

Exploring the meaning and productivity of a polysemous prefix

The case of the Modern Greek prepositional prefix *para-*

Angeliki Efthymiou

Democritus University of Thrace, Greece
aefthym@eled.duth.gr

Georgia Fragaki

University of Athens, Greece
efraga@phil.uoa.gr

Angelos Markos

Democritus University of Thrace, Greece
amarkos@eled.duth.gr

Abstract: This paper follows a corpus-based approach to the meaning and productivity of the Modern Greek prepositional prefix *para-*. A semantic categorization of the prefix is proposed and its productivity is measured across semantic categories, registers, text types and grammatical categories. *Para-* was found to be more productive in non-locational and evaluative meanings. Its most productive meaning is excess, while the locational meaning of proximity still remains strong. It is also more productive in written than spoken registers and the grammatical category of nouns. The findings of the study can have implications about the prefix's ongoing grammaticalization and its affixal status.

Keywords: Modern Greek; prepositional prefix; productivity; corpus-based study; polysemy

1. Introduction

This paper aims at exploring the meaning and productivity of the Modern Greek prepositional prefix *para-*, which is used to create adjectives (e.g., *para-kratikós* 'extra-governmental'), nouns (e.g., *par-onímio* 'by-name'), adjectival (e.g., *para-zalizménos* 'bedazzled') and adverbial participles (*para-pléondas* 'sailing by'), verbs (e.g., *para-káno* 'to overdo') and adverbs (e.g., *para-ékso* 'further out'). Our study follows similar research on other languages, which investigates a probabilistic aspect of morphological

productivity with a corpus-based methodology (e.g., Baayen & Lieber 1991; Baayen 1992; Plag et al. 1999; Gaeta & Ricca 2003). According to this approach, productivity can be operationalized by the concept of vocabulary growth, i.e., how frequently new word types that are formed by a morphological process are to be expected to appear when an increasing amount of text is sampled. This methodology has been followed in a study of Greek verb-forming suffixes (Efthymiou et al. 2012) and in a recent work on the meaning and productivity of *iper-*, another Modern Greek prepositional prefix (Efthymiou et al. 2014; 2015). The data used for the identification of meanings, as well as the measurement of the productivity of *para-* comes from the Corpus of Greek Texts (CGT), a reference corpus of Modern Greek, including approximately 28 million words from a variety of spoken and written text types.

There has been no extensive research on the productivity of the prefix or the various meanings associated with *para-*. The relevant literature on the semantics of *para-* is mostly concerned with the so-called intensifying meaning of the prefix (cf. Delveroudi & Vassilaki 1999; Efthymiou 2003; Gavriilidou 2013). In these studies the meaning and use of *para-* is analyzed along with other prefixes or items of “intensity”. There is also research concerning the semantics of the preposition *pará* and, in particular, its drifting from locational to non-locational meanings (e.g., Pouloupoulou 2002). These remarks on the semantics of the preposition *pará* can be useful in studying the meanings of the respective prepositional prefix. Equally useful are studies on the diachrony of the prefix (e.g., Karantzola & Giannouloupoulou 2000; Markopoulou 2014) and its morphological status (Anastassiadi-Symeonidi 1986; Ralli 2004). Significantly, there is a lack of corpus-based research on the field, which can provide a less biased view on the semantics and productivity of the prefix on the basis of authentic linguistic interaction.

This paper attempts to enrich the literature on the prefix *para-*, first, by identifying the different meanings of the prefix on the basis of authentic data from a large corpus of Greek and, second, by addressing the following research questions:

- (1) Is the prefix equally productive across (a) different meanings (semantic categories) (b) different grammatical categories of the derived items and (c) across spoken and written registers and text types of the corpus?
- (2) Are there any associations between the grammatical categories of the derived items and the meanings of *para-* in terms of productivity?
- (3) Are there any associations between the registers and the text types in which the derived items occur and the meanings of *para-*?

The following section is dedicated to a brief overview of the relevant literature of the prefix under investigation. In section 3 the data and the methodology used in this study are outlined. The results of the research are presented in section 4 and further discussed in section 5. Finally, section 6 summarizes the findings and draws the implications of the study.

2. The Modern Greek prepositional prefix *para-*

Para- mainly derives from the Ancient Greek preposition *pará* ‘close to (but falling short of)’ (Bortone 2010, 291), but, in the course of its grammaticalization into a prefix, it has also developed some additional non-locational meanings (e.g., temporal continuity: *para-méno* ‘to stay on, to remain’, parallelism: *pará-dromos* ‘side road’, violation or divergence: *par-erminévo* ‘to misread’, *pará-tipos* ‘irregular’, excess: *para-cimáme* ‘to oversleep’; Anastassiadi-Symeonidi 1986; Pouloupoulou 2002; Ralli 2004; Markopoulou 2014, cf. Schwyzer 1950). It is interesting that *para-* does not only exhibit extensive polysemy, expressing both locational (e.g., *para-thalásios* ‘seaside’) and non-locational meanings (cf. above; Triandafyllidis 1998; Babinotis 2002 and section 3), but was already used both as a preposition and a prepositional prefix in Ancient Greek (cf. Liddell & Scott 1961; for details on other locational prefixes that correspond to prepositions, see e.g., Talmy 2000; Tyler & Evans 2001; Lieber 2004; 2005).

According to Schwyzer (1950) and Markopoulou (2014), the main process involved in the semantic change of *para-* is metaphor, i.e., a semantic process, which is characteristic of grammaticalization (Heine et al. 1991, 79–86). Of special interest is the development of the “more subjective or evaluative” meanings (in the sense of Traugott & Dasher 2002) of *para-*, i.e., the meaning of violation or divergence and the meaning of excess. According to Markopoulou (2014), these figurative meanings appear in a few words already in the classical period, but become productive in Hellenistic and Medieval Greek, as well as later, in Modern Greek.

Furthermore, in contemporary Greek the productivity of *para-* seems to be further increased through its use in neological formations, especially in the semantic domain of excess, as well as in neological loan translations, especially in scientific and technical domains (e.g., *para-stratiotikós* ‘paramilitary’ *para-plirofórisi* ‘misinformation’).¹ Interestingly enough, the

¹ Cf. Triandafyllidis (1998); Anastassiadi-Symeonidi (1986); Ralli (2004); Markopoulou (2014). For a discussion of the diachronic development of affixal polysemy, see e.g., Rainer (2005; 2009).

productivity of *para-* seems to be expanding both in the non-locational semantic domains of diversity or excess and in the locational meaning of proximity (e.g., *paréŋçima* ‘parenchyma’; cf. Triandafyllidis 1998; Markopoulou 2014).

Given the recent development of the meaning of excess and the difficulty in reconciling the wide range of meanings associated with *para-*, Triandafyllidis (1998), following Anastassiadi-Symeonidi’s (1986) proposal, suggests two different homonymous lemmas, one for *para-1*, and another for the “excessive” *para-2*. More specifically, *para-1* and *para-2* are treated in the dictionary as homonymous, because they go back to different etyma; *para-1* derives from the ancient Greek preposition *pará* and *para-2* derives from the “intensifying” adverb *pára*. According to the etymological part of the same dictionary, *pára* also originates from the ancient Greek preposition *pará*.

Although such an analysis might seem reasonable from a strictly synchronic perspective, it could be argued that it overlooks the fact that there are recurrent patterns of derivational polysemy, which may be observed in numerous languages and can also partly apply to *para-*. A common case is that of intensifying affixes which have their origin in locative affixes expressing the transgression of a limit (cf. Eng. *extra-*, *ultra-*, *over-*, MG *iper-*, *kse-*, *apo-*; for discussion see e.g., Amiot 2012; Bauer et al. 2013; Rainer 2009). Along the same line of reasoning, Ralli (2002; 2004) proposes a unitary treatment of *para-* (cf. Pouloupoulou 2002). In particular, Ralli (2002) claims that there is a single *para-*, which expresses the basic notion of proximity or parallelism and suggests that this notion may trigger an excessive meaning.²

In the rest of the paper, following Ralli (2002; 2004), we adopt a unitary analysis and regard *para-* as one polysemous prefix. Nevertheless, we believe that further research is needed in order to account for the distribution of *para-1* and *para-2*. For instance, a problem with the unitary analysis is the fact that *para-* (i.e., *para-2*) seems to be excessive only when it derives verbs (and adjectival participles) on verbal bases (cf. Efthymiou 2003; Gavriilidou 2013).

Furthermore, the picture of *para-* becomes even more complicated if we consider the following fact. As regards verbal formations with *para-*, according to Ralli (2004), *para-* displays a dual character, since it behaves like an internal prefix (in the sense of Di Sciullo 1997) when it affects the

² For a different analysis, which considers the development of the meaning of excess as an extension of the meaning of violation or divergence, cf. Pouloupoulou (2002); Markopoulou (2014).

meaning of the base and like an external prefix when it brings only external specifications to it. More specifically, she claims that *para-* is internal when it expresses the basic idea of proximity or parallelism to the meaning of the base and external in its excessive meaning. Moreover, she states that the excessive *para-* is highly productive in Modern Greek and can be added at the left-hand edge of the parasynthetic verbs (e.g., *para-apovlakóno* ‘to make extremely stupid’). Finally, she observes that an obligatory vowel deletion is attested when *para-* is internal (e.g., *par-érxome* ‘to elapse, to pass by’) and an optional vowel deletion when *para-* is external (e.g., *para-éxo* ‘to over-have’). In simple terms, we believe that these facts raise an interesting theoretical issue: on the one hand, the semantic uniformity and robust productivity of the excessive *para-*, as well as its “loose” relation with its base, may be taken as evidence that *para-* has evolved into two distinct homophonous prefixes (i.e., the Triandafyllidis Dictionary suggestion). On the other hand, as already said, the cross-linguistically frequent polysemic constellation of both locative and quantitative meanings can suggest a polysemy analysis.

The examination of the morphological status of *para-* has raised another interesting theoretical issue. Given that *para-* derives from the Ancient Greek preposition *pará*, its morphological status is often described as ambiguous and the formations in which it participates can be considered as either compounds or derivatives.³ For instance, in most traditional grammars (cf. Triandafyllidis 1991 [1941]) the combination of an Ancient Greek preposition with a base is interpreted as part of a compounding process (see Ralli 2013, 124–126 for discussion). On the other hand, most linguists agree that Modern Greek prepositional prefixes should be synchronically analyzed as affixes (see e.g., Philippaki-Warbuton 1970; Malikouti-Drachman & Drachman 1992; Smirniotopoulos & Joseph 1998; Ralli 2004). The basic arguments for treating Modern Greek prepositional prefixes as affixes are the following: a) the use of the respective prepositions is mostly restricted to fixed or archaic expressions, b) Modern Greek prepositions are always free (e.g., *ja* ‘for’, *me* ‘with’) and c) the meaning of all prefixed verbs is not always compositional.⁴

More specifically, as regards the properties of the bound form *para-*, it seems that the MGr preposition *pará* is now only used with non-locational meanings (except for a few fixed phrases; e.g., *pará to prothipurýo* ‘close to

³ For discussion, see e.g., Ralli (2004).

⁴ For a discussion on the distinction between compounding and derivation see e.g., Iacobini (1999); Amiot (2005); Ralli (2007).

prime-minister’); its locational meanings are taken over by *kondá* (cf. Bortone 2010). In addition, *para-* still possesses some of its original Ancient Greek properties (i.e., the locational meaning), but has developed new meanings that do not correspond to the meanings expressed by the AG preposition. Thus, its meanings do not correspond to the homomorphous AG or MGr preposition (also see section 3). Furthermore, there is a large number of *para-* derivatives with a highly lexicalized meaning (*paravjéno* ‘to compete’ < *para+vjeno* ‘to go out’; cf. Ralli 2004; Markopoulou 2014). Finally, earlier *para-* combines with adjectival, nominal, verbal and adverbial bases (Markopoulou 2014), whereas the Modern Greek preposition only combines with nouns.

Most of the above properties can be used as criteria for assigning *para-* the functional status of derivational prefix. At the same time, the fact that *para-* still possesses some of its original Ancient Greek properties (i.e., the locational meaning), may indicate that its grammaticalization is not complete and that in the native speaker’s mind the meaning of the bound element can still be associated with that of the independent word (see e.g., Ralli 2004; Booij 2005; Amiot 2005). Interestingly, this last property of *para-* seems to support Ralli’s (2007) claim that Modern Greek prepositional prefixes are closer to the category of derivational prefixes than to the category of bound stems, but have not yet acquired full affixal status.

3. Data and methodology

The morphological productivity of *para-* was studied by using the *Corpus of Greek Texts* (CGT), a reference corpus of Modern Greek, which is synchronic (it contains texts from 1990 to 2010) and comprises a variety of spoken and written text types of approximately 28 million words. (For more details on the corpus, see Goutsos 2010). Specific information on the size of the text types included in the CGT is given in Table 1.

As can be seen in Table 1, spoken data, i.e., news, interviews, public speeches and everyday conversations, constitute 10.51% of the CGT, while written data, which include literature, journalistic texts such as news and opinion articles, information items, academic texts, popularized non-fiction texts, law and administration, private and procedural texts, etc., cover 89.49% of the corpus. There is obviously considerable imbalance both between the size of spoken and written data and that of the specific

Table 1: Size of text types in the *Corpus of Greek Texts*

Mode	Text type	Number of words	Percentage
Spoken	News	291,382	1
	Interview	592,584	2
	Public speech	1,839,766	6.75
	Conversation	207,548	0.76
	Total	2,931,280	10.51
Written	Literature	2,455,080	9
	News	4,764,337	17.5
	Opinion articles	3,189,132	11.7
	Information items	100,570	0.36
	Academic	3,994,277	14.67
	Popularized non-fiction	7,648,513	28
	Law and administration	1,472,700	5.4
	Private	186,210	0.68
	Procedural	145,770	0.53
	Miscellanea	335,906	1.65
	Total	24,292,495	89.49

text types represented in the corpus,⁵ which, however, does not affect the results of *para-*'s morphological productivity across text types, since the methodology used in its measurement can overcome this difficulty (see below). However, the small size of some text types proved to be a significant obstacle that was not easy to overcome. This is why the morphological productivity of *para-* across text types can only be measured for 10 text types (out of 14) in the CGT, while no measurements are possible for the spoken text types of news and conversation and the written text types of private and procedural texts.

Our study is corpus-based, since it starts from an already known item, *para-*, and seeks examples of its use and quantitative evidence in the CGT.⁶ First, a search for $\pi a q^*$ (par^*) was made in CGT, by using Concord in WordSmith Tools 5.0 (Scott 2008). This search provided us with concor-

⁵ It has to be noted that we have used the best available resource for our purposes, since the CGT contains the greatest number of spoken texts and the widest variety of text types than all other reference corpora of Greek (cf. Goutsos 2010, 30).

⁶ For the terms corpus-based and corpus-driven, see e.g., Tognini-Bonelli (2001, 17).

dances, i.e., lines of text containing the item queried (called **node word**) in its close context, including all tokens starting with *par-*, regardless of what follows.⁷ The Concord tool offers information on the total number of tokens found and the name of the file in which tokens are found. The latter can help us identify the text type in which tokens are used, since this information is coded in the CGT.

The result of this first search in the corpus was 238,103 tokens. These data were then cleared, according to the following criteria. All irrelevant items, consisting of occurrences of *pará* as a preposition (10,788 tokens), names of persons, mountains, monuments etc. (e.g., *parácelsos* ‘Paracelsus’, *parnasós* ‘Parnassus’), loan words (e.g., *parás* ‘money’: Turkish, *parádisos* ‘heaven’: Persian, *parkáro* ‘to park’: Italian) and other irrelevant items (e.g., *parðalós* ‘multicolour’, *parθénos* ‘virgin’) were excluded (cf. footnote 7). Fully lexicalized items whose base does not exist or whose meaning cannot be synchronically detected (e.g., *par-álios* ‘by the sea’, *para-tsúkli* ‘nickname’, *par-usjázo* ‘to present’, *par-oçiménos* ‘obsolete’) were also removed. Finally, tokens which are derived by items already prefixed by *para-* (e.g., *pará-dosi* ‘delivery’ > *para-dído* ‘to deliver’) were not included in the final list of *para-* tokens in this study (see e.g., Plag 1999; Plag et al. 1999; Gaeta & Ricca 2003; Fradin et al. 2008; Efthymiou et al. 2012). It is interesting to note that most of the types which were excluded were high frequency items, something which accounts for the comparably low number of tokens finally used in the study.

After the application of these criteria, the data were reduced to less than a quarter of the original data (238,103 tokens), i.e., to 53,596 tokens. The 53,596 tokens were manually lemmatized, resulting in 553 lemmas, and the number of hapax legomena (i.e., types that occur only once in the corpus) was counted. The tokens were classified into grammatical categories, i.e., verbs, nouns, adjectives, adverbs, adjectival participles⁸ and adverbial

⁷ The corpus is not tagged for morphological categories and as a result the outcome of the search is not restricted to types prefixed by *para-*, but also includes types starting with a homophonous string of segments which are completely different from the investigated prepositional prefix and thus had to be excluded.

⁸ This category includes types of three genders, ending in *-ménos-i-o*, *-ómenos-i-o*, *-ón-úsa-ón*, *ís-ísa-én* and *-as,-asa,-an*. The vast majority of them end in *-ménos-i-o* and are rarely found in the so-called periphrastic types of the present perfect and past perfect in the passive voice (*íme* ‘to be’ + past participle). However, the inclusion of these periphrastic types in the tense system of Greek is controversial. According to several researchers, these structures are mainly stative predicates, which means that they are not principally used for tense distinctions (see, among else, Moser 1994, 140ff). In general, the types of adjectival participles in *-menos-i-o* are considered to

participles.⁹ The occurrence of the 53,596 tokens in spoken and written data, as well as across the 14 text types of the corpus, was also found.

In addition, the meaning of the 553 lemmas prefixed by *para-* and their use in the texts were studied, resulting in a classification of 10 semantic categories. Dictionaries of Modern Greek (Triandafyllidis 1998; Babinotis 2002) and relevant research (e.g., Pouloupoulou 2002; Eftymiou 2003; Amiot 2004; Kennedy & McNally 2005; Paradis 2008; Gavriilidou 2013) were consulted. *Para-* meanings were classified first into LOCATIONAL and NON-LOCATIONAL meanings. The former include the meanings of PROXIMITY (e.g., *para-thalásios* ‘seaside’) and BEYOND SPATIAL BOUNDARY (e.g., *para-páno* ‘above, further up’), while the latter are divided into NON-EVALUATIVE and EVALUATIVE meanings. Non-evaluative meanings include the meanings of TEMPORAL CONTINUITY OR DURATION (e.g., *para-méno* ‘to stay on, to remain’, *para-therízo* ‘to spend the summer’), TRANSMITTAL (e.g., *para-dído* ‘to deliver’), RESEMBLANCE (e.g., *para-frázo* ‘to paraphrase’) and PARALLEL, SUBSIDIARY OR ACCESSORY ROLE (e.g., *pará-dromos* ‘side road’, *para-mána* ‘nanny, surrogate mother’). Evaluative meanings comprise the meanings of DIVERGENCE, ERROR OR VIOLATION (e.g., *para-vlépo* ‘to overlook’, *par-erminévo* ‘to misread’, *parátipos* ‘irregular’), EXCESS (e.g., *para-cimáme* ‘to oversleep’, *para-xaidévo* ‘to overindulge’), REINFORCEMENT (e.g., *par-efθís* ‘straight away’) and PERIPHRASTIC REINFORCEMENT (upgrading the determinacy of the propositional content, e.g., *íksere ce para-íksere* ‘he knew and he knew very well’). This semantic classification is illustrated in Figure 1.

From a quantitative point of view, morphological productivity can be defined in a probabilistic sense (Plag 2006). A number of distinct and complementary corpus-based statistical measures of productivity, which all rely on the availability of large electronic text corpora, have been proposed in the last few decades by Baayen and his collaborators (cf. Baayen 1992; 1993; 2001; see also Baayen & Lieber 1991; Baayen & Renouf 1996; Plag et al. 1999). These measures can be operationalized by the concept of vocabulary growth, i.e., how frequently new word types that are formed by a morphological process are encountered when an increasing amount of text

belong to a verb-adjective continuum (e.g., Lascaratou 1988–1989). Similarly, note that the *Reverse Dictionary of Modern Greek* (Anastassiadi-Symeonidi 2002) classifies the types in *-ón-úsa-ón*, *-ís-ísa-én* and *-as*, *-asa*, *-an* among nouns or adjectives. For this reason, we opted for treating them as a separate category.

⁹ This grammatical category includes non-declinable forms in *-ondas/-óndas* with adverbial use, which have been treated in the literature in various ways: as participles (Nakas 2012), gerunds (Holton et al. 1997) or converbs (Moser 2006).

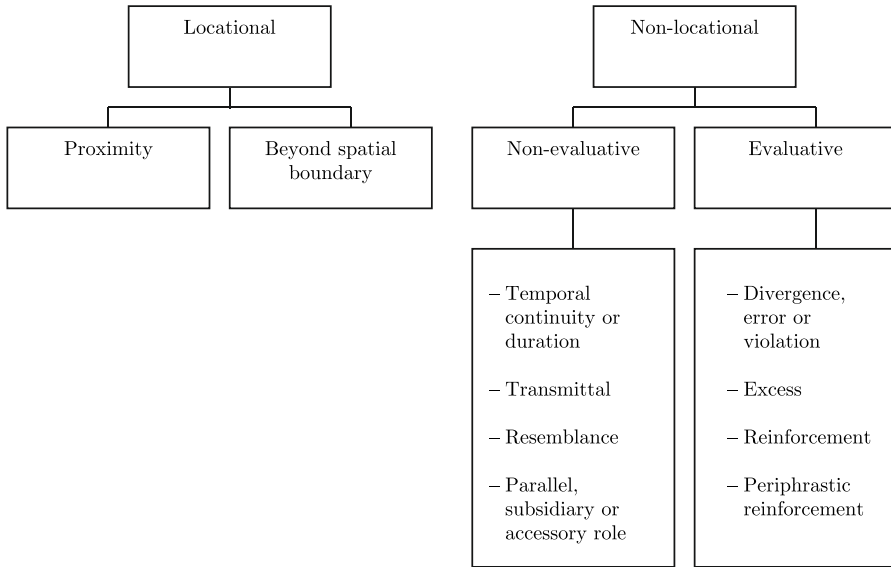


Figure 1: Classification of lemmas derived by *para-* into semantic categories

is sampled. This study relies on this probabilistic aspect of morphological productivity and employs a measure termed as “potential productivity”.

Potential productivity, P , is obtained as the ratio of hapax legomena, V_1 , to the total number of tokens, N , of all words with a morphological category (Baayen 2001). This type of productivity, also called “productivity in the narrow sense” (cf. Plag 1999), serves to estimate the rate at which new types of a given morphological category are to be expected to appear, when N tokens have been sampled (Baayen & Lieber 1991). The resultant figure is a decimal that has a value between 0 and 1. The higher the value, the more productive the category is, and vice versa. The rate at which new types appear in the corpus can be visualized via a vocabulary growth curve. This curve reports vocabulary size (number of types, V) as a function of sample size (number of tokens, N) and P can be also seen as the slope of the tangent to this curve at N .

Potential productivity is employed in this study to estimate the productivity of the prefix *para-* across the subcorpora of written and spoken registers, across the text types included in the corpus and across its grammatical and semantic categories. However, this type of productivity ignores type frequency and, since P is a function of N , its value depends on the

size of the sample of the corpus. It is therefore problematic to compare directly the potential productivity of a small subcorpus with that of a large subcorpus without distortion, due to the substantial differences in V and P (Baayen 1992; Plag et al. 1999). In our case, the differences, both in the corpus as a whole and within registers, text types and categories, cannot be ignored (see Table 1).

An efficient way to compare processes of different size is to extrapolate the values of potential productivity to larger sample sizes. For this purpose, we resort to Large-Number-of-Rare-Events (LNRE) models, a family of parametric statistical models appropriate for modelling word frequency distributions (Baayen 2001). An LNRE model attempts to estimate the expected number of types (the vocabulary size) both at smaller sample sizes (interpolation) and at larger sample sizes (extrapolation), based on the counts of low frequency types in the corpus (the frequency spectrum). The accuracy of this estimate of the number of types depends on how well each model fits the data. Three popular LNRE models considered in this study are the Generalized Inverse Gauss-Poisson (GIGP; Baayen 2001), the finite Zipf-Mandelbrot and Zipf-Mandelbrot (fZM and ZM ; Baroni & Evert 2006). These models are implemented in the package *zipfR* (Baroni & Evert 2006), a tool for lexical statistics in the R language.¹⁰ Although LNRE models provide a practical way to quantify the productivity of a morphological process and compare texts with respect to their lexical richness, caution is required when interpreting and generalizing their results, as they will remain approximate for many actual data sets (Baayen 2001, 236).

To address research questions 1 to 3, the three aforementioned LNRE models were first fitted to the data of each process (semantic category, grammatical category or text type). A multivariate chi-squared test was used to evaluate how well the predictions of each model fit the observed data (Baayen 2001, 118–122). Since satisfactory model fit is often difficult to attain, a pragmatic approach was followed, in the sense that further analyses were based on the best available model. A corresponding growth curve was obtained for the process with the largest number of tokens and the selected model was used to extrapolate the growth curves of the remaining processes up to the size of the largest one. For this range, the expected number of types and hapax legomena of a given process at 100 equally-spaced intervals was estimated. Then, average potential productivity (average P) was calculated. The corresponding growth curves were

¹⁰ For a formal introduction to LNRE models refer to Baayen (2001).

graphed for 100 equally sized intervals and 95% confidence intervals were plotted around the curves. Two curves are regarded as significantly different, if one is outside the confidence interval of the other. A flattening curve indicates an unproductive process, while an ascending curve suggests that a process is productive. In general, it is typical for a vocabulary growth curve to grow more steeply at smaller vocabulary size, where it is more likely to encounter new types. As the corpus grows, the likelihood decreases.

4. Results

This section discusses the productivity patterns of *para-* across its semantic categories (4.1.), spoken and written registers and text types (4.2.) and the grammatical categories of the types prefixed by it (4.3). Furthermore, it presents the findings from the association of both meanings and grammatical categories (4.4.) and meanings and registers and text types (4.5.) in terms of productivity.

4.1. Productivity patterns across semantic categories

Figure 2 shows the vocabulary growth curves of the four semantic super-categories of *para-* (LOCATIONAL and NON-LOCATIONAL, and within the latter EVALUATIVE and NON-EVALUATIVE). The figure illustrates how vocabulary size, i.e., the number of types, shown on the vertical axis, increases as one reads through the tokens of the corpus, plotted on the horizontal axis. The number of types plotted corresponds to the expected vocabulary size $E[V(N)]$, i.e., the number of different types one may expect to count on average for a great many different orderings of the text fragments in a given corpus. The growth curves are plotted for 100 equally sized intervals between 1 and 35,526, which stands for the number of tokens of the NON-LOCATIONAL meaning in the corpus. Figure 2 also plots 95% confidence intervals around the vocabulary growth curves.

Drawing on the information above, it is clear that the most productive meaning super-category of *para-* is NON-LOCATIONAL. This seems to be a result of the high productivity of the semantic category EVALUATIVE, which is a subcategory of the NON-LOCATIONAL meaning. Moreover, NON-EVALUATIVE is clearly less productive than LOCATIONAL meaning. Average potential productivity values of the four categories are presented in the last column of Table 2.

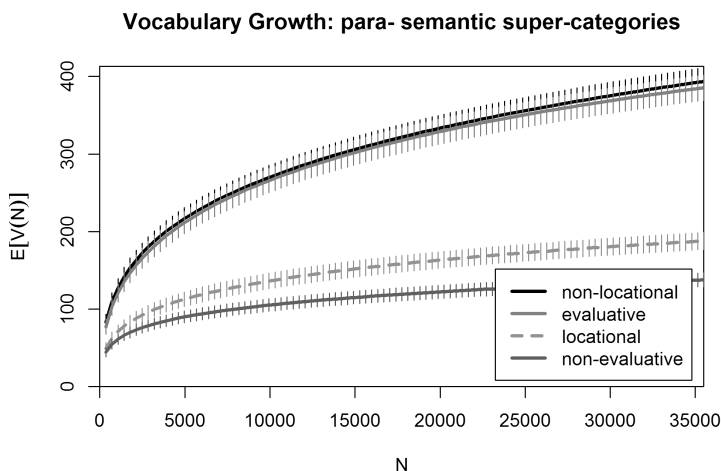


Figure 2: Expected vocabulary size $E[V(N)]$ for four semantic categories (super-categories) of *para-* as a function of the size in tokens N of the NON-LOCATIONAL meaning (35,526 tokens with 95% confidence intervals)

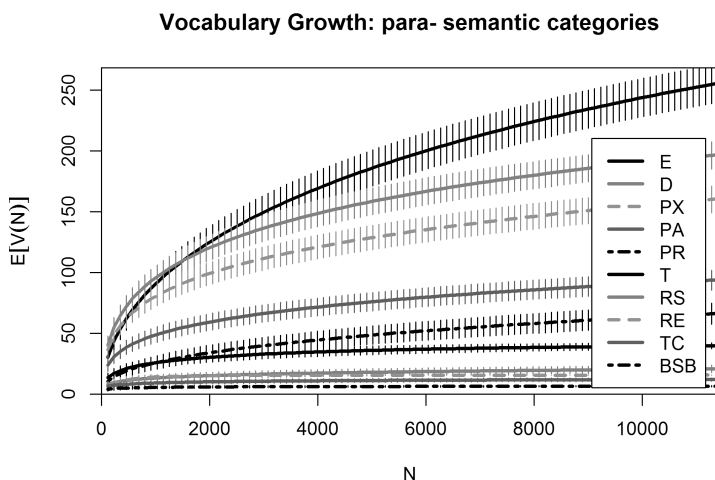


Figure 3: The expected vocabulary size $E[V(N)]$ for ten different meanings of *para-* as a function of the size in tokens N of TRANSMITTAL (first 11,466 tokens with 95% confidence intervals). Note: E = Excess, D = Divergence, PX = Proximity, PA = Parallel, subsidiary or accessory role, PR = Periphrastic reinforcement, T = Transmittal, RS = Resemblance, RE = Reinforcement, TC = Temporal Continuity or duration, BSB = Beyond Spatial Boundary.

Table 2: Types, tokens, hapax-legomena and potential productivity (growth rate) as a function of text type, averaged over 100 equally spaced measurement points in the interval [1, 15,763]

Semantic categories	Types (<i>V</i>)	Tokens (<i>N</i>)	Hapax-legomena (<i>V</i> ₁)	Avg. <i>P</i>
Locational	159	18,070	36	0.0041
Proximity	151	9,128	36	0.0117
Beyond spatial boundary	8	8,942	0	0.0003
Non-locational	394	35,526	110	0.0093
Non-evaluative	129	25,772	27	0.0028
Temporal continuity or duration	12	8,782	1	0.0006
Transmittal	40	11,466	3	0.0025
Resemblance	17	3,455	4	0.0012
Parallel, subsidiary or accessory role	60	2,069	19	0.0065
Evaluative	265	9,754	83	0.0092
Divergence, error or violation	184	8,708	46	0.0146
Excess	57	382	29	0.0205
Reinforcement	13	647	3	0.0009
Periphrastic reinforcement	11	17	5	0.0051

The growth curves of Figure 3 and the average productivity values in Table 2 suggest that the two most productive meanings of *para-* are EXCESS and DIVERGENCE, ERROR OR VIOLATION, which belong to the highly productive EVALUATIVE super-category. However, the other two semantic subcategories of EVALUATIVE are of medium productivity (PERIPHRASTIC REINFORCEMENT) and considerably low productivity (REINFORCEMENT). The third most productive meaning is PROXIMITY, even though LOCATIONAL are less productive than NON-LOCATIONAL meanings. Note that the LOCATIONAL super-category includes PROXIMITY, a very productive meaning and BEYOND SPATIAL BOUNDARY, the least productive meaning, which explains the low overall productivity of the super-category. Furthermore, the low productivity of the super-category NON-EVALUATIVE seems to be mirrored in the majority of its semantic subcategories: three out of four NON-EVALUATIVE meanings (TEMPORAL CONTINUITY OR DURATION, RESEMBLANCE and TRANSMITTAL) belong to the group of the least productive meanings. The only exception is the meaning of PARALLEL, SUBSIDIARY OR ACCESSORY ROLE, which is the fourth most productive meaning, showing a medium score in terms of productivity.

4.2. Productivity patterns across registers and text types

The productivity of *para-* has also been measured across spoken and written registers of the CGT, as well as across specific spoken and written text types. Following a similar methodology with the one presented in the previous section, a growth curve for each subcorpus was drawn up to the number of tokens for the prefix in the written registers. According to Figure 4 and Table 3 below, the prefix *para-* was found more productive in written registers.

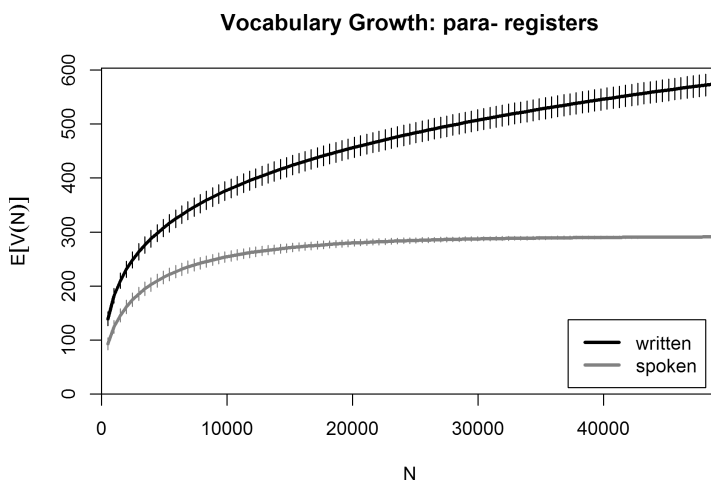


Figure 4: The expected number of types $E[V(N)]$ for *para-* (calculated by means of the finite Zipf-Mandelbrot (fZM) model) in the written and spoken subcorpora of the CGT as a function of the size in tokens N sampled for *para-* in the written subcorpus (first 9,159 tokens with 95% confidence intervals)

Figure 5 shows the growth curves corresponding to specific written and spoken text types. All the slopes follow the general tendency to grow more rapidly at low N and then gradually decrease. *Para-* appears to be more productive in opinion articles and popularized non-fiction texts, followed by academic texts, literature, news, miscellanea, public speech and information items. The prefix is least productive in law and administration texts and interviews. For a small number of text types with very low token frequency, the identification of productivity patterns via LNRE models was not possible (noted by “n/a” in Table 3).

Vocabulary Growth: para- text types

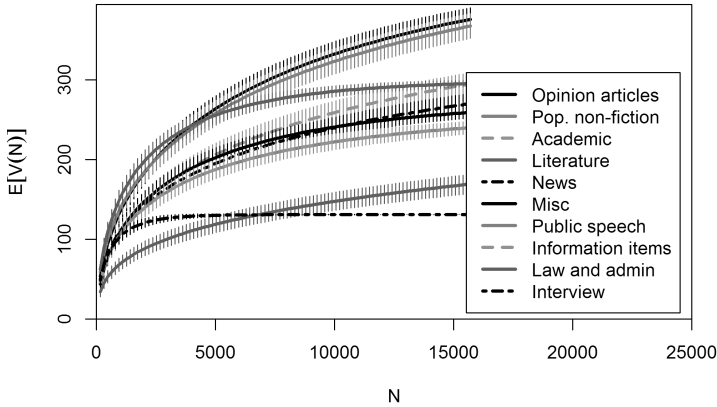


Figure 5: The expected vocabulary size $E[V(N)]$ for *para-* in ten text types as a function of the size in tokens N of pop. non-fiction (first 15,763 tokens with 95% confidence intervals)

4.3. Productivity patterns across grammatical categories

The vocabulary growth curves of the six grammatical categories are shown in Figure 6. Average potential productivity values based on the curves are summarized in Table 4. The prefix was found clearly most productive in nouns, which is shown by the steepest curve.¹¹ On the other hand, it has the lowest productivity in adverbial participles and adverbs, represented by almost flat curves. Average P values indicate that adjectival participles are slightly more productive than verbs and adjectives. However, note that the growth curve of adjectival participles grows steeply at lower sample sizes but soon becomes completely flat, indicating that it is unlikely to encounter new, unattested types for larger samples. On the contrary, verbs and adjectives still show some potential of new types even for the maximum number of tokens.

¹¹ This result is not an effect of classifying adjectival and adverbial participles in separate categories, rather than including them in verbs, as suggested by one anonymous reviewer. Nouns are the most productive category, even if the above categories were to be counted as one.

Table 3: Types, tokens, hapax-legomena and potential productivity (growth rate) as a function of text type, averaged over 100 equally spaced measurement points in the interval [1, 15,763].

Text type	Types (V)	Tokens (N)	Hapax-legomena (V_1)	Avg. P
Conversation	1	1	1	n/a
Interview	92	673	32	0.0061
News (spoken)	68	303	28	n/a
Public speech	168	3,456	55	0.0132
Spoken (total)	210	4,433	60	0.0045
Information items	18	159	7	0.0132
Law and administration	114	4,372	35	0.0091
Literature	230	3,248	64	0.0160
Private	50	267	22	n/a
Pop. non-fiction	368	15,763	101	0.0207
News	231	8,716	67	0.0148
Miscellanea	97	757	44	0.0145
Academic	255	9,488	77	0.0165
Opinion articles	283	6,081	98	0.0215
Procedural	33	312	7	n/a
Written (total)	575	49,163	142	0.0095

4.4. Productivity patterns across meanings and grammatical categories

The association between meanings and grammatical categories in terms of productivity is summarized in Table 5. The values correspond to the types, tokens, hapax legomena and average potential productivity of *para-* in the subcorpus defined by the grammatical category of the corresponding row and the meaning of the corresponding column. Each value of productivity is an average of 100 equally spaced measurement points up to the total number of tokens (53,596). It is important to outline that in most cases the size of the subcorpus did not allow for model estimation (“n/a” in Table 5). Thus, further discussion is based only on available data for each meaning and grammatical category and the results should be interpreted with caution.

Table 5 allows us to identify the most productive meanings across grammatical categories, but also the most productive categories across different meanings. As regards *LOCATIONAL* meanings, *PROXIMITY* was

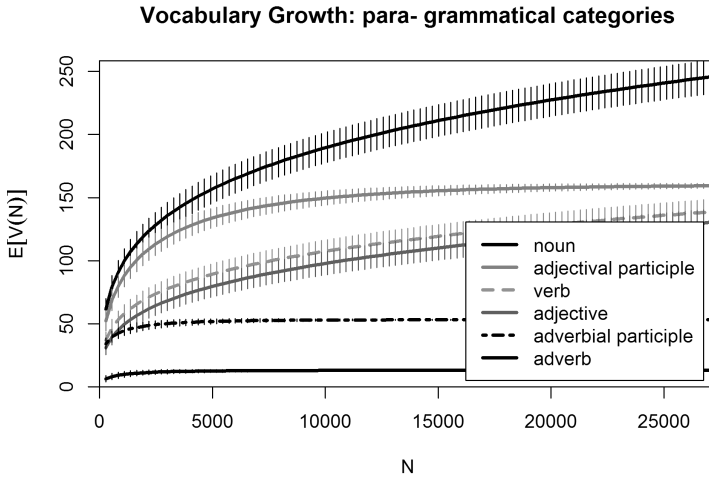


Figure 6: The expected vocabulary size $E[V(N)]$ for the six different grammatical categories of *para-* as a function of the size in tokens N of verbs (first 27,349 tokens with 95% confidence intervals)

Table 4: Types, tokens, hapax-legomena and potential productivity (growth rate) as a function of text type, averaged over 100 equally spaced measurement points in the interval $[1, 15,763]$.

Grammatical category	Types (V)	Tokens (N)	Hapax-legomena (V_1)	Avg. P
Adjective	12	9,037	29	0.0039
Noun	159	5,254	44	0.0072
Verb	139	27,349	39	0.0040
Adj. Participle	107	2,026	29	0.0044
Adv. Participle	43	855	7	0.0009
Adverb	12	9,075	0	0.0003

found to be most productive in adjectives and nouns, whereas BEYOND SPATIAL BOUNDARY in adverbs. Among NON-EVALUATIVE meanings, TEMPORAL CONTINUITY OR DURATION and TRANSMITTAL were found most productive in verbs, while RESEMBLANCE and PARALLEL, SUBSIDIARY OR ACCESSORY ROLE were most productive in nouns. In the case of EVALUATIVE meanings, DIVERGENCE, ERROR OR VIOLATION appears most productive in adjectival participles and nouns and the productivity of EXCESS is high in verbs.

Table 5: Productivity patterns across meanings and grammatical categories. Note: E = Excess, D = Divergence, PX = Proximity, PA = Parallel, subsidiary or accessory, PR = Periphrastic reinforcement, T = Transmittal, RS = Resemblance, RE = Reinforcement, TC = Temporal Continuity, BSB = Beyond Spatial Boundary.

Gram. Category	Meanings of <i>para-</i>											
	BSB				D				E			
	<i>V</i>	<i>N</i>	<i>V</i> ₁	Avg. <i>P</i>	<i>V</i>	<i>N</i>	<i>V</i> ₁	Avg. <i>P</i>	<i>V</i>	<i>N</i>	<i>V</i> ₁	Avg. <i>P</i>
Adj. Part.	0	0	0	n/a	30	357	12	0.0057	11	50	6	n/a
Adj.	0	0	0	n/a	34	2,998	7	0.0029	3	96	1	n/a
Adverb	8	8,945	0	0.0026	0	0	0	n/a	1	3	0	n/a
Adv. Part.	0	0	0	n/a	21	219	4	0.0009	1	1	1	n/a
Noun	0	0	0	n/a	55	1,870	14	0.0055	1	1	1	n/a
Verb	0	0	0	n/a	34	3,264	3	0.0017	40	231	20	0.0082
Gram. Category	PA				PR				PX			
	<i>V</i>	<i>N</i>	<i>V</i> ₁	Avg. <i>P</i>	<i>V</i>	<i>N</i>	<i>V</i> ₁	Avg. <i>P</i>	<i>V</i>	<i>N</i>	<i>V</i> ₁	Avg. <i>P</i>
	Adj. Part.	3	17	1	n/a	0	0	0	n/a	31	482	6
Adj.	14	593	5	0.0021	0	0	0	n/a	39	2,373	15	0.0051
Adverb	0	0	0	n/a	0	0	0	n/a	1	19	0	n/a
Adv. Part.	1	35	0	n/a	0	0	0	n/a	8	120	1	n/a
Noun	38	1,112	12	0.0047	0	0	0	n/a	42	1,336	7	0.0028
Verb	4	312	1	0.0002	11	17	5	n/a	31	4,798	6	0.0021
Gram. Category	RE				RS				TC			
	<i>V</i>	<i>N</i>	<i>V</i> ₁	Avg. <i>P</i>	<i>V</i>	<i>N</i>	<i>V</i> ₁	Avg. <i>P</i>	<i>V</i>	<i>N</i>	<i>V</i> ₁	Avg. <i>P</i>
	Adj. Part.	3	26	1	n/a	3	32	0	n/a	5	413	0
Adj.	1	574	0	n/a	3	2,403	1	0.0004	0	0	0	n/a
Adverb	1	2	0	n/a	1	106	0	n/a	0	0	0	n/a
Adv. Part.	1	2	0	n/a	2	39	0	n/a	2	102	0	n/a
Noun	2	27	0	n/a	4	816	2	0.0012	1	3	1	n/a
Verb	2	16	1	n/a	2	56	0	n/a	4	8,264	1	0.0003
Gram. Category	T											
	<i>V</i>	<i>N</i>	<i>V</i> ₁	Avg. <i>P</i>								
	Adj. Part.	16	649	0	0.0001							
Adj.	0	0	0	n/a								
Adverb	0	0	0	n/a								
Adv. Part.	7	337	1	0.0001								
Noun	7	89	1	n/a								
Verb	10	10,391	1	0.0005								

A reverse interpretation of the productivity patterns observed in Table 5 shows that nouns are highly productive in DIVERGENCE ERROR OR VIOLATION and PARALLEL, SUBSIDIARY OR ACCESSORY ROLE, verbs are most productive in EXCESS and highly productive in meanings especially suited to verbs (TEMPORAL CONTINUITY OR DURATION and TRANSMITTAL). Adjectives are most productive in PROXIMITY and both adjectival and adverbial participles in DIVERGENCE ERROR OR VIOLATION. The least productive grammatical category (adverbs) is related with the least productive meaning (BEYOND SPATIAL BOUNDARY).

4.5. Productivity patterns across meanings, registers and text types

Following the methodology of the previous section, we investigate the productivity patterns of meanings across registers and text types. Table 6 cross-tabulates the spoken and written text types by the different meanings of *para-* in terms of types, tokens, hapax legomena and average potential productivity. As can be seen, the order of the three most productive meanings of *para-* in the CGT (see 4.1 above) is the same: EXCESS is the most productive meaning in both spoken and written registers, followed by DIVERGENCE, ERROR OR VIOLATION and PROXIMITY.

EXCESS is highly productive in literature, whereas DIVERGENCE, ERROR OR VIOLATION is the most productive meaning in several text types, i.e., opinion articles, news, popularized non-fiction and academic texts. It is also notable that the meaning of TRANSMITTAL is highly productive in law and administration texts.

5. Discussion

On the basis of the results presented above we can now proceed to answer the research questions raised in the Introduction. As regards the productivity of the prefix across the meaning categories identified, we found that NON-LOCATIONAL meanings are more productive than LOCATIONAL meanings and that EVALUATIVE meanings are more productive than NON-EVALUATIVE meanings. It is interesting that the two most productive meanings, EXCESS and DIVERGENCE, ERROR OR VIOLATION, are both non-locational and evaluative meanings. This finding concurs with the observation that the preposition *pará* has developed non-locational meanings, while its initial locational meanings have been significantly reduced (Poulopoulou 2002). This can be accounted for in the frame of grammaticalization theory, which has identified the movement from an initial locational meaning to more abstract and evaluative (or subjective) meanings as a general pattern of semantic change (Traugott 1995, 31; cf. Amiot 2012).

In particular, EXCESS is found to be the most productive meaning of *para-*. This result seems to confirm the observation found in previous studies (e.g., Delveroudi & Vassilaki 1999; Ralli 2004; Markopoulou 2014) that *para-* became extremely productive with the recent meaning of excess. Furthermore, this finding seems to accord with Lieber and Baayen's claim (1993, 71–72) that an affix which is not very productive may in fact gather strength in some well-defined subset of formations and re-emerge as highly productive there. The presence of the meaning of PROXIMITY in the third rank of productivity shows that this locational meaning remains strong, even if, in general, LOCATIONAL are less productive than NON-LOCATIONAL meanings. This means that the preponderance of non-locational and evaluative meanings in terms of productivity is not absolute and that the prefix *para-* (in contrast with the preposition *pará*) actively preserves aspects of its ancient locational meaning.

Interestingly, enough, the examination of our data reveals that the locational meaning of PROXIMITY does not arise with every possible base, but usually occurs if the base is interpretable as location or position (e.g., *para-thalásios* 'seaside', *par-áktios* 'costal, inshore', *para-pléo* 'sail by'). It is also worth noticing that in our data the meaning of PROXIMITY appears mostly in [+learned] formations usually translated from French and English (e.g., *para-θέτο* 'to quote', *par-aortikós* 'paraaortal', *para-spondilikós* 'paravertebral') or in words of AG or Hellenistic origin (e.g., *par-ekvéno* 'digress').¹² In our opinion, this finding may indicate that the initial locational meaning of *para-* was rejuvenated in Modern Greek by terms adopted from foreign languages. Furthermore, this finding seems to accord with Luschützky and Rainer's (2011) claim that some cases of polysemy of affixes could be due to non-semantic mechanisms such as borrowing.

The results concerning the productivity of the prefix across written and spoken registers suggest that there is higher productivity in written registers. This may be related to the large number of [+learned] derivatives of *para-*, which are more suitable in formal contexts. The preference in written registers may also be affected by the higher productivity of the prefix in nouns and adjectival participles, which take part in nominalization processes, usually found in written texts. Furthermore, the findings confirm that written language is lexically richer than spoken language, as has already been suggested (Plag et al. 1999; cf. Biber et al. 1999, 53). It is interesting that written registers also show higher productivity in our

¹² It is also interesting to note that the locational *para-* often appears in parasynthetic constructions (e.g., *par-óρθ-ios* 'riparian'; for discussion, see Efthymiou (2015)).

previous studies on the productivity of the majority of Greek verb-forming suffixes (Efthymiou et al. 2012) and on the productivity of the Greek prefix *iper-* (Efthymiou et al. 2014).

Moving on to the productivity of the prefix *para-* across particular text types, the highest scores were found in opinion articles, popularized non-fiction and academic texts. In the case of opinion articles and popularized non-fiction texts, this may be associated with the high productivity of *para-* in evaluative meanings (e.g., *parerminévo* ‘to misread’, *parastratitikós* ‘paramilitary’), since evaluation is expected in subjective and popularized texts. In addition, the high productivity of *para-* in technical and academic terms (e.g., *parafisiología* ‘paraphysiology’, *parénçima* ‘parenchyma’) may account for its high scores in academic and popularized non-fiction texts, in which there is a clear need for establishing expertise in a field.

The measuring of productivity across grammatical categories revealed that *para-* is by far most productive in nouns. This preference may be related with the already mentioned nominalization processes in written registers, where *para-* is very productive. It can also be explained by the high classificatory value of nouns in scientific language, since *para-* is found to be very productive in academic and popularized non-fiction texts.

However, the picture is reversed, when the productivity of grammatical categories across meanings is concerned. Verbs, which are less productive than nouns and adjectival participles, show a one-to-one relation with the most productive meaning of *para-*, EXCESS, as can be seen in Table 7.

Table 7: Correlation of the most productive meanings with the most productive grammatical categories

	Excess	Divergence, error or violation	Proximity
Nouns		+	+
Adj. participles		+	+
Verbs	+	+	+
Adjectives		+	+

This finding supports the recurrent references in the literature to the close relation between verbs derived by *para-* and excessive meaning (Efthymiou 2003, 523; Fista 2007, 151; Gavriilidou 2013, 72; cf. Markopoulou 2014, 185). The second and third most productive meanings (i.e., DIVERGENCE, ERROR OR VIOLATION and PROXIMITY) are found to be productive (more or less) in all four most productive grammatical categories, something which suggests that they accommodate various formations and thus relate to

more established meanings. This difference between the meaning of EXCESS and the other two meanings may reflect the later development of the excessive meaning in contrast with the initial locational meaning of PROXIMITY and the relatively early meaning of DIVERGENCE ERROR OR VIOLATION and brings to mind the discussion in section 2 on the treatment of *para-* as a single polysemous prefix or as two homonymous prefixes, one of which denotes the excessive meaning.

As regards nouns, it is very interesting that they show a wide range of associations with meanings of *para-*, but not with the most productive one, EXCESS. However, they are highly productive in the second most productive meaning, that of DIVERGENCE ERROR OR VIOLATION, as well as in the meaning of PARALLEL, SUBSIDIARY OR ACCESSORY ROLE.

It is also notable that the meaning of PROXIMITY is productive in a large number of grammatical categories, i.e., adjectives, nouns, verbs, adjectival participles. However, the picture is not uniform with the other locational meaning, that of BEYOND SPATIAL BOUNDARY, which, along with being the least productive meaning of *para-*, it is related with the least productive grammatical category (adverbs).

The results on the productivity patterns of meanings across registers have shown that the productivity of the three most productive meanings (EXCESS, DIVERGENCE, ERROR OR VIOLATION and PROXIMITY) is not affected by the spoken or written mode of the texts. As regards particular text types, there are significant correlations of meanings with text types. For instance, the meaning of EXCESS is highly productive in literature. This can be accounted for by the expressive and evaluative role of narration, as well as by the [\pm learned] or [$-$ learned] derived verbs (e.g., *paraléo* ‘to blow out of proportion’, *paraksilóno* ‘to go overboard’), which are useful in imitating everyday style. Whereas EXCESS has a clear profile by basically relating with verbs occurring in literature, DIVERGENCE, ERROR OR VIOLATION draws a more complex picture as the most productive meaning in several text types, namely opinion articles, news, popularized non-fiction and academic texts. This may be reflected on the fact that the meaning is found both in terms like *parafasía* ‘paraphasia’, *paretimolojía* ‘folk etymology’ and in evaluative items such as *paraloyotexnía* ‘second rate literature’, *parasjopó* ‘to gloss over’, *parodiyó* ‘to mislead’. Finally, it is also notable that the meaning of TRANSMITTAL is highly productive in law and administration texts, possibly because of its use in the description of procedures, for example with terms like *parapembómenos* ‘sent to trial’.

6. Conclusions and further research

This paper has proposed a corpus-based semantic categorization for the Modern Greek prefix *para-* and, on the basis of this, measured its productivity across semantic categories, spoken and written registers, several text types and grammatical categories. It also investigated possible associations of these parameters in terms of the prefix's productivity. This is the first attempt to give a comprehensive picture of the productivity of *para-* by taking into account various aspects of the prefix, including its different meanings, its use in context, its preference for particular grammatical categories etc.

The findings on productivity across semantic categories suggest that *para-* is more productive in NON-LOCATIONAL and EVALUATIVE meanings, such as EXCESS and DIVERGENCE, ERROR OR VIOLATION. This preference for more subjective or evaluative meanings than more objective (i.e., locational) meanings can be considered an indication of its affixal status and thus can offer support to the suggestions mentioned in section 2 that *para-* should be regarded as a prefix. Having said this, the picture concerning the semantics of *para-* is more complicated, since the locational meaning of PROXIMITY still remains strong as the third most productive meaning in our data.

The association of the meanings of *para-* with the grammatical categories and the text types in which they occur has revealed an interesting finding. The most productive meaning of *para-*, EXCESS, was found to have a one-to-one relation with the grammatical category of verbs, as well as a clear preference for the text type of literature. On the other hand, other very productive meanings, such as DIVERGENCE, ERROR OR VIOLATION and PROXIMITY, show fewer restrictions by being productive in a wide range of grammatical categories and text types. This particularity of EXCESS may be related to the quite recent development of this meaning, which was not present in Ancient Greek, in contrast to the initial locational meaning of PROXIMITY and the relatively early meaning of DIVERGENCE, ERROR OR VIOLATION, which were possibly rejuvenated in Modern Greek by terms adopted from foreign languages. It is notable that the meaning of EXCESS is not only particular in terms of productivity, but also in terms of its phonology (cf. the observation in section 2 on the optional vowel deletion of the excessive meaning). In our opinion, these findings can be related to the discussion about the treatment of *para-* as a single polysemous prefix or as two homonymous prefixes, one of which denotes the excessive meaning; however, no definitive argument can be provided in

favour of one or the other suggestion, since choosing between them would seem to depend on the theoretical framework adopted.

The findings on the productivity of the meanings of *para-* can pave the way for a more fine-grained semantic analysis, which will not overestimate meanings that have limited or no productivity. It remains to be examined whether a general meaning (e.g., a highly abstract semantic skeleton in the sense of Lieber 2004 or a word formation schema in the sense of Booij 2010) can accommodate for all the (highly or least productive) meanings of the prefix. Another issue that needs further investigation is the type of restrictions (e.g., structural, pragmatic etc.) that take part in word formation with *para-* and their correlation with the productivity of the prefix.

Finally, it would be interesting to analyze the productivity of *para-* in relation to that of other prefixes denoting both locational and intensifying meanings, e.g., *iper-* (*ipér-jios* ‘aboveground’, *iper-ayorá* ‘supermarket’, *iper-plíris* ‘super full’, *iper-forologyó* ‘to overtax’, *iper-esiodoksó* ‘to be over-optimistic’), *kata-* (*kata-céo* ‘to burn all over’, *kat-edafízo* ‘to demolish’), *kse-* (*kse-ðondjázó* ‘to take one’s teeth out’, *kse-kuféno* ‘to make somebody completely deaf’). The comparison to *iper-* seems to be of particular interest, since in their intensifying meaning *para-* and *iper-* are both related to the sense of a limit (Efthymiou 2003, 525). On the basis of our previous work on the productivity and semantics of *iper-* (Efthymiou et al. 2014; 2015) we could suggest that there is a complementary relation between the two prefixes as regards their intensifying meaning. For instance, *para-* and *iper-* seem to be more productive with the meaning of excess in different grammatical categories, namely *para-* in verbs and *iper-* in nouns and adjectival participles. In addition, it is interesting that *para-* is usually found in the meaning of excess with [\pm learned] bases, denoting subjectivity, while *iper-* mostly with [+learned] bases and terminology (cf. Efthymiou 2003, 525). Comparative analyses of this kind between prefixes are found in the literature (Delveroudi & Vassilaki 1999; Karantzola & Giannouloupoulou 2000; Ralli 2002; Efthymiou 2003; Gavriilidou 2013) and can offer useful insight that can be further explored through corpus-based research in large amounts of data.

Acknowledgements

We would like to thank the two anonymous reviewers for their helpful comments and suggestions.

References

- Amiot, Dany. 2004. Haut degré et préfixation. In F. Lefevre and M. Noailly (eds.) *Intensité, Comparaison, Degré. Travaux linguistiques du Cerlico 17*. Rennes: Presses Universitaires de Rennes. 91–104.
- Amiot, Dany. 2005. Between compounding and derivation. Elements of word-formation corresponding to prepositions. In Dressler et al. (2005, 183–195).
- Amiot, Dany. 2012. De la localisation à l'évaluation: des verbes préfixés évaluatifs au sens bien particulier. *Corela* [En ligne], HS-12. <http://corela.revues.org/2775>
- Anastassiadi-Symeonidi, Anna. 1986. Η νεολογία στην Κοινή Νεοελληνική [Neology in Modern Greek]. Thessaloniki: Aristotle University of Thessaloniki.
- Anastassiadi-Symeonidi, Anna. 2002. Αντίστροφο λεξικό της Νέας Ελληνικής [Reverse dictionary of Modern Greek]. Thessaloniki: Institute of Modern Greek Studies.
- Baayen, Harald and Rochelle Lieber. 1991. Productivity and English derivation: A corpus-based study. *Linguistics* 29. 801–843.
- Baayen, Harald and Antoinette Renouf. 1996. Chronicling “The Times”: Productive innovations in an English newspaper. *Language* 72. 69–96.
- Baayen, R. Harald. 1992. Quantitative aspects of morphological productivity. In G. Booij and J. van Marle (eds.) *Yearbook of morphology 1991*. Dordrecht: Kluwer. 109–150.
- Baayen, R. Harald. 1993. On frequency, transparency and productivity. In G. Booij and J. van Marle (eds.) *Yearbook of morphology 1992*. Dordrecht: Kluwer. 181–208.
- Baayen, R. Harald. 2001. *Word frequency distributions*. Dordrecht: Kluwer.
- Babiniotis, George. 2002. *Λεξικό της Νέας Ελληνικής Γλώσσας* [Dictionary of Modern Greek]. Athens: Lexicology Centre.
- Baroni, Marco and Stefan Evert. 2006. The zipfR package for lexical statistics: A tutorial introduction. <http://zipfr.r-forge.r-project.org/materials/zipfr-tutorial.pdf>
- Bauer, Laurie, Lieber Rochelle and Ingo Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Eduard Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Booij, Geert. 2010. *Construction morphology*. Oxford: Oxford University Press.
- Booij, Geert E. 2005. Compounding and derivation: Evidence for Construction Morphology. In Dressler et al. (2005, 109–32).
- Bortone, Pietro. 2010. *Greek prepositions. From antiquity to the present*. Oxford: Oxford University Press.
- Delveroudi, Rea and Sophie Vassilaki. 1999. Préfixes d'intensité en grec moderne: *para-*, *kata-*, *poly-* et *olo-*. In A. Deschamps and J. Guillemin-Flescher (eds.) *Les opérations de détermination: Quantification/qualification*. Paris: Orphys. 149–167.
- Di Sciullo, Anna Maria. 1997. Prefixed-verbs and adjunct identification. In A. M. Di Sciullo (ed.) *Projections and interface conditions*. Oxford: Oxford University Press. 52–73.
- Dressler, Wolfgang U., Dieter Kastovsky, Oskar E. Pfeiffer and Franz Rainer (eds.). 2005. *Morphology and its demarcations*. Amsterdam & Philadelphia: John Benjamins.
- Efthymiou, Angeliki. 2003. Προθήματα ή ά συνθετικά που δηλώνουν επίταση στη Ν.Ε. [Prefixes and first constituents denoting intensity in Modern Greek]. *Studies in Greek Linguistics* 23. 519–528.

- Efthymiou, Angeliki. 2015. Modern Greek parasynthetic verbs. A hierarchical relationship between prefixes and suffixes? In S. Manova (ed.) *Affix ordering across languages and frameworks*. Oxford: Oxford University Press. 82–107.
- Efthymiou, Angeliki, Georgia Fragaki and Angelos Markos. 2012. Productivity of verb forming suffixes in Modern Greek: A corpus-based study. *Morphology* 22. 515–545.
- Efthymiou, Angeliki, Georgia Fragaki and Angelos Markos. 2014. Aspects of productivity of the Modern Greek prefix *iper-*: A corpus-based study. In G. Kotzoglou, K. Nikolou, E. Karantzola, K. Frantzi, I. Galantomos, M. Georgalidou, V. Kourti-Kazoullis, C. Papadopoulou and E. Vlachou (eds.) *Proceedings of the 11th ICGL*. Rhodes: University of the Aegean. 373–383.
- Efthymiou, Angeliki, Georgia Fragaki and Angelos Markos. 2015. Exploring the polysemy of the Modern Greek prefix *iper-*. *Morphology* (Online First). DOI 10.1007/s11525-015-9269-9.
- Fista, Evangelia. 2007. Μέθοδος αυτόματης αναπαράστασης προθηματοποιημένων και συνθέτων ρημάτων της Νέας Ελληνικής [A method for automatic representation of prefixed verbs of Modern Greek]. Doctoral dissertation. Aristotle University of Thessaloniki.
- Fradin, Bernard, Georgette Dal, Natalia Grabar, Fiammetta Namer, Stéphanie Lignon, Delphine Tribout and Pierre Zweigenbaum. 2008. Remarques sur l'usage des corpus en morphologie. *Langage* 171. 34–59.
- Gaeta, Livio and Davide Ricca. 2003. Frequency and productivity in Italian derivation: A comparison between corpus-based and lexicographical data. *Rivista di Linguistica* 15. 63–98.
- Gavriilidou, Zoe. 2013. Όψεις επίτασης στη Νέα Ελληνική [Aspects of intensity in Modern Greek]. Thessaloniki: Kiriakidis.
- Goutsos, Dionysis. 2010. The corpus of Greek texts: A reference corpus for Modern Greek. *Corpora* 5. 29–44.
- Heine, Bernd, Ulrike Claudi and Friederike Hünemeyer. 1991. *Grammaticalization. A conceptual framework*. Chicago: University of Chicago Press.
- Holton, David, Peter Mackridge and Irene Philippaki-Warbuton. 1997. *Greek: A comprehensive grammar of the modern language*. London: Routledge.
- Iacobini, Claudio. 1999. Distinguishing derivational prefixes from initial combining forms. In G. Booij, A. Ralli and S. Scalise (eds.) *Proceedings of the first Mediterranean Morphology Meeting*. Patras: University of Patras. 132–140.
- Karantzola, Eleni and Giannoula Giannoulou. 2000. Σημσιολογικά στοιχεία για τη σύνθεση και την παραγωγή στην πρώιμη νεοελληνική: Α' σύνθεση με προθέσεις [Semantic elements about compounding and derivation in Early Modern Greek: A compounding with prepositions]. *Studies in Greek Linguistics* 20. 193–202.
- Kennedy, Christopher and Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81. 1–37.
- Lascaratou, Chryssoula. 1988–1989. How 'adjectival' are adjectival passive participles in Modern Greek and English? *Glossologia* 7–8. 87–97.
- Liddell, Henry G. and Robert Scott. 1961. *A Greek–English lexicon. A new edition, revised and augmented throughout by Sir Henry Stuart Jones*. Oxford: Clarendon.
- Lieber, Rochelle. 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press.

- Lieber, Rochelle. 2005. English word-formation processes: Observations, issues, and thoughts on future research. In P. Štekauer and R. Lieber (eds.) *Handbook of word-formation*. Dordrecht: Springer. 375–427.
- Malikouti-Drachman, Angeliki and Gaberell Drachman. 1992. Greek clitics and lexical phonology. In W. U. Dressler, H. C. Luschützky, O. E. Pfeiffer and J. R. Rennison (eds.) *Phonologica 1988*. Cambridge: Cambridge University Press. 197–206.
- Markopoulou, Despina. 2014. Γραμματικοποίηση και λεξικοποίηση: Αναψηλαφώντας μια αμφισβητούμενη σχέση [Grammaticalization and lexicalization: Reconsidering a controversial relation]. In Z. Gavriilidou and A. Revithiadou (eds.) [Studies dedicated to Anna Anastassiadis-Symeonidis, Emeritus Professor of Aristotle University of Thessaloniki]. Kavala: Saita. 179–192.
- Moser, Amalia. 1994. The interaction of lexical and grammatical aspect in Modern Greek. In I. Philippaki-Warbuton, K. Nicolaidis and M. Sifianou (eds.) *Themes in Greek linguistics*. Amsterdam & Philadelphia: John Benjamins. 137–144.
- Moser, Amalia. 2006. The Greek forms in *-ontas*: A study in “converbiality”, temporality, aspectuality and finiteness. *Glossologia* 17. 43–67.
- Nakas, Thanasis. 2012. Για το νεοελληνικό «γερούνδιο» ή «γερουντίβο» ή «γερουνδιακό» (άλλως πως, «far from the madding gerund!») [On the Modern Greek “gerund” or “gerundium” or “gerundivum” (in other words, “far from the madding gerund!”)]. In Z. Gavriilidou, A. Efthymiou, E. Thomadaki, and P. Kambakis-Vougiouklis (eds.) *Selected papers of the 10th ICGL*. Komotini: Democritus University of Thrace. 990–999.
- Paradis, Carita. 2008. Configurations, construals and change: expressions of DEGREE. *English Language and Linguistics* 12. 317–43.
- Philippaki-Warbuton, Irene. 1970. *On the verb in Modern Greek*. Bloomington, IN: Indiana University Publications.
- Plag, Ingo. 1999. *Morphological productivity. Structural constraints in English derivation*. Berlin & New York: Mouton de Gruyter.
- Plag, Ingo. 2006. *Productivity*. In B. Aarts and A. McMahon (eds.) *Handbook of English linguistics*. Malden, MA & Oxford: Blackwell. 537–556.
- Plag, Ingo, Christiane Dalton Puffer and R. Harald Baayen. 1999. *Morphological productivity across speech and writing*. *English Language and Linguistics* 3. 209–228.
- Pouloupoulou, Maria. 2002. Από τις τοπικές στις μη τοπικές σχέσεις: Η περίπτωση των *παρά, αντί, εκτός/έξω* [From locational to non-locational relations: The case of *pará, andí, ektós/ékso*]. Doctoral dissertation. Aristotle University of Thessaloniki.
- Rainer, Franz. 2005. Semantic change in word formation. *Linguistics* 43. 415–441.
- Rainer, Franz. 2009. Polysemy in derivation. In R. Lieber and P. Štekauer (eds.) *The Oxford handbook of compounding*. Oxford: Oxford University Press. 338–353.
- Ralli, Angeliki. 2002. Prefixation vs. compounding. The case of Greek preverbs. In A. M. D. Sciullo (ed.) *Asymmetry in grammar: Phonology, morphology and language acquisition*. Amsterdam & Philadelphia: John Benjamins. 37–65.
- Ralli, Angeliki. 2004. Stem-based versus word-based morphological configurations: The case of Modern Greek preverbs. *Lingue e Linguaggio* 2. 241–275.
- Ralli, Angeliki. 2007. Η σύνθεση λέξεων. Διαγλωσσική μορφολογική προσέγγιση [Compounding: A cross-linguistic morphological approach]. Athens: Patakis.
- Ralli, Angeliki. 2013. *Compounding in Modern Greek*. Dordrecht: Springer.

- Schwyzler, Eduard. 1950. Griechische Grammatik 2. Syntax und Syntaktische Stilistik. München: Beck.
- Scott, Mike. 2008. WordSmith Tools, version 5. Liverpool: Lexical Analysis Software.
- Smirniotopoulos, Jane and Brian Joseph. 1998. Syntax versus the lexicon: Incorporation and compounding in Modern Greek. *Journal of Linguistics* 34. 447–488.
- Talmy, Leonard. 2000. *Toward a cognitive semantics*. Vol. 1: Concept structuring systems; Vol. 2: Typology and process in concept structuring. Cambridge, MA: MIT Press.
- Tognini-Bonelli, Elena. 2001. *Corpus linguistics at work*. Amsterdam & Philadelphia: John Benjamins.
- Traugott, Elizabeth Closs. 1995. Subjectification in grammaticalisation. In D. Stein and S. Wright (eds.) *Subjectivity and subjectification*. Cambridge: Cambridge University Press. 31–54.
- Traugott, Elizabeth Closs and Richard B. Dasher. 2002. *Regularity in semantic change* (2nd ed.). Cambridge: Cambridge University Press.
- Triandafyllidis, Manolis. 1991 [1941]. *Νεοελληνική γραμματική (της Δημοτικής)* [Modern Greek Grammar of the Demotic]. Thessaloniki: Institute of Modern Greek Studies, Manolis Triandafyllidis Foundation.
- Triandafyllidis, Manolis. 1998. *Λεξικό της Κοινής Νεοελληνικής* [Standard Modern Greek dictionary]. Thessaloniki: Institute of Modern Greek Studies, Aristotle University of Thessaloniki.
- Tyler, Andrea and Vyvyan Evans. 2001. Reconsidering prepositional polysemy networks: The case of *over*. *Language* 77. 724–765.