Tanulmány

Melinda Majtényi

The productivity of the prefix *über-* in English and Hungarian: A cross-linguistic, corpus-based study[*]

Abstract

The aim of this paper is to provide a corpus-based, cross-linguistic investigation of the prefix *über-*. Using the Corpus of Contemporary American English and a Hungarian webcorpus, I have collected data containing instances of *über-* and assigned a rating of productivity based on Baayen's (2009) productivity measures. Additionally, the bases to which *über-* is attached, the spelling of these constructions, genre in which they occur and their meaning have been examined in the corpus data set.
Keywords: productivity, English, Hungarian, corpus, prefix

1   Introduction

Language changes as time passes by. Among the different levels of language, morphology is exposed to language change as well. New morphological categories may come into existence and established ones may disappear. In this study, a recent change will be described both in English and Hungarian. The case of *über-* is a good illustration of how changeable natural language is.

The prefix *über-* is well-known from German as a verbal particle in cases such as *übergehen* or *übersehen*. The examples below taken from the data set collected within the framework of this study illustrate that *über-* cannot be found only in German any more, but also in English and Hungarian.

(1) The movie version of this book of the <u>uber-popular</u> trilogy by Suzanne Collins arrives in theaters March 23.
(2) Minni's Haight Street location may be <u>uberhip</u>, but Kantor's succulent smoked ribs and sausages are unabashedly traditional.
(3) Két darab <u>überintenzív</u> próbái kezdődtek meg…
(4) <u>Übermájer</u> módon most magamtól idézek…

2   Theoretical framework

2.1   *Previous approaches to morphological productivity*

In linguistics, when we use the word "productivity", it generally refers to the productivity of morphological processes. As Plag (2006: 547-549) highlights morphological productivity is a multifaceted phenomenon that is a derived notion, but essential in the description of different word formation processes. This section aims to clarify the concept of morphological productivity, distinguishing two basic approaches to the notion, and provide a huge variety of productivity measures.

In the linguistic literature there have been a number of attempts to define the term "morphological productivity" (Jespersen 1942, Hockett 1958, Aronoff 1976, Baayen & Lieber 1991, Bauer 2001, Baayen 1993, 2009 among others). Jespersen (1942: 4) was the first who mentioned productive rules in English word-formation. He equated productive rules with "living" formations. Hockett (1958: 575) uses the label "productivity" for the feature of language that allows people to say things which have never been said before. Chomsky (1965: 6) calls this property of natural languages "creativity". Bauer (2001) divides productivity into two different phenomena: availability and profitability. Availability means that new words can be produced with the help of available morphological processes. It is a binary, basically qualitative notion: a process is either available or not. Profitability, however, is quantitative in nature and means the extent to which a morphological process is able to form new formations.

In the English and Germanic linguistic literature two basic approaches to productivity can be differentiated. In a qualitative approach linguists (Kastovsky 1986, Dressler 1997, Ladányi 2001, Booij 2002, Kiss 2011 among others) think that all morphological processes should be described in terms of rules. In this rule-governed view, the concept of productivity is determined by constraints that are imposed on a given rule. Thus, productivity is explained by the semantic features of the domain the rule applies to and not by the number of newly coined words (Ladányi: 2001: 233 among others). These linguists approach productivity in a natural framework and propose that productivity and frequency do not always occur together. It is claimed that productivity is independent of frequency. Bauer (1992), as an advocate of natural morphology, emphasizes that the concept of productivity cannot be explained via type or token frequency,[1] but with the help of different constraints that mean how big difficulties are that a coinage needs to overcome. It is held that the bigger the obstacle to overcome the more productive the process.

Others (Aronoff 1976, Hay and Baayen 2002, Baayen 2009 among others), on the other hand, view productivity as a quantitative notion, that is, it is a matter of degree. They claim that certain morphological processes are more productive or less productive than others. Morphological productivity is not a black-or-white phenomenon and as such morphological processes are not clearly divided into productive or unproductive ones (Baayen 2009). Based on this concept of productivity, the advocates of this quantitative approach have proposed several methods of measuring productivity, which will be elaborated on in the next section.

---

[1]   The term "types" refers to different words in a corpus and tokens are occurrences of a type. For example, *Sue travelled to Mexico last year and wanted to go back there last month* includes two tokens of the word form *last*.

Rainer (1987) summarizes six different types of the definitions that appear in the literature:

a) a definition in terms of the frequency of the output words
b) a definition in terms of the number of available bases
c) a definition in terms of the proportion of words actually used to the number of words potentially created by a particular process
d) a definition in terms of the possibility of forming new words
e) a definition in terms of the probability of new forms occurring
f) a definition in terms of the number of new forms occurring in a specified period of time

Definition a, b, c and f are quantitative, whereas definition d and e are qualitative in nature. Definition f relies on a diachronic perspective, unlike the rest that are basically synchronic. Definition a, b and f are based on "existing words", others on "potential words".

There are many competing, fundamentally different definitions and theories of morphological productivity. Bauer (2001: 25) mentions that this amount of variation "leaves studies of productivity […] in a rather poor state".

For the purposes of this research, the quantitative notion of productivity described above will be used. Productivity is perceived as a scalar term here, because as several examples have shown in the linguistic literature (Baayen & Lieber 1991, Baayen 1992, Plag-Dalton-Puffer-Baayen 1999 among others), this approach can be applied successfully and validly to analyses of corpora. This notion of productivity is similar to Rainer's (1987) definition (c). We have arrived at the definition of morphological productivity which will be used as a guideline for the analysis.

2.2   *Measuring productivity*

Along with the difficulties of defining productivity, there are a number of methods of measuring productivity. As one of the earliest ways to measure productivity, Aronoff (1976: 36) calculates the ratio of actual words to possible words. Possible words are the formations that can be produced by pertinent morphological rules. In contrast, actual words include existing established words in a given language. Thus, a morphological process is highly productive if this ratio is high.

Plag (2004: 6), however, reveals the drawbacks of this model. It makes wrong predictions for example in the case of –*ness*. The endless number of potential words yields an extremely low productivity index. In Baayen's (1989) opinion it should be renamed as "an index of unproductivity". Lieber (1992) also advises us against defining productivity in terms of actual words because there are major problems with the notion of actual words. Actual words are commonly equated with the words which are listed in dictionaries, but not every actual word can be listed in dictionaries.

Baayen (1989, 1993 and 2009) proposes a number of statistical, corpus-based measures that can help us reveal different aspects of morphological productivity. Baayen (2009: 902-905) summarizes the following mathematical formalizations of productivity.

The first measure defines productivity in terms of how many types of a given morphological category can be found in a corpus. It is called realized productivity because it basically looks at "realized", past use. As Baayen (2009: 902) puts it "the realized productivity of a morphological category C is estimated by the type count V(C,N) of its members in a corpus with N tokens". It is a type-based estimate of productivity. In his previous work, Baayen

(1993) calls it the extent of use. It totally ignores token frequency. Hence, productive morphological categories can be described by high type frequency. For instance, the English regular past tense ending *–ed* can be attached to thousands of verbs, whereas irregular schemas are very rare. The regular past tense schema has a higher realized productivity than the irregular one.

Realized productivity, however, represents only one aspect of productivity. As Baayen (2009: 902) highlights, this first approximation of productivity is problematic in some cases. First, it does not account for the similarity between words that are produced with the help of the same schema. Second, it does not take into account the lower weight of low-frequency words. Third, type-based counts do not do justice to different morphological categories.

Before we go on with the presentation of the second measure of productivity, we need to pay some attention to the notion of hapax legomena. Hapax legomena are types which occur exactly once in a corpus. These are words in the data-set with a frequency of 1. The number of hapax legomena in a corpus of size N is $V(1,N)$, whereas the number of hapax legomena for the morphological category C in a corpus of size N can be described as $V(1,C,N)$. Baayen (2009) warns us against confusing hapax legomena with native neologisms. Even in a large corpus, there may occur words among hapax legomena with a basically well-established form. We should also keep in mind that hapax legomena are only tools for statistical measurements.

The second measure of productivity is referred to as the hapax-conditioned degree of productivity (Baayen 1993). It shows how many new forms are contributed by the morphological category C. Thus, it assesses the rate at which a morphological category C is expanding. Expanding productivity is similar to Bauer's (2001) profitability. It can be estimated as follows: $P=V(1,C,N)/V(1,C)$ where $V(1,N)$ denotes the total number of hapax legomena in a corpus. It captures the differences by comparing counts of hapax legomena and considers a "category's contribution to the growth rate of the vocabulary in corpus". (Baayen 2009: 902) The problem with this second measure is that the total number of hapax legomena is needed in the complete corpus.

The third type of measure is called potential productivity. Baayen (1993) calls it "productivity in the narrow sense". It is based on how many tokens of a morphological category are new forms. This ratio is known as the "category-conditioned degree of productivity" (Baayen 2009: 903). It is the ratio of hapax legomena in a given morphological category to the total number of tokens in that category: $P*=V(1,C,N)/N(C)$. If $P*$ is higher than that of simplex words, the process is productive. On the contrary, the schema is unproductive, if the value of $P*$ is lower than that of simplex words. In other words, it is easier to produce a new word than to form a new item with the help of the schema. The first study which validated this measure was Baayen (1994).

On the other hand, some linguists do not consider this measure appropriate. One fault is that "it is not possible to weight the relative contributions of [two phenomena]" (Bauer 2001: 154). If the productivity of two or more affixes is compared, it renders the question meaningless. Furthermore, linguists approaching productivity in a natural framework notice that it focuses on performance-level probability and totally ignores competence-level potentiality (Ladányi 2001: 234).

However, the fact has never been denied that $P*$ and Baayen's concept of quantitative productivity is applied successfully to research in big corpora. In this paper I will be using the measure $P*$ to gauge the productivity of the prefix *über-* in English and Hungarian corpora.

3   Data analysis

In both analyses fairly large corpora have been used to extract the instances of über-constructions. After searching the data set, lists of all types were compiled with frequency data (see Appendix). These lists have been investigated focusing on the following research questions:

Question 1   What is the part of speech of bases to which *über-* is attached?
Question 2   How are über-constructions spelled and what determines their spelling?
Question 3   In which genres does it occur most frequently?
Question 4   With the help of Baayen's (2009) P* how can the morphological productivity of *über-* be measured?
Question 5   What kind of meaning is conveyed by the prefix *über-*?

3.1   *über- in English*

3.1.1   *Research data and methods*

For the English part of this research, the Corpus of Contemporary American English (COCA) has been used. It was created by Mark Davis, and it is one of the largest freely-available online corpora. It contains more than 450 million words in 189,431 texts. Each year 20 million words were collected from 1990-2012. The most recent texts were added in June 2012. It is a balanced corpus, i.e. it covers a range of text categories. In COCA each year, about four million words are collected per genre: spoken, fiction, magazine, newspaper and academic texts. Out of the roughly 450 million words, the spoken part adds up to 95 million words. It makes it possible for scholars to investigate linguistic phenomena in terms of different extra-linguistic variables. It enables the user to view the context in which words appear. Moreover, it provides different search options allowing the collection of affix data. Due to its design, it suits our purposes best. It is appropriate for looking at current, ongoing changes in the English language.

One of the greatest advantages of COCA is that it can be searched quite easily. I looked for word forms which included the prefix *über-*. In COCA the string uber* must be typed as the search string and it provides us with a nice list of all the instances where *über-* is involved. However, the whole process cannot be automatized. Word forms such as *ubermensch* and *ubergang* had to be omitted manually from the list in order to include only the tokens which are formed in English. *Ubermensch* and *ubergang* are direct loans from German and would cause a distortion of data, if they were left in the list.

3.1.2   *Results*

Regarding question 1, I have found the following type and token frequencies: 257 tokens out of which there were 187 noun tokens and 70 adjectives; 170 types with 118 noun types and 52 adjective types. The most frequent bases were *nanny* (47), *uberman* (10), *hip* (5), *model* (4) and *liberal* (4). Therefore, *über-* is only attached to nouns and adjectives in English and as Figure 1 demonstrates it, it occurs more often as part of nouns than that of adjectives. Nouns add up to nearly 70% of the types.
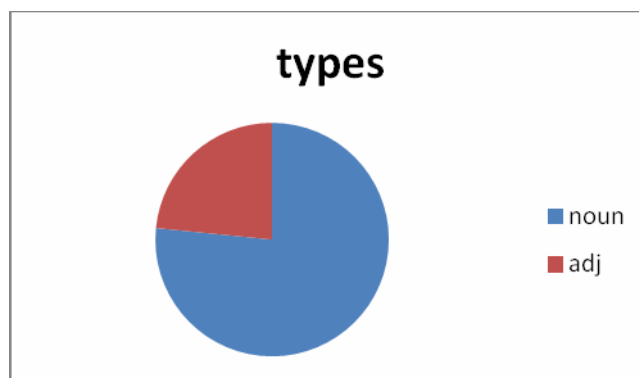
Fig. 1: Distribution of noun and adjective bases in COCA

In the list of data there are three different ways of spelling. The construction can be spelled as one single word, with a hyphen separating the prefix and the base or as two separate words. The distribution of the spelling schemata is demonstrated in Table 1.

Ways of spelling	Number of tokens
one word	114
with hyphen	101
two words	39

Table 1: Distribution of spelling in COCA

COCA enables us to carry out diachronic research. If we look at spelling diachronically, we can see that as we proceed in time, a change in the tendency of spelling can be observed after 2007. Before 2007, the majority of über-constructions were written with a hyphen. After it more and more forms are written as one word and the proportion of hyphenated forms is declining.

	<1998	1998-2002	2003-2007	2008-2012
one word	43.75%	34%	31%	55%
with hyphen	50%	53%	52%	28%
two words	6.25%	13%	17%	17%

Table 2: Diachronic Distribution of spelling in COCA
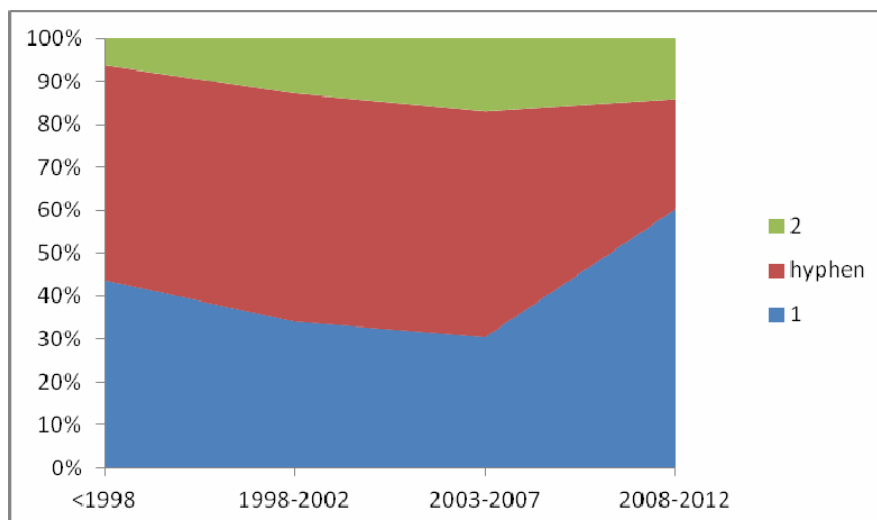
Fig. 2: Diachronic Distribution of spelling in COCA

The distribution of genres can highlight some important facts about the pragmatic side of the phenomenon. As Figure 3 illustrates, *über-* can be found mostly in magazines and newspapers. A word form with the prefix *über-* is the least likely to occur in academic genres. The contribution of spoken texts cannot be ignored either. As we can see, *über-* is far more frequent in written than in spoken language. It emerged to be highly freuquent in magazines, fiction and newspapers, but fairly frequent in academic texts.
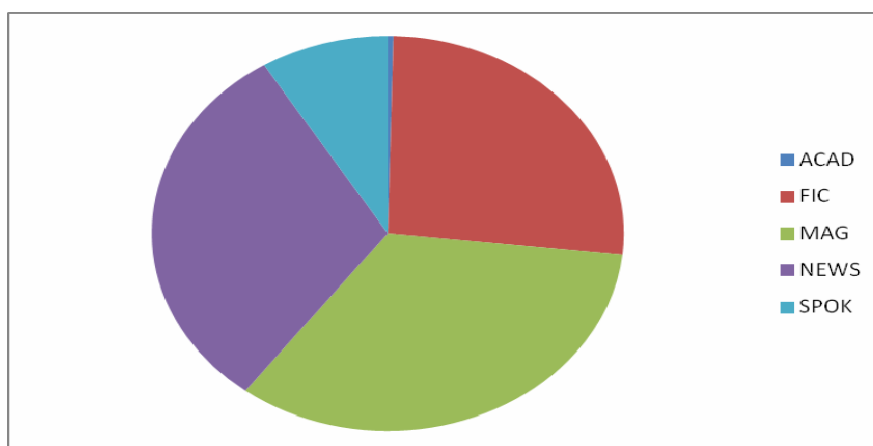
Fig. 3: Über-constructions in different genres in COCA

Question 4 is concerned with productivity. Below the morphological productivity of *über-* will be gauged with the help of P*. As we have already seen, the token frequency is 257. The number of hapax legomena is 140. Finally, we should calculate "productivity in the narrow sense". It assigns the potential productivity index of 0.54 to *über-*. Table 3 summarizes the results obtained.

token frequency	N(C)=257
hapax legomena	V(1,C,N)=140
P*	V(1,C,N)/N(C)=140/257=0.54

Table 3: Productivity index (based on COCA)

Out of 257 tokens more than half are hapax legomena. P*= 0.54 is indicative of high productivity. These results are statistically significant.

The question arises whether a change of productivity can be identified over time. Productivity through time has been elaborated on in several corpus-based studies, for instance in Meibauer/Guttopf/Scherer (2004) or Scherer (2005). Based on our data set, productivity can be measured diachronically as follows.

The texts in COCA are taken from the time period between 1990 and 2012. This time span can be divided into four shorter periods and token frequency and P* can be counted. Table 4 includes the productivity indexes obtained.

	<1998	1998-2002	2003-2007	2008-2012
N (C)	21	57	76	103
P*	1.00	0.86	0.93	0.81

Table 4: Diachronic measures of productivity (based on COCA)

As we can see, token frequency rises steadily over time. It can mean that word forms with *über-* are becoming more and more widespread in English. On the other hand, P* shows that the prefix *über-* was and still is very productive in English, although in the last few years its productivity has been decreasing. Nowadays it is not as productive as it was between 1990 and 1998.

Following from these data, we arrive at the question whether there is a difference between the change in the productivity of adjective and that of noun bases diachronically. Table 5 and Table 6 below represent separate productivity indexes for these two sets of words. They show that the potential productivity of adjectives and nouns with the prefix *über-* is getting smaller and smaller. Über-constructions are still used, but only a small number of new forms are built. This fall is attested especially among nouns.

Adjectives	<1998	1998-2002	2003-2007	2008-2012
N (C)	3	10	31	26
P*	1.00	0.80	0.81	0.65

Table 5: Diachronic measures of productivity among adjectives

Nouns	<1998	1998-2002	2003-2007	2008-2012
N (C)	18	47	45	77
P*	1.00	0.87	0.64	0.27

Table 6: Diachronic measures of productivity among nouns

Before the meaning of the prefix *über-* in English is discussed, it is advisable to point out some basic meanings in German. Due to limitations of space, only a short and sketchy grammatical and semantic description can be provided about *über-* in German. Besides an adverbial, adjectival and prepositional function, *über-* can occur as a prefix as well. It can be attached to nouns, adjectives and verbs. On the basis of the Duden online dictionary, the following distinct meanings can be differentiated. In cases like *Überkontrolle* 'too much control' or *Übersubventonierung* 'too much subsidization', the prefix expresses the sense that something is too much. In *Überministerium*, however, it refers to superiority. If it is attached to adjectives as in *überdeutlich* 'absolutely obvious' it functions as an intensifier. Attached to verbs, it can occur as an inseparable verb prefix or as a separable verb particle. It is sometimes not easy to make a distinction between these two types of verbal constructions in German.

If we have a look at the list of the examples taken from COCA it can be stated that the prefix *über-* carries two meanings. It either functions as an augmentative prefix as in example (5) and (6) and conveys the meaning 'large' or 'great', or it refers to superiority as in (7) and (8). If it is attached to adjectives, it has an augmentative meaning and in the case of nouns it involves superiority.

(5) President Bush will have to decide between continuing his futile war policy or taking money from his <u>uber-rich</u> buddies.

(6) So sounded a query at one of the first panels at YearlyKos, the second annual terrestrial meeting this August of devotees of the <u>uber-liberal</u> DailyKos website.

(7) Still, people in the fashion world will probably be interested in the (endless) complaints of the fashionable Andrea, since the book's author, Lauren Weisberger, used to be the assistant of Vogue's <u>uber-editor</u> Anna Wintour.

(8) I think that it seems to ebb and flow, this concept of having a systemic risk regulator. This would be sort of an <u>uber-regulator</u>, overseeing the others.

Albair's (2010) research focuses on the comparison of thirteen evaluative affixes. His paper comprises both augmentative prefixes like *hyper-*, *mega-* or *tera-* and diminutive prefixes such as *micro-* or *nano-*.

The last point raises the issue of comparing the productivity of über-constructions with that of other augmentative suffixes. I have carried out a further study in COCA to investigate it. The productivity of adjectival bases with augmentative prefixes has been compared with the bases taken from recent texts between 2005 and 2012. Six augmentative prefixes have been studied: *über-*, *super-*, *mega-*, *ultra-* and *hyper-*. Table 7 demonstrates the results. As can be seen from these results, *über-* shows the highest rate of productivity in comparison with the other augmentative prefixes. However, in terms of token frequency it has the lowest index among them. It reveals that über-constructions are rare to find, but they are formed highly productively.

	N(C)	P*
uber-	58	0.64
super-	11876	0.03
mega-	515	0.19
ultra-	1513	0.15
hyper-	921	0.17

Table 7: Productivity of augmentative prefixes in English

As Baayen (2009: 903) points out, "potential productivity is highly sensitive to markedness relations". It means that the higher the index the more marked the process. Thus, based on the results in table 7 it can be claimed that *über-* can be seen as a marked form compared to other augmentative suffixes in English.

3.2 Data analysis in Hungarian

3.2.1 Research data and methods

In the second part of the research I have used a Hungarian webcorpus (Halácsy et al. 2004) containing 1.5 billion words unfiltered (600 million words fully filtered) without annotations. The corpus consists of texts downloaded from the .hu domain, thus representing common written Hungarian fairly extensively. Texts that were present multiple times and files which contained no useable text were filtered out.

Unlike in COCA, a string search function is not provided, therefore the texts had to be searched one by one partly manually and a list of über-constructions has been compiled (see Appendix). All of the instances of *übermensch* have been omitted as in the COCA list. Furthermore, it could be argued that the word *über-realtime-cool* is a loan translation, not a genuine Hungarian formation, so it was deleted from the Hungarian data.

Unfortunately, because of the lack of annotations, search options and any additional information such as genre of the different texts, period of time we could not conduct as many queries as in COCA.

3.2.2 Results

As an answer to question 1, the following type and token frequencies have been found: 48 tokens out of which there were 15 noun tokens and 33 adjectives. The list contained 41 different types including 15 noun types and 26 adjectives. The most frequent bases were *fasza* (4), *okos* (2), *macsó* (2) and *frankó* (2). Hence, *über-* is primarily attached to adjectives in Hungarian.
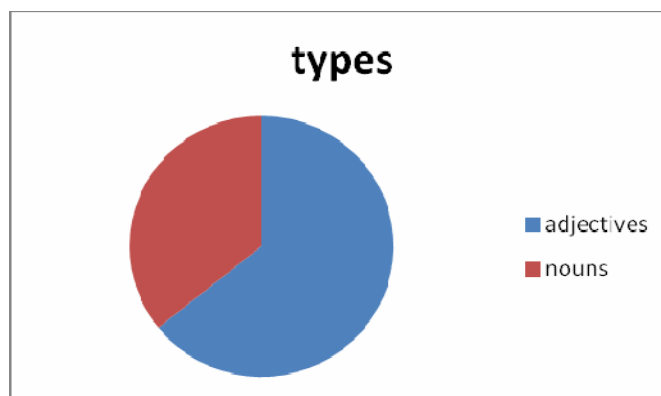
Fig. 4: Distribution of noun and adjective bases in Hungarian webcorpus

As far as spelling is concerned, the list shows clearly that all of the words are spelled as one word only with three exceptions: *über csilivili* is written in two separate words and *über-szuper* and *über-realtime-cool* with a hyphen. One-word spelling can be explained either by the fact that in German such expressions are always written as one single word, or by the fact that Hungarian tends to write augmentative prefixes together with the base, for example *szuper-nagyi*.

Unfortunately, the Hungarian webcorpus does not include data on genres or registers. Texts have been collected from the internet, which involves a wide range of genres. However, if we have a closer look at the examples extracted, we can make some assumptions. As a native speaker of Hungarian, I suppose that mainly blogs and comments made on the internet were the main sources of über-constructions. Here are some further instances to support this claim:

(9)   S úgy vélte, volna itt valami helyrekalapálni-való, az <u>überdemokratikus</u> rizsa jegyé-ben.

(10)   Főhősünk most is az <u>übermacsó</u> kategóriából kerül ki, azonban ellentétben elődjével van neve (ha ez netán még nem derült volna ki :-), s a párbeszédek közben képes beszélni is, ezt Max Payne-re hajazó elődje nem mondhatta el magáról.

(11)   … pedig nem vagyok <u>übernegatív</u> ember,de mióta suli van,még anniyra se tudom el-viselni …

Based on example (9), (10) and (11), it can be stated that the typical register of the texts in which the prefix *über-* occurs in Hungarian is slang. This claim can be supported by the fact that Kövecses' (2002) dictionary of slang lists for example *überokos* as a slang expression.

As in English, the productivity measure was P*. Table 8 illustrates the results.

token frequency	N(C)=48
hapax legomena	V(1,C,N)=36
P*	V(1,C,N)/N(C)=36/48=0.75

Table 8: Productivity index in Hungarian webcorpus

Out of 48 tokens 36 words were hapax legomena. This ratio in itself reveals that *über-* is widely used in Hungarian to coin new words. P*= 0.75 is indicative of high morphological productivity.

From a semantic point of view, *über-* conveys an augmentative meaning in all of the cases in Hungarian. For instance, *überférfi* means 'a great man' or *übererős* means 'extremely strong'. There is no example of other kinds of meaning.

3.3   Comparison

After having carried out research in COCA and the Hungarian webcorpus, the results need to be compared.

Generally, the type and token frequencies differ in the two languages. In the English data set altogether 257 tokens were found, whereas in Hungarian only 48. Even if the two corpora are different from each other, it can be figured out that *über-* is a more common and well-established form in English. In Hungarian it occurs more marginally. This can be supported by the fact that the online *Oxford Advanced Learner's Dictionary* has a separate entry for *über-*, while *Magyar Értelmező Kéziszótár* [Hungarian Explanatory Dictionary] (2003) does not include it at all. Probably, it is due to the fact that online OALD is regularly updated and more recent than the Hungarian dictionary published in 2003.

If type frequencies are compared, it can be seen that English has far more types as well. English had 170 different über-constructions, whereas Hungarian only 42. It can be inferred that not only the overall number of occurrences of über-constructions is higher in English than in Hungarian, but also there are far more different types as well.

The ratio of adjective and noun bases is also different. Comparing the English noun list with the Hungarian noun list and the English adjective list with its Hungarian counterpart we can see that in COCA 70% of the types were nouns. In Hungarian this ratio is only 36%. In English predominantly nouns can take the prefix *über-*, while in Hungarian *über-* is attached rather to adjectives.

Concerning spelling there is no big difference. In English über-constructions are most commonly spelled with a hyphen or as one word. There is a growing trend towards single-word spelling. In Hungarian, however, they are spelled as one word nearly without any exceptions.

Also the distributions of genres seem to be similar. In both languages *über-* is preferred in informal text types. Unfortunately, in Hungarian only the native speaker's intuition is available and suggests that.

Table 9 demonstrates the number of hapax legomena and the productivity rates in COCA and the Hungarian webcorpus.

	COCA	Hungarian webcorpus
hapax legomena	140	36
P*	0.54	0.75

Table 9: Comparison of productivity indexes

In English more hapax legomena were found than in Hungarian. It is still Hungarian that has a higher productivity index. On the basis of the diachronic results, in English the productivity

of *über-* seems to be decreasing. In Hungarian where there are so many new forms with *über-*, it appears to be increasing.

Finally, we should compare the meanings. In English, two basic meanings can be attested. In adjectives *über-* conveys an augmentative meaning, while in nouns it refers to superiority. In contrast, in Hungarian *über-* can have only an augmentative meaning.

4   Summary and conclusion

In this paper, I have presented a cross-linguistic study of the prefix *über-*. Data were collected from COCA and a Hungarian webcorpus and they were investigated with a close look at the type and token frequencies, the possible bases, the spelling, the genres, the productivity indexes and meaning. The results of the present study clearly proved that *über-* is a highly productive prefix in both languages although diachronically it shows a downward trend in English. The direction of this development in Hungarian needs to be investigated further.

References

Albair, Joshua (2010): What is the State of Evaluative Affixes in Contemporary English? *Debut Journal Volume 1 Number 1 Spring 2010.* downloaded from http://www.llas.ac.uk/resour cedownloads/3088/debut_vol_1_albair.pdf.

Aronoff, Mark (1976): *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

Baayen, R. Harald (1989): *A corpus-based approach to morphological productivity: statistical analysis and psycholinguistis interpretation.* Unpublished doctoral dissertation. Free University of Amsterdam.

Baayen, R. Harald (1992) Quantitative aspects of morphological productivity. In: Booij, G. & van Marle, J. (eds.): *Yearbook of Morphology 1991*. Dordrecht: Kluwer, 109-149.

Baayen, R. Harald (1993): On frequency, transparency, and productivity. In: Booij, G. & van Marle, J. (eds.): *Yearbook of Morphology 1992*. Dordrecht: Kluwer, 181-208.

Baayen, R. Harald (1994): Productivity in production. *Language and Cognitive Processes* 9, 447-469.

Baayen, R. Harald (2009): Corpus linguistics in morphology: morphological productivity. In: Lüdeling, A. & Kyto, M. (eds.): *Corpus Linguistics. An international handbook*. Berlin: Mouton De Gruyter, 900-919.

Baayen, R. Harald & Rochelle Lieber (1991): Productivity and English derivation: a corpus based study. *Linguistics* 29, 801-844.

Bauer, Laurie (1992) Scalar productivity and -lily adverbs. In: Booij, G. & van Marle, J. (eds.): *Yearbook of Morphology 1991*. Dordrecht: Kluwer, 185-191.

Bauer, Laurie (2001): *Morphological Productivity*. Cambridge: Cambridge University Press.

Booij, Geert (2002): *The Morphology of Dutch.* Oxford: Oxford University Press.

Chomsky, Noam (1965): *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Dressler, Wolfgang U. (1997): On productivity and potentiality in inflectional morphology. *CLASNET Working Papers* 7, 2-22.

Halácsy, Péter, Kornai, András, Németh, László, Rung, András, Szakadát, István & Trón, Viktor (2004): Creating open language resources for Hungarian. In: *Proceedings of the 4th international conference on Language Resources and Evaluation (LREC2004).*

Hay, Jennifer & Baayen, R. Harald (2002): Parsing and productivity. In: Booij, G. & van Marle, J. (eds.): *Yearbook of Morphology 2001.* Dordrecht: Kluwer, 203-235.

Hocket, Charles F. (1958): *A Course in Modern Linguistics.* New York: Macmillan.

Jespersen, Otto (1942): *A Modern English Grammar on Historical Principles, Part VI: Morphology.* London: George Allen and Unwin; Copenhagen: Munskgaard.

Juhász, József & Pusztai, Ferenc (2003): *Magyar Értelmező Kéziszótár.* Budapest: Akadémiai Kiadó.

Kastovsky, D. (1986): The problem of productivity in word-formation. *Linguistics* 24, 585-600.

Kiss, Katalin (2011): *Contributions to a semantic-contrastive analysis of verb-particle constructions in English and verbs with coverbs in Hungarian.* Dissertation. University of Debrecen.

Kövecses, Zoltán (2002): *Magyar szlengszótár.* Budapest: Akadémiai Kiadó.

Ladányi, Mária (2001): Szempontok a morfológiai produktivitás megállapításához. In: Bakró-Nagy M., Bánréti Z. & É. Kiss K. (eds.): *Újabb tanulmányok a strukturális Magyar nyelvtan és a nyelvtörténet köréből. Kiefer Ferenc tiszteletére barátai és tanítványai.* Budapest: Osiris, 232-246.

Lieber, Rochelle (1992): *Deconstructing morphology.* Chicago: University of Chicago Press.

Meibauer, Jorg, Guttropf, Anja & Scherer, Carmen (2004): Dynamic aspects of German -er nominals: a probe into the interrelation of language change and language acquisition. *Linguistics 42*, 155-193.

Plag, Ingo (2004): Productivity. In: *Encyclopedia of Language and Linguistics.* 2nd ed. Amsterdam: Elsevier, 1-26.

Plag, Ingo (2006): Productivity. In: Aarts, B. & McMahon, A. (eds.): *Handbook of English Linguistics.* Oxford: Blackwell, 537-556.

Plag, Ingo, Dalton-Puffer, Christiane & Baayen, R. Harald (1999): Productivity and register. *Journal of English Language and Linguistics* 3, 209-228.

Rainer, Franz (1987): Produktivitätsbegriffe in der Wortbildungstheorie. In Dietrich, W., Gauger, H-M. & Geckeler, H. (Hrsg.): *Grammatik und Wortbildung romanischer Sprachen.* Tübingen: Narr 187-202.

Scherer, Carmen (2005): *Wortbildungswandel und Produktivität. Eine empirische Studie zur nominalen ‚-er'-Derivation im Deutschen.* Tübingen: Niemeyer.

Online sources:

http://corpus.byu.edu/coca/

http://www.duden.de/suchen/dudenonline/%C3%BCber

http://oald8.oxfordlearnersdictionaries.com/dictionary/uber

Melinda Majtényi
University of Debrecen
Institute of English and American Studies
Pf. 73
H-4010 Debrecen
melinda.majtenyi@gmail.com

Appendix

1   List of types with *über-* from COCA

nouns

1 uber alarm 1
2 uberagent 1
3 uber-american 1
4 uber-athletes 1
5 uber-babe 2
6 uberbabies 1
7 uber-baddie 1
8 uber bar 1
9 uber billionaire 1
10 uber-bitch 1
11 uber-boss 1
12 uber-breed 1
13 uber-brit 1
14 uber-cafe 1
15 uber celebrity 1
16 uberchallenge 1
17 uber-character 1
18 uber chick 2
19 uberclub 1
20 uber-computer 1
21 uber cup 1
22 ubercutter 1
23 uber design 1
24 uberdesigner 1
25 uber-director 1
26 uberdose 1
27 uber DVD 1
28 uber edition 1
29 uber-editor 2
30 uber-electronics 1
31 uberenergy 1
32 uberexperts 1
33 uberfan 1
34 uberfashion 1
35 uberfemale 1
36 uber-fitness 1
37 uberfood 1
38 uber-foodies 1
39 uber football 1
40 uber-fundraiser 2
41 uber geek 3

42 uber-green house 1
43 uber-guitarist 1
44 uberhiker 1
45 über hit 1
46 uber-hobbies 1
47 uber hostess 1
48 uber-hunk 1
49 uber-icon 1
50 uber id figure 1
51 uber-idiot 1
52 uber-idiotism 1
53 uber jam 1
54 uber-jew 2
55 uber-keynesians 1
56 uber-lawyers 1
57 uber-lizard 1
58 uberloser 1
59 ubermales 1
60 uberman 10
61 uber-marionette 1
62 uber-mayor 1
63 uber millionaire 1
64 uber model 4
65 uber-mogul 1
66 uber-moisurizer 1
67 uber-movement 1
68 ubernanny 47
69 uber national security
   adviser 1
70 ubernetwork 1
71 uber-orcs 1
72 uber-overarchievers 1
73 uber-overdog 1
74 uber-parents 1
75 uber-patriotism 1
76 uber-phone 1
77 uberpopularity 1
78 uber-portal 1
79 uber premium 1
80 uber producer 3
81 uber-prospect 1

82 uberrace 1
83 uber-regulator 2
84 uber-roundedness 1
85 ubersatori 1
86 uber sausage 1
87 uber sauvage 2
88 uber-science 1
89 uber-scientist 1
90 uber-screens 1
91 uberscribe 1
92 ubersexuality 1
93 ubersexuals 1
94 uber shareholder
   activist 1
95 uber-slackers 1
96 uber-soccer 1
97 uber-soldiers 1
98 uber soprano 1
99 uber sports 1
100 uber-storm trooper
   1
101 ubersuburb 1
102 uber-success 1
103 ubertechnocrats 1
104 uber-teen 1
105 uber-terrorist 1
106 uber-thug 1
107 uber-trend 1
108 uber-trendiness 1
109 uber-ubiquity 1
110 uber-users 1
111 ubervenue 1
112 ubervirtue 1
113 uber-wafflehaus 1
114 uberwaif 1
115 uber-wanker 1
116 uberwoman 1
117 uber-yang 1
118 uber-yin 1

adjectives

1 uber-active 1	45 uber-sweet 1
2 uber-aggressive 1	46 ubertalented 1
3 uber-athletic 1	47 uber-thinky 1
4 uber-back 1	48 ubertrendy 2
5 uber-casual 2	49 uber-urban 1
6 uber-catholic 1	50 uber-wealthy 2
7 uberchic 2	51 uber-western 1
8 uber-closer 1	52 uberwholesome 1
9 uber-competitive 1	
10 uber-conditioning 1	
11 uberconfident 1	
12 uber cool 1	
13 ubercrushing 1	
14 uber dapper 1	
15 uber-dark 1	
16 uber-driven 1	
17 uber dry 1	
18 uber-energizing 1	
19 uber entitled 1	
20 uberexpensive 1	
21 uber famous 2	
22 uberfashionable 1	
23 uberfast 2	
24 uber-heterosexual 1	
25 uberhip 5	
26 uber-hot 1	
27 uber-hyped 1	
28 uber-integrated 1	
29 uberintense 1	
30 überintensive 1	
31 uber Kobachan 1	
32 uberliberal 4	
33 uberlight 2	
34 uber-observant 1	
35 uber-popular 3	
36 uber-posh 1	
37 uber-rich 2	
38 uber-ripe 1	
39 uber-sexy 2	
40 ubersoft 1	
41 uber-soulful 1	
42 uber-stud 1	
43 uberstylish 1	
44 uber-successful 1	

2   List of types with *über-* in Hungarian webcorpus

nouns
1 überállat 1
2 überbefektetés 1
3 übercsónak 1
4 überférfi 1
5 überfíling 1
6 übergebe 1
7 überindividualizmus 1
8 überkanca 1
9 überlobbista 1
10 übermozi 1
11 überpápa 1
12 übersztárköltő 1
13 übervélemény 1
14 übervilág 1
15 überziher 1

adjectives
1 tökszuperübercsászárkirályágos 1
2 überciheres 1
3 überciki 1
4 über csilivili 1
5 überdemokratikus 1
6 überdízel 1
7 überdomináns 1
8 übererős 1
9 überfasza 4
10 überfrankó 2
11 überfüggetlen 1
12 übergagyi 1
13 übergonosz 1
14 überhochdeutsch 1
15 überintenzív 1
16 überkirály 1
17 übermacsó 2
18 übermájer 1
19 übernegatív 1
20 überokos 2
21 überpláne 1
22 überpolitikusi 1
23 überprimitív 1
24 überrulez 2
25 über-szuper 1
26 übervicces 1