190–199.

- Laczkó, Krisztina and Attila Mártonfi. 2004. Helyesírás [Spelling]. Budapest: Osiris Kiadó.
- Lengyel, Klára. 2000. A szóösszetétel [Compounds]. In B. Keszler (ed.) Magyar grammatika [Hungarian grammar]. Nemzeti Tankönyvkiadó: Budapest. 324–348.
- Maclaury, Robert E., Judit Almási and Zoltán Kövecses. 1997. Hungarian *piros* and *vörös*: Color from points of view. *Semiotica* 114. 67–81.
- Moss, Anthony E. 1989. Basic colour terms: Problems and hypotheses. *Lingua* 78. 313–320.
- Papp, Eszter. 2012. A magyar színnevek terminológiai elemzése [A terminological analysis of Hungarian colour terms]. Doctoral dissertation. University of Pécs.
- Philip, Gill. 2006. Connotative meaning in English and Italian colour-word metaphors. *Metaphorik* 10. 59–93.
- Pusztai, Ferenc (ed.). 2011. Magyar értelmező kéziszótár [A concise dictionary of Hungarian]. Budapest: Akadémiai Kiadó.
- Ratliff, Floyd. 1976. On the psychophysiological bases of universal color terms. *Proceedings of the American Philosophical Society* 120. 311–330.
- Selényi, Pál. 1948. *Piros és veres* [*Piros* and *veres*]. *Magyar Nyelvőr* 72. 12–14.
- Sweetser, Eve. 1999. Compositionality and blending: Semantic composition in a cognitively realistic framework. In T. A. Janssen and G. Redeker (eds.) *Cognitive linguistics: Foundations, scope, and methodology*. Berlin & New York: Mouton de Gruyter. 129–162.
- Szlávi, Anna. 2008. The conceptualization of black and white in adjective + noun compounds. MA thesis. Eötvös Loránd University.
- Uusküla, Mari. 2008. Basic colour terms in Finno-Ugric and Slavonic languages: Myths and facts. Tartu: Tartu University Press.
- Uusküla, Mari and Urmas Sutrop. 2007. Preliminary study of basic colour terms in modern Hungarian. *Linguistica Uralica* 43. 102–122.
- Wierzbicka, Anna. 1990. The meaning of color terms: Semantics, culture, and cognition. *Cognitive Linguistics* 1. 99–150.
- Wierzbicka, Anna. 1996. *Semantics: Primes and universals*. Oxford: Oxford University Press.
- Wong, May L.-Y. 2012. Corpus linguistics and cognitive linguistics: A convergence of basic assumptions. *International Journal of Cognitive Linguistics* 3. 43–68.
- Zaicz, Gábor (ed.). 2004. Etimológiai szótár. Magyar szavak és toldalékok eredete [Etymological dictionary: The origins of Hungarian words and suffixes]. Budapest: Tinta Könyvkiadó.

Appendix 1

Type and token frequencies of the *piros* + N and *vörös* + N sequences, in descending order of frequency

Number of types	<i>Piros</i> + N	Number of tokens	<i>Vörös</i> + N	Number of tokens
1	lámpa 'lamp/light'	356	kereszt 'cross'	1,872
2	pont 'point/mark'	228	bor 'wine'	1,157
3	paprika 'paprika'	198	ördög 'devil'	414
4	lap 'card'	114	hagyma 'onion'	398
5	szín 'color'	113	csillag 'star'	329
6	zászló 'flag'	102	zászló 'flag'	222
7	rózsa 'rose'	95	bolygó 'planet'	192
8	alma 'apple'	93	khmer 'khmer'	186
9	felt 'stain/blot'	85	réz 'copper'	138
10	virág 'flower'	82	katona 'soldier'	116
11	jelzés 'sign'	73	posztó 'cloth'	113
12	ruha 'dress'	64	szín 'color'	109
13	tojás 'egg'	61	haj 'hair'	106
14	arc 'face'	60	fény 'light'	106
15	ceruza 'pencil'	54	begy 'robin redbreast'	94
16	szalag 'ribbon'	53	here 'clover'	90
17	gomb 'button'	53	rózsa 'rose'	88
18	csík 'bar/strip'	53	fenyő 'larch'	86
19	betű 'letter'	53	felt 'stain/blot'	82
20	szem 'eye'	52	véresejt 'blood-cell'	79
21	sapka 'hat'	52	terror 'terror'	79
22	vonal 'line'	51	köd 'fog'	77
23	vér 'blood'	46	szőnyeg 'carpet'	71
24	fény 'light'	45	hadsereg 'army'	68
25	festék 'paint'	42	vértest 'blood cell'	66
26	szőnyeg 'carpet'	40	eltolódás 'shift'	65
27	csizma 'boot'	39	arc 'face'	61
28	autó 'car'	39	márvány 'marble'	60
29	szív 'heart'	37	fej 'head'	51
30	labda 'ball'	36	festék 'paint'	50
31	szegfű 'carnation'	34	szem 'eye'	49
32	bor 'wine'	34	tégla 'brick'	45
33	betli ²⁶	33	bársony 'velvet'	42
34	nadrág 'trousers'	32	szegfű 'carnation'	41
35	pötty 'dot'	31	lobogó 'flag'	39
36	tulipán 'tulip'	30	gárdista 'guardist'	38
37	száj 'mouth'	30	nyakkendő 'tie'	37
38	ing 'shirt'	29	óriás 'giant'	35
39	csőr 'beak'	29	kakas 'cock'	33

40	sál 'scarf'	28	csík 'strip/bar'	33
41	szoknya 'skirt'	28	iszap 'sludge'	28
42	kendő 'scarf'	28	hangya 'ant'	26
43	kabát 'coat'	28	kánya 'kite'	26
44	sáv 'zone'	27	vonal 'line'	24
45	orr 'nose'	27	postakocsi 'stagecoach'	24
46	könyv 'book'	27	kő 'stone'	24
47	pulóver 'sweater'	26	salak 'sludge'	22
48	nap 'sun/day'	26	mező 'field'	22
49	bogyó 'berry'	26	zóna 'zone'	21
50	pántlika 'ribbon'	25	hús 'meat'	21
51	nyakkendő 'tie'	25	betű 'letter'	21
52	láng 'flame'	25	virág 'flower'	20
53	kör 'circle'	24	ruha 'dress'	20
54	papucs 'slipper'	22	nap 'day/sun'	20
55	ló 'horse'	22	gróf 'count'	20
56	kocsi 'car'	22	ember 'person'	20
57	alap 'base'	22	lámpa 'lamp/light'	19
58	sarok 'corner'	21	fonál 'thread'	19
59	madár 'bird'	21	szalag 'ribbon'	18
60	háromszög 'triangle'	21	selyem 'silk'	17
61	lé 'juice'	20	kutya 'dog'	17
62	fej 'head'	20	por 'dust'	16
63	bársony 'velvet'	20	kép 'image'	15
64	ász 'ace'	19	hó 'snow'	15
65	paradicsom 'tomato'	19	föld 'soil'	15
66	műanyag 'plastic'	19	bőr 'skin/leather'	15
67	cipő 'shoe'	19	agyag 'clay'	15
68	bőr 'skin/leather'	19	vércse 'kestrel'	14
69	ajak 'lip'	19	maffia 'mafia'	14
70	útlevel 'passport'	18	kéz 'hand'	14
71	tinta 'ink'	18	kokárda 'rosette'	14
72	nyíl 'arrow'	18	homokkő 'sandstone'	14
73	hó 'snow'	18	báró 'baron'	14
74	toll 'pen/plume'	17	sarok 'corner'	13
75	szandál 'sandal'	17	kendő 'scarf'	13
76	pecsét 'seal'	17	izzás 'glow'	13
77	masni 'bow'	17	homok 'sand'	13
78	gömb 'button'	16	farok 'tail'	13
79	fez 'fez'	16	drapéria 'drapery'	13
80	csillag 'star'	16	diktatúra 'dictatorship'	13
81	pad 'bench'	15	káposzta 'cabbage'	13
82	könyvecske 'booklet'	15	róka 'fox'	12
83	király 'king'	15	pont 'point/mark'	12
84	kereszt 'cross'	15	köröm 'nail'	12
85	blúz 'blouse'	15	korong 'disc'	11

86	öv 'belt'	14	félhold 'crescent'	11
87	zokni 'sock'	14	órség 'guard'	10
88	trikó 'T-shirt'	14	áfonya 'lingonberry'	10
89	selyem 'silk'	14	szőr 'hair'	10
90	petty 'dot'	14	szikla 'cliff'	10
91	nyelv 'tongue'	14	macska 'cat'	10
92	gyertya 'candle'	14	kockakő 'flagstone'	10
93	cseréptető 'tiled roof'	14	fal 'wall'	10
94	tégla 'brick'	13	bársonyfüggöny 'velvet curtain'	10
95	pipacs 'poppy'	13	alap 'base'	10
96	kő 'stone'	13	abrosz 'tablecloth'	10
97	fazék 'pot'	13		
98	vászon 'linen'	12		
99	karika 'ring'	12		
100	fotel 'armchair'	12		
101	csikó 'colt'	12		
102	bogár 'bug'	12		
103	szám 'number'	11		
104	metró 'metro'	11		
105	mellény 'vest'	11		
106	hal 'fish'	11		
107	gyümölcs 'fruit'	11		
108	függöny 'curtain'	11		
109	filctoll 'marker pen'	11		
110	cserép 'tile'	11		
111	cseresznye 'cherry'	11		
112	topánka 'slipper'	10		
113	tető 'roof'	10		
114	tenyér 'palm'	10		
115	szírom 'petal'	10		
116	ribizli 'redcurrant'	10		
117	posztó 'cloth'	10		
118	mező 'field'	10		
119	mez 'jersey'	10		
120	lada 'Lada car'	10		
121	kalap 'hat'	10		
122	fakanál 'wooden spoon'	10		
123	cérna 'thread'	10		
124	bojt 'tassel'	10		
125	bicikli 'bicycle'	10		

²⁶ *betli*: an expression used in the card game *ulti*, when no scores are expected by the player in a particular round.