

LOW PRODUCTIVITY INDEXES: THE OLD ENGLISH VERBAL SUFFIXES *-ETTAN* AND *-LÆCAN*

GEMA MAÍZ VILLALTA*

ABSTRACT. *The aim of this journal article is to measure the productivity of two Old English verbal suffixes that display a low index of productivity, namely -ettan and -læcan. The main sources of this research are the lexical database of Old English Nerthus and the Dictionary of Old English Corpus. Methodologically, two indexes of productivity are measured, productivity in the narrow sense P and global productivity P* (Baayen 1989, 1992, 1993), as well as frequency, which is calculated in terms of type and token (Bauer 2005). Two conclusions are reached. Firstly, the indexes of productivity P and global productivity P* of -ettan are higher than the ones of -læcan. After a comparison with other Old English affixes, it turns out that -ettan is relatively productive while the suffix -læcan is practically unproductive. Secondly, the indexes of productivity and the frequencies of type and token of these affixes coincide and are in accordance with Bauer's (2004) generalization that a low type frequency matched by a high token frequency implies low productivity.*

KEY WORDS. *Old English, word-formation, frequency, productivity, lexical database of Old English Nerthus, corpus.*

RESUMEN. *El objetivo de este artículo es el análisis de la baja productividad de dos de los sufijos verbales del inglés antiguo -ettan y -læcan. Las principales fuentes utilizadas en este estudio son la base de datos léxica del inglés antiguo Nerthus y el diccionario online del Dictionary of Old English Corpus. En cuanto a la metodología, se tiene en cuenta las medidas de productividad P en un sentido limitado y la productividad P* en un sentido más amplio (Baayen 1989, 1992, 1993), al igual que se ha tenido en cuenta el índice de frecuencia de clase y la frecuencia de tipo respectivamente (Bauer 2005). Se llega a la conclusión de que los índices de productividad P y productividad global P* son más altos en el sufijo -ettan que en -læcan, por lo que el sufijo -ettan es relativamente productivo, mientras que el sufijo -læcan apenas es productivo. En segundo lugar, los índices de productividad y de frecuencia de clase y de tipo de los afijos en cuestión están en consonancia con la afirmación de Bauer (2004) de que una frecuencia de clase baja coincide con una frecuencia de tipo alta, lo que a su vez implica una productividad baja.*

PALABRAS CLAVE. *Inglés antiguo, formación de palabras, frecuencia, productividad, base de datos léxica Nerthus, corpus.*

1. INTRODUCTION

Fernández-Domínguez *et al.* (2007: 51) comment on the tendency in word-formation research to focus on high productivity measures to the exclusion of low indexes of productivity. These authors (2007: 35) distinguish three major models of productivity assessment: frequency models, which centre on type frequency, token frequency and relative frequency; probabilistic models, which concentrate on productivity on the narrow sense, global productivity and the degree of productivity based on the count of *hapax legomena* or unique formations; and Štekauer's (2005) onomasiological model, in which word-formation patterns are regular, predictable and productive. After applying these models to a corpus of nearly six thousand items, these authors reach the conclusion that the interpretation of figures for low productivity is ambiguous. As Fernández-Domínguez *et al.* (2007: 51) put it, "whereas high figures unequivocally correspond to high productivity, it is not entirely clear whether low figures correspondingly match low productivity or whether they imply a decrease in measurement accuracy". These authors identify issues in assessing the productivity of processes of word-formation. In the first place, frequency models are restricted to affixation measurement, which puts aside other morphological processes of lexical creation. Secondly, probabilistic models rely almost exclusively on figures to interpret the data. Finally, the onomasiological model is hardly compatible with the others because it considers non-quantitative aspects of word-formation processes, such as the need of the speech community for a given neologism.

Although this journal article cannot solve all the problems raised by the authors cited above, it aims at contributing to the current debate on morphological productivity in general and the assessment of low productivity in particular by taking a look at a historical language. More specifically, the aim of this work is to measure the productivity of the Old English verbal suffixes *-ettan* and *-læcan* by means of a methodology based on the following premises (Maíz Villata, *fc.*; Mateo Mendaza *fc.-a*, *fc.-b*): (i) the assessment of productivity requires a distinction between the type frequency (dictionary-based) and the token frequency (corpus-based) of the word-formation processes under scrutiny; (ii) an accurate measurement of productivity calls for a combination of frequency and probabilistic models of productivity; (iii) the quantitative data obtained from the measurement of productivity must be interpreted in terms of explanatory principles such as lexicalization and grammaticalization. As regards (i), in the framework of a historical language, the morphological productivity of a word-formation process has to make reference to the number of attested types and tokens produced by the process in question. With reference to (ii), Bauer (2005) has distinguished two different approaches to the study of productivity: a qualitative approach that makes reference to the property of a given word-formation process or affix to be used to derive new words in a systematic way, and a quantitative approach, whereby the productivity of an affix can be measured by counting the number of attested types with that affix. Thus, the methodology adopted in this piece of research adopts a qualitative and quantitative approach to the question of productivity. Regarding (iii), some theories of productivity highlight the predictability of

productivity assessments. For instance, for Plag (1999: 6) productivity is “the possibility of coining new complex words according to the word formation rules of a given language, as these rules may predict the existence of forms which are unattested or whose status as well-formed derivatives is more than doubtful”. A study in the productivity of a historical language can yield conclusions on the reconstruction of unattested forms, but cannot take issue with the possibility of coining new words. On the other hand, a historical language makes allowance for judgements of loss of analysability (lexicalization) and loss of lexical status (grammaticalization) that can be made from a diachronic perspective only. When a lexical item has disappeared from the lexicon throughout diachronic evolution, as is the case with the two affixes under scrutiny in this article, the historical perspective draws a picture of decaying productivity, but it is still necessary to measure the productivity of the process at a given synchronic stage.

Given this background, this article is organized as follows. Section 2 raises the methodological questions involved in the assessment of morphological productivity, with special reference to a historical language, and offers a brief morphological description of the two Old English suffixes under analysis. Section 3 focuses on the type and token frequency as well as the indexes of productivity of the suffixes and offers a comparison with other Old English suffixes with low productivity index. Finally, section 4 draws the main conclusions of this work. The data of this research have been retrieved from the lexical database of Old English *Nerthus* (www.nerthusproject.com) and the *Dictionary of Old English Corpus* (Healey *et al.* 2004).

2. SOME METHODOLOGICAL REMARKS

Whereas productivity has been central to the debate in morphological theory (Aronoff 1976; Baayen 1989, 1992, 1993; Baayen and Lieber 1991; Baayen and Renouf 1996; and Bauer 2001, 2005, among others), few steps have been taken in the area of Old English productivity. Kastovsky (1992) and Lass (1994) stress the difficulty of assessing the productivity of the word-formation processes of a historical language as a result of the coexistence of productive and unproductive patterns in the lexicon. Kastovsky (1992: 356-358) finds three main problems when dealing with Old English productivity. First, there is no direct way of testing productivity, which implies that we have to rely on indirect evidence such as the number of occurrences in a text or the continuity of a given process of word-formation. Second, productivity and transparency can vary diachronically. In Kastovsky’s (1992: 357) words, “when one has to deal with a linguistic period such as Old English, stretching over some 600 years, there are bound to have been many changes. Only the output of the patterns recorded in the later documents is available for study”. This goes hand in hand with Lass (1994: 193), who remarks that it is difficult to determine whether a given occurrence of a derived form represents an institutionalized lexical item or not, or whether it is a new formation. And third, when a given word-formation process loses its productivity, it may leave at least some of its output as part of the vocabulary.

In short, the question of the productivity of affixes and processes of Old English word-formation remains unexplored, in spite of the remarkable generalization and transparency of this phenomenon. In this respect, Kastovsky (1992), who offers the most systematic and comprehensive account of Old English word-formation to the date, insists on the associative character of the lexicon, which is comprised of Germanic lexical items, as well as on the derivational families that result from the operation of word-formation processes. Kastovsky (1986, 1989, 1990, 2005, 2006) also identifies a typological shift in Old English from stem-based morphology to word-based morphology and reaches the conclusion that by the end of the Old English period morphophonological alternations have lost their motivation completely. From a more theoretical perspective, Martín Arista (2008, 2009, 2010a, 2010b, 2010c, 2011a, 2011b, 2011c, fc.-a, fc.-b, fc.-c) has devised a model of functional morphology called *Layered Structure of the Word* and applied it to a number of Old English word-formation phenomena, including recursivity, feature percolation and cumulation, paradigmatic organization, etc.

Among the morphological processes of Old English word-formation that result in the associative and homogeneous character of the lexicon referred to by Kastovsky (1992), we find the attachment of the suffixes *-ettan* and *-læcan*. They produce denominal, deadjectival and deverbal verbs (Jember *et al.* 1975; Kastovsky 1992; Quirk and Wrenn 1994) as can be seen in examples (1a) and (1b), respectively. These examples offer a strictly synchronic description of the morphological relation holding between the base and the derivative of each word-formation process (by drawing on the lexical database of Old English *Nerthus*, I use numbered entries to mark morphologically or lexically relevant contrasts such as *sār 1* ‘sickness, wound, suffering’ vs. *sār 2* ‘sore, sad, grievous’):

(1)

- a. *beðettan* ‘to bath, foment’ (< *bæð* ‘bath’)
sārettan ‘to grieve, lament’ (< *sār 2* ‘sore, sad, grievous’)
bliccettan ‘to glitter, quiver’ (< *blīcan* ‘to glitter, shine’)
- b. *æfenlæcan* ‘to grow towards evening’ (< *æfen* ‘even, evening’)
(ge)cneordlæcan ‘to be diligent, study’ (< *gecneord* ‘eager, diligent’)
gewerodlæcan ‘to make sweet or pleasant’ (< *werodian* ‘to grow sweet’)

For Kastovsky (1992: 391), “-ett(an) seems to have frequentative or intensifying meaning, and is added to nominal and deadjectival, but primarily to verbal bases”. Regarding *-læcan*, Kastovsky (1992: 391) remarks that “læc(an) forms deadjectival verbs with the meaning ‘be, become, make’ and denominal verbs with the meaning ‘produce, grow, become’”. Additionally, *-ettan* and *-læcan* produce deadverbial derivatives, such as *(ge)efnettan* ‘to equal, emulate’ (< *efne 1* ‘even, evenly’), or *gesamodlæcan* ‘to bring together’ (< *samod 1* ‘simultaneously’). The assessment of productivity, therefore, will make reference to the verbal output of these processes, irrespective of the lexical categories of the bases of derivation.

To measure the productivity of these affixes, I combine the textual data retrieved from *The Dictionary of Old English Corpus*, an *online* text collection comprising ca. 3,060 different texts belonging to different categories, such as prose, poetry, glosses to Latin and inscriptions, with a total of approximately 3 million words; and the lexicological information provided by the lexical database of Old English *Nerthus*, which contains a total of ca. 30,000 entries, taken primarily from Clark Hall's *A Concise Anglo-Saxon Dictionary* (1996), and secondarily from Bosworth and Toller's *An Anglo-Saxon Dictionary* (1973) and Sweet's *The Student's Dictionary of Anglo-Saxon* (1976). Lexicographical sources provide lemmatized forms (types), while textual sources display unlemmatized forms (tokens).

Turning to the formulae for measuring productivity, Baayen and Lieber (1991) define the global productivity of a word formation process in terms of the number of different types V and the probability of encountering new types (S). This can be seen in figure 1:

$$I = V / S$$

Figure 1. Index of productivity (Baayen and Lieber 1991)

To refine this approach, Baayen (1989, 1992, 1993) has advanced the notions of productivity (in the narrow sense) P and global productivity P^* . For Baayen, productivity in the narrow sense is defined as the quotient of the number of hapax legomena n_1 with a given affix and the total number of tokens N of all words with that affix, as is shown by figure 2:

$$P = n_1 / N$$

Figure 2. Productivity in the narrow sense (Baayen 1989, 1992, 1993)

As regards hapax legomena and their role in measuring productivity, there is a certain degree of consensus among theoreticians on the importance of unique formations. Although I concur with Lass (1994) on the fact that it is not clear if the existence of a hapax legomenon represents a piece of solid linguistic evidence or simply a question of language survival, I have used the formula displayed by figure 3, which is based on Baayen (1991, 1993), to calculate affix productivity:

$$\text{Index of productivity} = \frac{\text{Number of hapax legomena of the suffix}}{\text{Number of suffixed derivative tokens in the corpus}}$$

Figure 3. Index of productivity (based on Baayen 1991, 1993)

Considering the divisor of the formula in figure 3, the other concept required for the assessment of productivity in this approach is frequency. Frequency can be broken

down into paradigmatic frequency and syntagmatic frequency. Paradigmatic frequency makes reference to the attested derivations on a single base by means of different affixes, whereas syntagmatic frequency refers to the different textual occurrences of a certain attested derivation of a given base. In other words, paradigmatic frequency is based on lemmatized forms as presented by dictionaries whereas syntagmatic frequency measures unlemmatized forms as found in texts. At this particular point, I follow Bauer's (2001, 2005) distinction between type and token frequency. For Bauer (2001), the concept of type frequency refers to the number of items in a dictionary, whereas token frequency represents the number of occurrences of a particular affix in a corpus or a given text. In effect, for Bauer (2004: 102), the "token frequency of a particular affix is calculated from the number of times that an affix appears in a text", while the "type frequency of an affix in a given text is calculated from the number of different lexemes in which the affix occurs". Following Bauer (2001, 2005), I apply the formulae given in figures 4 and 5 to calculate, respectively, type frequency and token frequency:

$$\text{Type frequency} = \frac{\text{Number of suffixed derivatives}}{\text{Number of headwords}}$$

Figure 4. Type frequency (based on Bauer 2001, 2005)

$$\text{Token frequency} = \frac{\text{Number of suffixed derivative tokens}}{\text{Number of words in corpus}}$$

Figure 5. Token frequency (based on Bauer 2001, 2005)

Summarizing, the sort of account of productivity with lexicographical and textual sources adopted in this research requires to take the following steps: (i) to calculate the type frequency of derivatives in *Nerthus*; (ii) to calculate the token frequency and productivity in *The Dictionary of Old English Corpus*; (iii) to count the number of words in *The Dictionary of Old English Corpus*, that is, by prose, poetry and gloss; and (iv) to assess the global productivity of *-ettan* and *-laecan*, both in general and by text type. In the remainder of this section, I offer the details of the analysis.

3. THE PRODUCTIVITY OF THE SUFFIXES *-ETTAN* AND *-LAECAN*

Beginning with the type frequency of the affixes *-ettan* and *-laecan*, 74 types of *-ettan* are provided by *Nerthus*, including a few additions from Marckwardt (1942), while this lexical database turns out a total of 32 types of *-laecan*. They are offered, respectively, by (2a) and (2b), with the number of tokens of each verb between brackets:

(2)

- a. *āgnettan* ‘to appropriate’ (3), *bealcettan* ‘to belch, utter’ (32), *bedettan* ‘to bathe’ (1), *blægettan* ‘to cry’ (0), *bliccettan* ‘to glitter’ (10), *boretan* ‘to brandish’ (2), *bōtettan* ‘to improve’ (2), *brogdettan* ‘to shake’ (11), *cancettan* ‘to cry out’ (2), *ceahhettan* ‘to laugh loudly’ (3), *clæppettan* ‘to palpitate’ (5), *cloccettan* ‘to palpitate’ (1), *crācettan* ‘to croak’ (2), *doppettan* ‘to plunge in’ (3), *droppetan* ‘to drop’ (5), *(ge)efnettān* ‘to equal’ (18), *fāgettān* ‘to change colour’ (3), *feallettan* ‘to fall down’ (3), *firnettān* ‘to ask’ (1), *flogettān* ‘to fluctuate’ (2), *fnārettān* ‘to breathe heavily’ (1), *gealchattān* ‘to ordain, devise’ (1), *gealpettan* ‘to boast’, *gorrettan* ‘to gaze’ (5), *grunnettān* ‘to grunt’ (1), *grymettan* ‘to grunt, roar’ (100), *gyrrettan* ‘to roar (of lions)’ (2), *hafettan* ‘to clap’ (13), *(ge)hālettān* ‘to greet’ (12), *(ge)hāmettān* ‘to domicile’ (4), *hlēapettan* ‘to leap up’ (1), *hloccettan* ‘to utter a sound’ (0), *hoppettan* ‘to hop, leap’ (4), *hospettan* ‘to ridicule’ (1), *hrācettan* ‘to eructate’ (1), *huncettan* ‘to limp, halt’ (1), *lādettan* ‘to hate, loathe’ (11), *lēasettan* ‘to pretend’ (2), *lēaslicettan* ‘to dissemble’ (1), *licettan* ‘to feign’ (42), *lyfettan* ‘to flatter’ (5), *manettān* ‘to admonish’ (0), *miscrōcettan* ‘to croak or shrink horribly’ (1), *mūðettan* ‘to blab out’ (1), *onōrettan* ‘to perform with effort’ (3), *orrettan* ‘to put confusion’ (8), *plicettan* ‘to play with, expose to danger’ (2), *rāscettan* ‘to crackle’ (7), *rāsettān* ‘to rage (of fire)’ (2), *roccettan* ‘to eructate’ (11), *sallettan* ‘to sing psalms’ (1), *sārettan* ‘to grieve’ (2), *scofettan* ‘to drive hither and thither’ (1), *scottettan* ‘to move about quickly’ (6), *scrallettan* ‘to sound loudly’ (2), *siccettan* ‘to sigh’ (4), *slecgettān* ‘to beat’ (1), *spigettān* ‘to spit’ (1), *sporettan* ‘to kick’ (1), *spornettan* ‘to spurn’ (2), *sprangettān* ‘to quiver’ (2), *stammētān* ‘to stammer’ (0), *stefnettān* ‘to stand firm’ (2), *swolgettān* ‘to gargle’ (1), *swōrettan* ‘to breathe hard’ (7), *togettān* ‘to twitch, be spasmodic’ (1), *tolcettān* ‘to be luxuriant’ (2), *towettān* ‘to associate with’ (1), *trepettan* ‘to trip, dance’ (0), *þafettān* ‘to consent’ (1), *þamettān* ‘to clap (the hands)’ (1), *þodettān* ‘to strike’ (2), *wincettān* ‘to wink’ (1), *wīnrēafettān* ‘to gather grapes’ (1).
- b. *æfenlæcan* ‘to grow towards evening’ (5), *fālæcan* ‘to be at enmity with’ (1), *fremedlæcan* ‘to alienate’ (1), *(ge)drīstlæcan* ‘to presume’ (59), *(ge)ðwærlæcan* ‘to agree’ (33), *(ge)cneordlæcan* ‘to be diligent’ (15), *(ge)cūdlæcan* ‘to make known’ (3), *(ge)cȳdlæcan* ‘to become known’ (4), *(ge)cyrtenlæcan* ‘to beautify’ (3), *(ge)dyrstlæcan* ‘to presume’ (132), *(ge)efenlæcan* ‘to be like’ (118), *(ge)lōmlæcan* ‘to frequent’ (21), *(ge)rihtlæcan* ‘to make straight’ (101), *(ge)swæslæcan* ‘to wheedle’ (3), *gedrēoglæcan* ‘to put in order’ (3), *gehīwlæcan* ‘to form’ (0), *gemetlæcan* ‘to moderate’ (2), *geonglæcan* ‘to pass one’s youth’ (4), *gesamodlæcan* ‘to bring together’ (1), *geswētlæcan* ‘to batten’ (6), *gesyndlæcan* ‘to cause to prosper’ (3), *gewærlecan* ‘to warn’ (1), *gewerodlæcan* ‘to make sweet or pleasant’ (1), *gewistlæcan* ‘to feast’ (1), *gewundorlæcan* ‘to make wonderful’ (1), *limplæcan* ‘to unite’ (2), *loflæcan* ‘to praise’ (2), *nālēacan* ‘to come or draw near’ (564), *sumorlæcan* ‘to draw on towards summer’ (1), *swæðlæcan* ‘to search for’ (1), *wīðerlæcan* ‘to deprive’ (1), *winterlæcan* ‘to grow wintry’ (8).

If we compare the type frequency of *-ettan* and *-læcan* verbs, we get the result given in table 1. Notice that the figure of headwords in the lexical database *Nerthus* differs slightly due to the addition, in the case of *-ettan*, of the verbs taken from Marckwardt (1942):

Table 1. Type frequency based on the lexical database *Nerthus*.

Affix	types	<i>Nerthus</i> headwords	type frequency
<i>-ettan</i>	74	30,203	0.00245
<i>-læcan</i>	32	30,180	0.00106

Table 1 shows that the type frequency of *-ettan* verbs is higher than the one of *-læcan* verbs. In other words, the higher the number of types, the higher the frequency of type, given the same number of headwords in the lexical database.

For calculating the token frequency and productivity of *-ettan* and *-læcan* in *The Dictionary of Old English Corpus*, I have searched the corpus for all the unlemmatized (inflective) forms of these weak verbs. I have found a total of 396 occurrences of the 74 *-ettan* verbs and 1,096 occurrences of the 32 *-læcan* verbs, as has been shown in (2) above. Along with the number of tokens, it is necessary to count the hapax legomena to measure the index of productivity as described by figure 3. The token frequency of these affixes in the *The Dictionary of Old English Corpus* is as follows in table 2. The exact figure of words in *The Dictionary of Old English Corpus* used for this calculation is 2,952,566 words.

Table 2. Token frequency based on *The Dictionary of Old English Corpus*

Affix	tokens	DOEC words	token frequency
<i>-ettan</i>	396	2,952,566	0.000134
<i>-læcan</i>	1,096	2,952,566	0.000371

As can be seen in table 2, *-læcan* displays a higher token frequency. As is the case with type frequency, the higher the number of tokens the higher the token frequency of the affix, provided that the number of words in the corpus remains constant. It is remarkable, however, that *-ettan* is more frequent in terms of type whereas *-læcan* is more frequent if the measure of frequency is based on token. In this line, Bauer (2004: 102) remarks that “lack of productivity of an affix is said to lead to a high token frequency of that affix, but a low type frequency”. This means that *-ettan* is more productive than *-læcan*, even though the indexes of productivity thrown by both affixes are considerably low.

Corpus searches have turned out a total of 26 unique formations or hapax legomena with *-ettan* and 10 hapaxes displaying *-læcan*. They are listed under (3.a) and (3.b) respectively. It is important to bear in mind that the concept of *hapax legomenon* is used here in the sense of a single textual occurrence of a base of

derivation to which the affix in question is attached. The infinitival form of reference appears between brackets:

(3)

- a. *beðete* (*beðettan* ‘to bathe’), *cloccet* (*cloccettan* ‘to palpitate’), *firmetton* (*firmettan* ‘to ask’), *fnærettende* (*fnærettan* ‘to breathe heavily’), *gealchatte* (*gealchattan* ‘to ordain’), *galpettað* (*gealpettan* ‘to boast’), *grunnettan* (*grunnettan* ‘to grunt’), *hliapettan* (*hlīapettan* ‘to leap up’), *hospetet* (*hospettan* ‘to ridicule’), *hræctað* (*hræcetan* ‘to eructate’), *huncetton* (*huncettan* ‘to limp’), *leasliccettan* (*lēasliccettan* ‘to dissemble’), *miscrocctetan* (*miscrōcctetan* ‘to croack horribly’), *mudetton* (*mūðettan* ‘to blab out’), *sallettað* (*sallettan* ‘to sing psalms’), *scofet* (*scofettan* ‘to drive hither and thither’), *slecgete* (*slecgettan* ‘to beat’), *spigette* (*spigettan* ‘to spit’), *sporetteð* (*sporettan* ‘to kick’), *swolgettan* (*swolgettan* ‘to gargle’), *togetteð* (*togettan* ‘to twitch’), *towettan* (*towettan* ‘to associate with’), *gedæfetað* (*ðafettan* ‘to consent’), *þametað* (*þamettan* ‘to clap’), *winhreaftetið* (*wīnrēafetian* ‘to gather grapes’), *wincettað* (*wincettan* ‘to wink’).
- b. *fælæce* (*fālæcan* ‘to be an enmity with’), *fremedlæcende* (*fremedlæcan* ‘to alienate’), *gesamodlæceð* (*gesamodlæcan* ‘to bring together’), *gewærlæht* (*gewærlæcan* ‘to warn’), *gewerdlæhþ* (*gewerodlæcan* ‘to make sweet’), *gewistlæcan* (*gewistlæcan* ‘to feast’), *Gewundorlæc* (*gewundorlæcan* ‘to make wonderful’), *sumorlæhð* (*sumorlæcan* ‘to draw towards summer’), *swæðlæhte* (*swæðlæcan* ‘to search for’), *wyþerlecað* (*wiðerlæcan* ‘to deprive’).

Given the token frequency and the figure of hapax legomena given above, the index of productivity P of the suffixes *-ettan* and *-læcan* is shown in table 3:

Table 3. Index of productivity P based on *The Dictionary of Old English Corpus*

Affix	N	n ₁	P
<i>-ettan</i>	396	26	0.065656
<i>-læcan</i>	1,096	10	0.009124

Table 3 is telling us that *-ettan* is more productive than *-læcan*, if the index of productivity is calculated by dividing the number of hapax legomena by the total number of tokens of all words with the affix. This assessment coincides with the one of frequency. This research has shown that the type frequency of *-ettan* is higher than its token frequency, whereas the opposite holds for *-læcan* verbs. These facts indicate that *-ettan* is more productive than *-læcan*, although for a more accurate assessment a comparison with other affixes displaying low indexes of productivity is needed. Mateo Mendaza (fc.-a, fc.-b) has measured the productivity of the Old English adjectival suffixes *-isc*, *-cund* and *-ful(l)* as well as the prefix *ful(l)-*. Mateo Mendaza’s results are summarized by table 4 and 5:

Table 4. Index of productivity of *-isc* and *-cund* (Mateo Mendaza fc.-a)

Affix	N	n ₁	P
<i>-isc</i>	3,971	256	0.064
<i>-cund</i>	1,021	41	0.04
P = n ₁ / N	–	–	1.00000

Table 5. Index of productivity of *-ful(l)* and *ful(l)-* (Mateo Mendaza fc.-b)

Affix	N	n ₁	P
<i>-ful(l)</i>	3,887	198	0.05
<i>ful(l)-</i>	748	0	0
P = n ₁ / N	–	–	1.00000

As the comparison with the data in table 4 and 5 evidences, *-ettan* is slightly more productive than *-isc* (0.065 vs. 0.064) and clearly more productive than *-cund* (0.065 vs. 0.04), while *-laecan* is even less productive than *-isc*, *-cund* and *-ful(l)* (0.009 vs. 0.064, 0.04 and 0.05), but still more productive than the prefix *ful*, which Mateo Mendaza (fc.-b) considers totally unproductive. It can be said, on the face of the evidence, that the suffix *-ettan* is relatively productive, while *-laecan* is nearly unproductive.

To finish up this section, the index of global productivity of the Old English verbal suffixes *-ettan* and *-laecan* is offered by text category. Indeed, *-ettan* and *-laecan* verbs appear in the corpus in prose, glosses and poetry as in shown in table 6. In *The Dictionary of Old English Corpus* there are a 2,952,566 words, and, by text category, 2,105,363 words in prose texts, 729,558 in glosses and 117,645 words in poetry:

Table 6. Token frequency of *-ettan* / *-laecan* by text type based on *The Dictionary of Old English Corpus*

Text type	Affix tokens	Word total	Token frequency	no. of texts
Prose	171 / 723	2,105,363	0,000081 / 0,000343	32 / 60
Glosses	196 / 360	729,558	0,00027 / 0,000493	30 / 39
Poetry	29 / 13	117,645	0,00025 / 0,000111	16 / 6

As the figures in table 6 indicate, glosses have the highest token frequency of *-ettan* verbs, followed by poetry and prose texts, which exhibit the lowest frequency of token. On the other hand, gloss texts display the highest token frequency of *-laecan*

verbs, followed by prose and poetry texts, the latter yielding the lowest rate of token frequency. Comparing both affixes, the affix *-læcan* exhibits more tokens in prose and glosses than *-ettan*, but very few occurrences in poetry, whereas the affix *-læcan* displays the highest token frequency in glosses and in prose, but the lowest one in poetry. Table 7 offers the indexes of productivity of *-ettan* and *-læcan* by text type. Notice that V stands for the number of different inflective forms found for each verb:

Table 7. Index of productivity of *-ettan* / *-læcan* by text type based on *The Dictionary of Old English Corpus*

Text type	Hapaxes	Affix tokens	P	Types (V)
Prose	15 / 4	171 / 723	0,087 / 0,0055	118 / 163
Glosses	10 / 6	196 / 360	0,051 / 0,016	137 / 145
Poetry	1 / 0	29 / 13	0,034 / 0	27 / 10

The measure on productivity P by text coincides with the general measure given above, that is, the suffix *-ettan* is more productive than *-læcan* in the three text types considered. The reason, however, is different from the one found for the productivity of these affixes in general. Considering productivity in general, *-ettan* presents more types while *-læcan* has more tokens. By text type, the higher productivity of *-ettan* is the result of the higher figure of hapax legomena that *-ettan* displays, although *-læcan* yields more types than *-ettan*.

Before presenting global productivity by text type, I offer global productivity in general. Figure 6 represents global productivity P*, which relates the types found in *The Dictionary of Old English Corpus* (V) to the index of productivity:

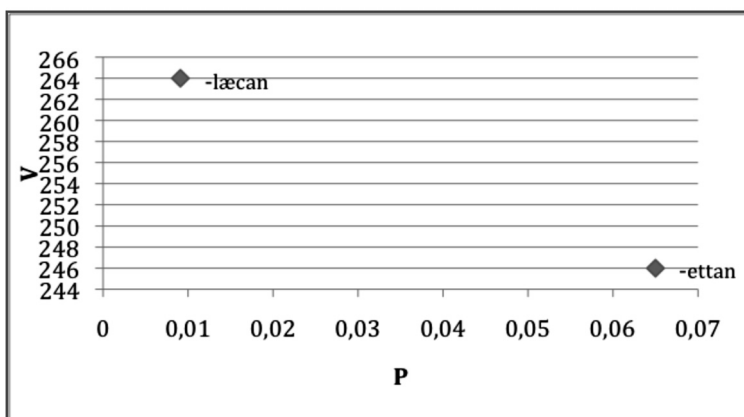


Figure 6. Global productivity of *-ettan* and *-læcan* verbs (based on *The Dictionary of Old English Corpus*)

In figure 6, V represents the number of types of *-ettan* and *-læcan* in general, without distinguishing by text type. The suffix *-læcan* displays the highest index of V (there are 264 types of *-læcan* and 246 of *-ettan*) while the suffix *-ettan* yields the highest P figure (0,065 as opposed to 0,009 in *-læcan*). Interestingly, the verb with the higher index of V throws the lower index of P and, conversely, the lower index of P corresponds to the higher index of V. That is, *-ettan* is more productive, while displaying fewer types, than *-læcan*, which throws a higher rate of V. The situation is comparable by text type. Figure 7 represents the global productivity P* by text type (P):

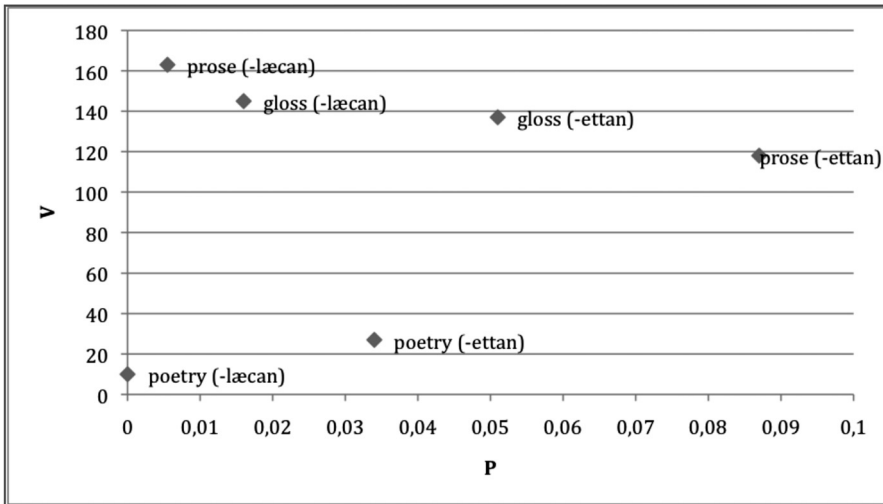


Figure 7. Global productivity of *-ettan* and *-læcan* by text type of *The Dictionary of Old English Corpus*

As can be seen in figure 7, the suffix *-ettan* in prose texts is the most productive (0,087), while the affix *-læcan* in poetry is not productive at all, that is, it turns out a measure of 0. Gloss texts occupy an intermediate position, given that they are less productive than *-ettan* in prose but more productive than *-læcan* in poetry. All productivity measures, therefore, coincide in that the affix *-ettan* is relatively productive while the suffix *-læcan* is practically unproductive. Overall, *-læcan* has a higher V value than *-ettan*.

4. CONCLUSION

I have begun this article by reviewing the study in low morphological productivity by Fernández-Domínguez *et al.* (2007). These authors raise the question of whether low productivity figures correspond to low productivity or simply indicate lack of

measurement accuracy. On this query, this research has shown a remarkable coincidence in the accounts on the productivity of the Old English verbal affixes *-ettan* and *-læcan*. As a general assessment of the productivity of these affixes, it can be said that *-ettan* is more productive than *-læcan*, even though the indexes of productivity thrown by both affixes are extremely low. In terms of type, the frequency of *-ettan* is higher than the one of *-læcan*. Regarding token, *-læcan* displays a higher token frequency. That is, *-ettan* is more frequent in terms of type whereas *-læcan* is more frequent if the measure of frequency is based on token. These findings are in accordance with Bauer's (2004: 102) remark that "lack of productivity of an affix is said to lead to a high token frequency of that affix, but a low type frequency". While frequency is directly proportional to items (types or tokens), productivity P is in direct proportion to hapax legomena, with which the index of productivity P of *-ettan* is higher than the one of *-læcan*, as a result of the more numerous hapaxes of the former suffix. A comparison with other Old English suffixes with low productivity indexes based on Mateo Mendaza (fc.-a, fc.-b) evidences that *-ettan* is slightly more productive than *-isc* and far more productive than *-cund*, while *-læcan* is even less productive than *-isc*, *-cund* and *-ful(l)*, but more productive than the zero productivity prefix *ful*. This comparison leads to the conclusion that the suffix *-ettan* is relatively productive, whereas *-læcan* is nearly unproductive. This statement is confirmed by the analysis productivity P by text type: the suffix *-ettan* is more productive than *-læcan* in prose, glosses and poetry, although for a different reason. If productivity P is measured in general, *-ettan* presents more types while *-læcan* has more tokens. By text type, the higher productivity of *-ettan* results from the higher number of hapax legomena of *-ettan*, although *-læcan* presents more types than *-ettan*. Finally, the measures of global productivity P*, in general and by text type, coincide in that the affix *-ettan* is relatively productive while the suffix *-læcan* is practically unproductive. That is, *-ettan* is more productive than *-læcan* in terms of frequency, productivity P and global productivity P*.

NOTE

- * Correspondence to: Gema Maíz Villalta. E-mail: gema.maiz@gmail.es
This research has been funded through the project FFI2008-04448/FILO.

REFERENCES

- Aronoff, M. 1976. *Word Formation in Generative Grammar*. Cambridge, Mass: MIT Press.
- Baayen, H. 1989. *A Corpus-Based Approach to Morphological Productivity*. Doctoral dissertation. Free University of Amsterdam.
- Baayen, H. 1992. "Quantitative aspects of morphological productivity". *Yearbook of Morphology* 1991. Eds. G. Booij and J. van Marle. Dordrecht: Kluwer. 109-149.

- Baayen, H. 1993. "On frequency, transparency and productivity". *Yearbook of Morphology* 1992. Eds. G. Booij and J. van Marle. Dordrecht: Kluwer. 181-208.
- Baayen, H. and R. Lieber. 1991. "Productivity and English derivation: a corpus-based study". *Linguistics* 29, 801-843.
- Baayen, H. and A. Renouf. 1996. "Chronicling *The Times*: productive lexical innovations in an English newspaper". *Language* 72, 69-96.
- Bauer, L. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press
- Bauer, L. 2004. *A Glossary of Morphology*. Edinburgh: Edinburgh University Press.
- Bauer, L. 2005. "Productivity: Theories". *Handbook of Word-Formation* 2005. Eds. P. Stekauer, and R. Lieber. Springer.
- Bosworth, J. and T. N. Toller. 1973 (1898). *An Anglo-Saxon Dictionary*. Oxford: Oxford University Press.
- Clark Hall, J. R. 1996 (1896). *A Concise Anglo-Saxon Dictionary*. Toronto: University of Toronto Press.
- Fernández-Domínguez, J., Díaz-Negrillo, A. and P. Štekauer. 2007. "How is low morphological productivity measured?". *Atlantis* 29 (1): 29-54.
- Healey, A., ed., with J. Price Wilkin and X. Xiang. 2004. *The Dictionary of Old English Web Corpus*. Toronto: Dictionary of Old English Project, Centre for Medieval Studies, University of Toronto. [available online in <http://tapor.library.utoronto.ca/doecorpus/>]
- Jember, G. K. et al. 1975. *English-Old English, Old English-English Dictionary*. Boulder, Colorado: Westview Press.
- Kastovsky, D. 1986. "Deverbal nouns in Old and Modern English: from stem-formation to word-formation". *Historical Semantics-Historical Word Formation*. Ed. J. Fisiak. Berlin: Mouton de Gruyter. 221-261.
- Kastovsky, D. 1989. "Typological changes in the history of English morphology". *Meaning and Beyond. Ernst Leisi zum 70. Geburtstag*. Eds. U. Fries and M. Heusser. Tübingen. 281-293.
- Kastovsky, D. 1990. "The typological status of Old English word formation". *Papers from the 5th International Conference on English Historical Linguistics*. Eds. S. Adamson, V. Law, N. Vincent and S. Wright. Amsterdam: John Benjamins. 205-224.
- Kastovsky, D. 1992. "Semantics and vocabulary". *The Cambridge History of the English Language I: The Beginnings to 1066*. Ed. R. Hogg. Cambridge: Cambridge University Press. 290-408.
- Kastovsky, D. 2005. "Conversion and/or zero: word-formation theory, historical linguistics, and typology". *Approaches to Conversion/Zero Derivation*. Eds. L. Bauer and S. Varela. Münster: Waxmann. 31-50.
- Kastovsky, D. 2006. "Typological changes in derivational morphology". *The Handbook of The History of English*. Eds. A. Van Kemenade and B. Los. Oxford: Blackwell. 151-177.
- Lass, R. 1994. *Old English: A Historical Linguistic Companion*. Cambridge: Cambridge University Press.

- Maíz Villata, G. fc.
- Marckwardt, A. H. 1942. "The verbal suffix *-ettan* in Old English". *Language* 18: 275-281.
- Martín Arista, J. 2008. "Unification and separation in a functional theory of morphology". *Investigations of the Syntax-Semantics-Pragmatics Interface*. Ed. R. Van Valin. Amsterdam: John Benjamins. 119-145.
- Martín Arista, J. 2009. "A typology of morphological constructions". *Deconstructing Constructions*. Eds. C. Butler and J. Martín Arista. Amsterdam: John Benjamins. 85-115.
- Martín Arista, J. *et al.* 2009. *Nerthus: An Online Lexical database of Old English*. [Available online in <http://www.nerthusproject.com>].
- Martín Arista, J. 2010a. "OE strong verbs derived from strong verbs". *SKASE Journal of Theoretical Linguistics* 7 (1): 36-56.
- Martín Arista, J. 2010b. "Lexical negation in Old English". *NOWELE-North-Western European Language Evolution* 60/61: 89-108.
- Martín Arista, J. 2010c. "Building a lexical database of Old English: issues and landmarks". *Current Projects in Historical Lexicography*. Ed. J. Considine. Newcastle: Cambridge Scholars Publishing. 1-33.
- Martín Arista, J. 2011a. "Projections and constructions in functional morphology: The case of Old English *HREOW*". *Language and Linguistics* 12 (2): 393-245.
- Martín Arista, J. 2011b. "Adjective formation and lexical layers in Old English". *English Studies* 92 (3): 323-344.
- Martín Arista, J. 2011c. "Morphological relatedness and zero alternation in Old English". *Morphosyntactic Alternations in English*. Eds. C. Butler and P. Guerrero Medina. London: Equinox.
- Martín Arista, Javier. fc a. "Parasynthesis in Old English word-formation".
- Martín Arista, J. fc b. "Old English lexical primes: Corpus analysis and database compilation". *Creation and Use of Historical Linguistic Corpora in Spain*. Ed. N. Vázquez. Newcastle: Cambridge Scholars.
- Martín Arista, J. fc c. "The Old English prefix *ge-*: A panchronic reappraisal".
- Mateo Mendaza, R. fc a. "The Old English adjectival suffixes *-cund* and *-isc*: textual occurrences and productivity".
- Mateo Mendaza, R. fc b. "The Old English adjectival affixes *-ful* and *ful-*: a text-based account on productivity".
- Plag, I. 1999. *Morphological Productivity. Structural Constraints in English Derivation*. Berlin/New York: Mouton de Gruyter.
- Quirk, R. and C. L. Wrenn. 1994. *An Old English Grammar*. London: Methuen.
- Štekauer, P. 2005. *Meaning Predictability in Word-formation: Novel, Context-Free Naming Units*. Amsterdam/Philadelphia: John Benjamins.
- Sweet, H. 1976 (1896). *The Student's Dictionary of Anglo-Saxon*. Cambridge: Cambridge University Press.

