

# Productivity in Italian word formation: a variable-corpus approach\*

LIVIO GAETA AND DAVIDE RICCA

## Abstract

*The quantitative approach to morphological productivity first proposed by Baayen crucially refers to the relation between the number of hapax legomena formed with a given affix occurring in a sufficiently large corpus and the total number of tokens of that affix sampled in the corpus. Most criticism against this measure focuses on its neglecting the role played by frequency in the evaluation of productivity. As an improvement of Baayen's procedure, a variable-corpus approach is proposed. Accordingly, the productivity values should be calculated at equal token numbers for different affixes instead of taking the different token numbers which result from sampling the whole corpus for all affixes, as in Baayen's works. This implies that variably-sized subcorpora must be sampled to compare affixes displaying different frequencies. On the basis of a 75-million-token newspaper corpus, the productivity values for several Italian affixes in the deverbal and deadjectival domain are calculated. The resulting rank proves linguistically plausible, avoiding the overestimation of productivity for low-frequency affixes typically occurring in fixed-corpus calculations. As a further advantage, the procedure proposed here makes it possible to deal satisfactorily with two problematic aspects usually neglected in previous investigations, namely, the quantitative impact of (i) allomorphies and lexicalizations and (ii) inner-cycle derivations on productivity measures.*

## 1. Introduction: productivity as a quantitative notion

In a number of recent contributions, Baayen (1989, 1992, 1993, 2001; see also Baayen and Lieber 1991; Baayen and Renouf 1996; Plag et al. 1999) has suggested relating the notion of productivity to the number of hapax legomena, that is, words with frequency 1, occurring in a sufficiently large corpus. The proposed measure of productivity  $P$  for a given affix is the

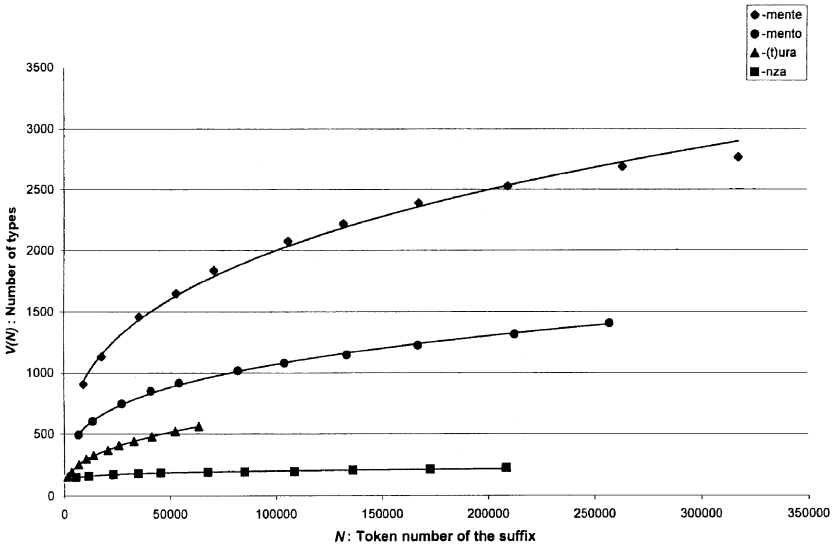


Figure 1. *La Stampa* 1996–1998: types growth curve as a function of  $N$

ratio between the number  $h$  of hapax legomena derived by that affix and the number  $N$  of all tokens of that affix occurring in the corpus:

$$(1) \quad P = h/N$$

In mathematical terms, it can be shown (Baayen 1989: 104) that the index (1) is the derivative at point  $N$  of the curve  $V(N)$ , which plots the type number  $V$  for a given affix (i.e. the number of different words derived by that affix) as a function of the token number  $N$  of the same affix. To get a concrete illustration, four instances of the curve  $V(N)$  are reported in Figure 1, taken from our data: they refer to the Italian suffixes *-mente*, forming adverbs, and *-mento*, *-(t)ura*, and *-nza*, forming action nouns, sampled from three years of the Italian newspaper *La Stampa*.<sup>1</sup>

In simpler terms, the ratio in (1) measures the probability of encountering a new type not attested before, namely, a hapax legomenon, after that  $N$  tokens of a given affix have been sampled (Baayen 1989: 99–100, 2001: 156–157). The curve  $V(N)$  in Figure 1 describes the growth of the lexical inventory of an affix. The measure of the slope of the curve, that is, the derivative at a certain point, gives the speed at which new types of a certain affix come out from the sample.

If an affix is even minimally productive, new types will be encountered: the value of  $V$  may only increase as  $N$  increases — mathematically it is a

nondecreasing monotonic function. However, for every affix the increasing rate of  $V(N)$  will decrease as we proceed in the sample, since it will become more and more probable that new tokens of the affix will be occurrences of already attested types. Hence, the productivity  $P(N)$  is a decreasing monotonic function of  $N$ .

It is evident from Figure 1 that the curves  $V(N)$  for the four suffixes increase at different rates, thus qualifying for different values of productivity. Whereas the curve of the suffix *-nza* immediately reaches almost the whole number of possible types and then remains stable, approximating a horizontal line, for the other suffixes the curve was clearly still increasing, although with different slopes, when we ended our sampling.

This is in essence the quality of the index  $P$  proposed by Baayen: investigating the increasing rate of new types formed with a certain affix in a corpus provides a clue for measuring the availability (*disponibilité* in Corbin 1987: 177) of a certain word formation rule. From this perspective, the basic objection raised against this value by van Marle (1992: 156), who does not “see what kind of direct relationship there is between the chance that a given rule is put into action and the frequency with which the words that have already been produced by that rule are used” is not really relevant. The point is not that “[o]nce a word is coined, the frequency of the use of that word [...] is more or less irrelevant to the degree of productivity of that rule.” The value  $P$  rather aims at measuring the speed at which the lexical inventory of the types formed with that affix is enriched.

Notice, moreover, that the curves  $V(N)$  display different lengths. This is due to the different token frequency of the four suffixes in the corpus: whereas *-mento* and *-mente* rank among the most frequent Italian affixes, *-nza* occupies an intermediate position and *-(t)ura* is much less frequent. As we will see later in section 2, this aspect is not taken into account in Baayen’s procedure, where the index  $P$  is always calculated — no matter how frequent an affix is — referring to the number of tokens  $N$  sampled in the whole corpus.

Several criticisms have been raised against this approach. First, the size of the corpus is questionable. A minimal threshold is made necessary by the nature of hapax legomena of being rare events, but not necessarily new formations (cf. Dal 2003: 17–18). Only when the corpus becomes large enough (as in the case of Baayen and Renouf’s [1996] *Times* corpus), we expect new sampled types to be mostly true neologisms, or at least words not attested in a large dictionary.<sup>2</sup>

A number of further problems arises about what to include in the sample. First, it is not always evident what has to be considered as a distinct

type of a certain affix. As observed by Bauer (2001: 151), should we treat the occurrences of *-ment* in *development* and *underdevelopment* as belonging to one single type, or should we rather posit an autonomous type for the latter ones (considering them as derivations from a distinct verb *underdevelop*)? The question is crucial because the second analysis provides us with a large source of potential hapax legomena that could increase the productivity rate of an affix. Moreover (cf. Plag 1999: 28), are words like *entity* and *celebrity* to be included in the sample of the suffix *-ity*? Although such words are to a certain extent analyzable, they cannot be related to existing bases, at least synchronically. A similar question can be raised for morphologically transparent, but semantically opaque words, like for instance *professor* with respect to *profess*. Since such words are usually quite frequent, their inclusion into the sample can strongly alter the productivity rate.

Similar problems also arise for the tokens. In particular, Plag (1999: 29) has put forward the question whether one should also include in the counts the occurrences of an affix in inner derivational cycles: for instance, an occurrence of *conventionalization* clearly counts as a token of the suffix *-ation*, but should it also be considered as a token of the inner suffix *-ize*? Plag answers negatively; however, to what extent is this procedure justified not only for suffixation, but especially for prefixation? Thus, should really a word like *reprintable* (clearly a token of the suffix *-able*) not be counted also as a token of the prefix *re-*?

Probably, the main objection raised against Baayen's approach concerns his free comparison of affixes with very different type and token frequency, which has relevant effects on productivity. For instance, van Marle (1992: 154–155) has observed that, with respect to the Dutch agent noun suffixes *-er/-ster*, the masculine suffix *-er* turns out to display a much lower productivity rate with respect to *-ster*, its much less frequent feminine counterpart. Since both suffixes select the same base domain and fulfill the same function, these diverging productivity rates are quite unexpected. Similar considerations would apply for the Italian pair *-(t)ore/-trice*, as will be seen later in section 3.

In response to van Marle's objections, Baayen (1994: 458) supports the asymmetry between *-ster* and *-er* by experimental evidence, showing that in a production task along the lines of Anshen and Aronoff (1988), subjects produce many more novel forms (i.e. not attested in a large corpus and/or in a comprehensive dictionary) for *-ster* than they do for *-er*. These data point to a rather different interpretation of *P*: namely, the higher value of *P* for *-ster* is related to the fact that it occurs in a much lower number of established formations than *-er*. Hence, in the given experimental conditions there is a relatively low probability for *-er* of giving

rise to new formations, while most of the potential domain for *-ster* is not (yet) exploited.

Thus, *P* data actually portrait a higher degree of saturation (*Sättigungsgrad* in Rainer 1993: 32) for *-er* formations referred to the domain of established base words. However, the different degree of saturation does not tell much about the relative probability of the two suffixes to apply to a new lexical entry selected as possible input, other things being equal (*Wahrscheinlichkeit der Regelanwendung* in Rainer 1993: 32). It seems to us that this latter facet of productivity matches more closely most approaches to the concept.

Baayen has proposed two further quantitative approaches to productivity. First, the notion of “global productivity” is introduced by plotting a number of English affixes in a two-dimensional space whose coordinates are precisely *P* — the productivity as defined in (1) — and *V* — the number of different types (see Baayen and Lieber 1991: 819, Figure 3). The global productivity should capture not only the availability of an affix, but also its profitability (*rentabilité* in Corbin 1987: 177), that is, the extension of the base domain to which it applies. However, by this method it is possible to rank the productivities only for those affixes where *P* and *V* correlate: namely, for those located approximately along a diagonal line from the bottom left-hand corner of the chart (the unproductive ones) up to the top right-hand corner (the most productive ones). There is no way, as Baayen (1992: 124) admits, to compare quantitatively the productivities of affixes with high *V* but low *P* (say, *-ity* in Baayen and Lieber’s 1991 data) with those exhibiting the opposite pattern (as *-ish* in the same source). Thus, as Bauer (2001: 154) notes, this procedure “fails to show that there is a vital relationship between the two measures, and rather implies that the two should be kept entirely separate. Yet, intuitively, the type-frequency would be thought to influence the probability of encountering new types.”

A third measure proposed by Baayen (1993: 192) quantifies the notion of degree of productivity in terms of the contribution of a given affix to the growth rate of the vocabulary as a whole. This is the “hapax-conditioned degree of productivity” *P\**, which considers the number of hapaxes formed with a certain affix in a corpus divided by the global number of hapaxes within the corpus. Since the latter remains constant for the different affixes, in relative terms *P\** turns out to coincide with the simple number of hapaxes. Baayen (1993: 205) sees this measure as “particularly suited to ranking productive processes according to their degree of productivity.”

Against the linguistic significance of *P\**, Bauer (2001: 155) objects that *P\** “asks ‘What proportion of new coinages use affix A?’ rather than

asking ‘What proportion of words using affix A are new coinages?’. It is this latter which seems a more relevant question to ask.’ In our opinion, there is no doubt that  $P^*$  makes full sense as a fair measure of the number of neologisms found in a corpus (Baayen and Renouf 1996). Moreover, as Baayen (pers. comm.) pointed out to us, in practice  $P^*$  often displays a strong correlation with the measure we are going to propose in this work (see [4] below). However, from a conceptual point of view, we share Bauer’s idea that it should be definitely desirable to maintain the insightful notion of productivity as related to the growth curve of each given affix, as expressed by Baayen’s  $P(N)$ , rather than by  $P^*$ .

In section 2 we will propose a different procedure for calculating  $P(N)$ , namely, to evaluate it for different affixes at equal values of  $N$ . We are confident to show that in this way the linguistic significance of Baayen’s  $P$  is kept intact, and at the same time, the shortcomings of his original approach can be avoided. In section 3 we will apply our procedure to a number of Italian affixes, providing a quantitative ranking of their productivities. Section 4 deals with the question of what to include into the counts, especially in relation to allomorphies and lexicalizations. Section 5 tackles the problematic aspect of the impact of inner cycle derivations on the evaluation of productivity. The final section 6 draws the general conclusions and raises some open questions for further research.

## 2. A variable-corpus approach

As hinted at above, many objections to the first measure of productivity proposed by Baayen, namely the ratio  $P = h/N$ , focus on the same shortcoming: the ratio  $h/N$  does not seem to give meaningful results if, *in a given corpus*, one compares the results obtained for affixes with very different token frequencies. However, in what follows we hope to show that this sort of objections does not invalidate the method itself, but only the way of applying it. The point is that, for each affix,  $P(N)$  is not constant, but is a decreasing function of  $N$ , even tending to zero when  $N$  approaches infinity (Baayen and Lieber 1991: 837). The shape of the function  $P(N)$  is shown in Figure 2, for the same suffixes whose type curves  $V(N)$  have been represented in Figure 1. As in Figure 1, the horizontal axis reports the number of tokens of the four suffixes, and consequently the endpoint of the curves lies at different values of  $N$ , since the more frequent suffixes obviously exhibit a higher value of  $N$  when the whole corpus is sampled.

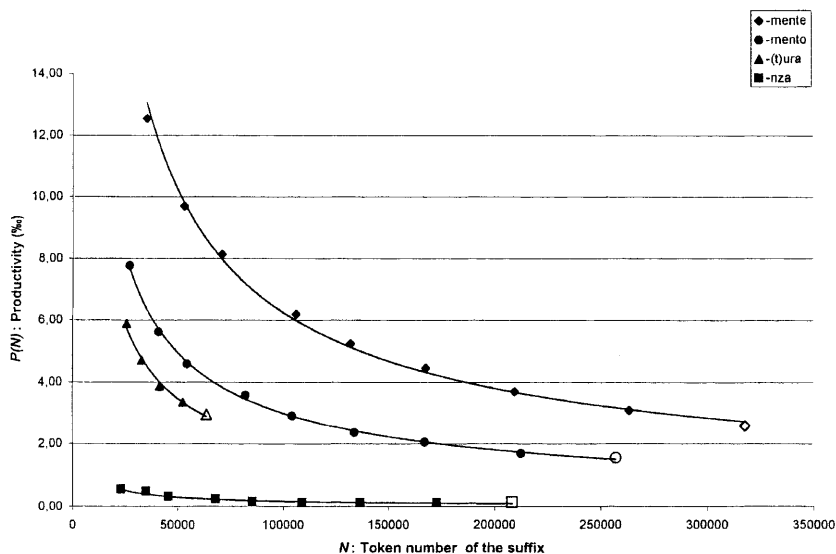


Figure 2. *Productivity as a function of N*

Now, comparing the  $P$  values referring to the whole corpus for all suffixes would mean comparing the values at the endpoint of each curve: in Figure 2 they have been emphasized by a bigger size for clarity. However, the endpoints correspond to different values of the independent variable  $N$ , that is, the number of tokens of the suffixes under investigation. And due to the decreasing character of all  $P$  curves, such a procedure will always imply an overestimation of the values of  $P$  for the less frequent suffixes, which can reach dramatic proportions if the affixes to be compared show great difference in token frequency, as is the case for van Marle's example of Dutch *-er* vs. *-ster*.

Precisely a direct comparison of affixes with very different frequency has been actually made in Baayen and Lieber (1991), and is responsible for the less convincing results obtained there: for instance, referring to the 18-million CELEX corpus,<sup>3</sup> Baayen and Lieber found out that among the deverbal suffixes, *-ee* outranks *-er* in productivity by a factor of about two. Despite the discussion given by the authors, this result looks plainly at odds with most linguists' expectations, but it is simply the consequence of the great unbalancing in token frequency between the two suffixes involved (a ratio of about 1:48). We will show that these counterintuitive results disappear if  $P$  is calculated for equal values of  $N$ .<sup>4</sup>

Since the affixes do not have the same token frequency, to compute  $P$  for equal values of  $N$  means to perform calculations for different

affixes on corpora of different size. Therefore, we had first of all to choose a global corpus which could easily be divided in smaller subcorpora keeping intact its textual character. A corpus of this kind (matching the choice made by Baayen and Renouf 1996) is given by the continuous issues of a daily newspaper: we chose three years from *La Stampa*, from 1996 to 1998, available on compact disc and easily exportable on ASCII files to be treated with a text analysis software (DBT<sup>TM</sup> by E. Picchi, CNR Pisa). Moreover, we judged a newspaper-based corpus to be particularly adequate for quantitative studies in derivational morphology, because it contains a mixture of very different speech registers and text types.

The full corpus, after elimination of spurious material as databases, short summaries, and titles of the articles, contains around 75,000,000 tokens (proper nouns, dates, and some other morphologically nonrelevant items included). It has been structured in 36 chunks — hereafter termed “subcorpora” — of progressively increasing size (1 to 36 months). For each subcorpus independently, through the DBT software and further computer manipulation,<sup>5</sup> we built a complete list of wordforms in direct and inverse alphabetical order, with each wordform carrying its token frequency. From these lists, all the occurrences of a given affix in a given subcorpus could be extracted and lemmatized, and finally made ready for type/token/hapax calculations, after an unavoidable and much time-consuming manual check.<sup>6</sup> This last stage is obviously necessary to eliminate all endings which are not suffixes, to group all misprints together with their correct type (otherwise they would have heavily distorted the data, being mostly hapaxes!) and, last but not least, to provide with identifying labels all those formations which, although relatable to the given affix, do not exhibit full semantic and/or morphotactic transparency (for some illustration, see section 4 below).

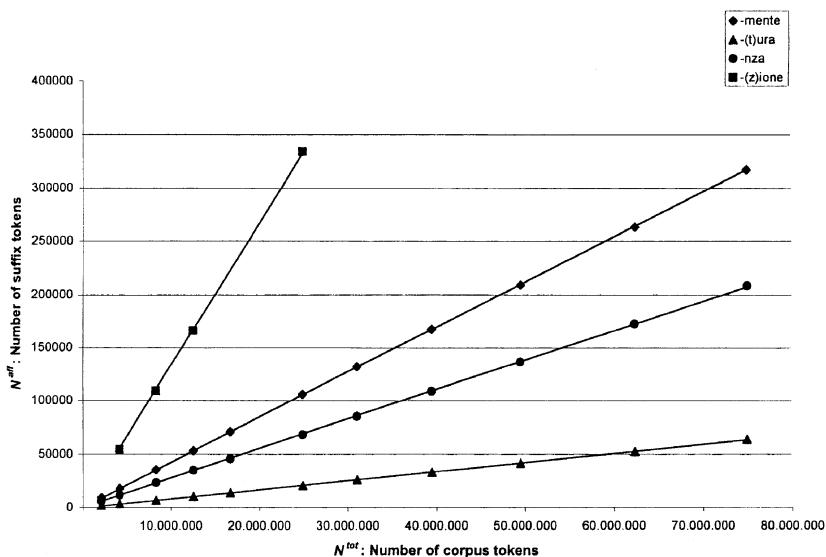
As said above, data obtained by means of our “variable corpus” procedure will be linguistically meaningful only if the subcorpora are uniform in terms of their textual typology. It is to be expected that the distribution of textual types found, say, in six months of *La Stampa* will not differ appreciably from the one found in the entire three-year corpus, and an important consequence of this reasonable assumption can also be quantitatively checked. Table 1 shows, for some of the suffixes investigated, that their token frequency remains substantially stable while the chunk size increases from a minimal subcorpus of just two months up to the full 36-month corpus.

To be sure, it is not to expect that each single lexical item is so evenly distributed throughout the corpus. Table 1, however, effectively shows that the averaging contribution of the different types belonging to the



Table 1. *Checking subcorpus uniformity*

Suffix	Token frequency (%) for different subcorpus sizes					
	2 months 4,162,397 tok.	4 months 8,302,320	6 months 12,535,480	12 months 24,915,369	24 months 49,485,568	36 months 74,917,798
-(z)ione	13.0	13.1	13.3	13.4	13.7	13.9
-mente	4.32	4.29	4.24	4.26	4.23	4.24
-nza	2.83	2.81	2.80	2.73	2.76	2.78
-(t)ura	0.82	0.84	0.84	0.83	0.84	0.85

Figure 3. *Stability of suffix frequencies*

same affix leads very early to stable data concerning its overall token frequency. In other words, for every given affix, the number of its tokens can be safely taken as being directly proportional to the total number of tokens in the (sub)corpus. The graphical equivalent of Table 1 is given in Figure 3, which plots the number of tokens of each affix ( $N^{affix}$ ) against the total number of tokens of the currently sampled corpus ( $N^{tot}$ ) as long as the sampling goes on. The higher the slope, the greater the token frequency of the affix; but the linear ratio between  $N^{affix}$  and  $N^{tot}$  holds for all the affixes investigated throughout the sampling process.

For each affix, a curve  $P(N)$  can be drawn by fitting the discrete values calculated on the subcorpora. Then the values for a fixed value of  $N$ , say

$N_0$ , can be obtained by interpolation<sup>7</sup> and compared for the different affixes. The same procedure may be repeated for different values of  $N$ , thus allowing us to compare productivity ratios of two affixes at different points of their curves.

### 3. Ranking productivities for Italian deverbal and deadjectival word formation

In this section the main results of our investigation will be presented. We dealt with two of the three main categorial domains of Italian word formation, namely, deverbal and deadjectival derivation. Denominal derivation has been left out for the sake of maximizing comparability between the data. In fact, the denominal domain as a whole is far larger than the other two (as can be roughly estimated by means of a large dictionary<sup>8</sup>), which might in principle have relevant consequences on the linguistic evaluation of productivity data.<sup>9</sup> More data on other Italian derivational processes, including many denominal affixes, can be found in Gaeta and Ricca (2003a). However, the number of affixes included here is largely sufficient to discuss the methodological aspects of our work, which is the main focus of this article.

Within both domains, we selected a significant group among the most frequent affixes that could be segmented via a partially automatic procedure, namely:

- (2) For the deverbal domain:
- a. the suffixes *-mento*, *-(z)ione*, *-(t)ura*, *-aggio*, *-nza* forming action nouns:
 

<i>cambiare</i>	→	<i>cambiamento</i>	‘change’
<i>trasformare</i>	→	<i>trasformazione</i>	‘transformation’
<i>mappare</i>	→	<i>mappatura</i>	‘mapping’
<i>lavare</i>	→	<i>lavaggio</i>	‘washing’
<i>decadere</i>	→	<i>decadenza</i>	‘decay’
  - b. the adjectival suffixes *-bile* ‘-able’ and *-evole*:
 

<i>lavare</i>	→	<i>lavabile</i>	‘washable’
<i>mancare</i>	→	<i>manchevole</i>	‘faulty’
  - c. the prefix *ri-* ‘re-’:
 

<i>giocare</i>	→	<i>rigiocare</i>	‘play again’
<i>dare</i>	→	<i>ridare</i>	‘give back’/‘give again’
  - d. the suffixes *-(t)ore* and *-trice* forming masculine/feminine agent and instrument nouns and also deverbal adjectives:

- giocare* → *giocatore/giocatrice* ‘player’/‘player (f.)’  
*calcolare* → *calcolatore* ‘computer’  
                   → *calcolatrice* ‘pocket calculator’  
*uno sguardo rivelatore* ‘a revealing (m.) glance’  
*un’osservazione rivelatrice* ‘a revealing (f.) observation’
- (3) For the deadjectival domain:
- the suffixes *-ità*, *-ezza* forming quality nouns:  
*vero* → *verità* ‘truth’  
*bello* → *bellezza* ‘beauty’
  - the negative prefix *in-* ‘un-’/‘in-’:  
*utile* → *inutile* ‘useless’
  - the adverbializing suffix *-mente* ‘-ly’:  
*fermo* → *fermamente* ‘firmly’
  - the elative suffix *-issimo*:  
*lungo* → *lunguissimo* ‘very long’

All affixes listed above are rather frequent in texts, but certainly not to the same extent: the ratio between the highest and the lowest token frequencies (those of *-(z)ione* and *-evole* respectively) is more than 50:1. The complete list of token frequencies for the affixes investigated is given in Table 2. All data are calculated referring to occurrences in the outmost derivational cycle only. As discussed later in section 5, the inclusion of

Table 2. *Token frequency for some important Italian derivational affixes*

Affixes	$N_{max}$ (= $N$ in the whole corpus)	Token frequency (%)
<i>-(z)ione</i>	1,043,979	13.9
<i>-ità/-età</i>	356,857	4.8
<i>-mente</i>	317,725	4.2
<i>-(t)ore</i>	273,706	3.7
<i>ri-</i>	270,066	3.6
<i>-mento</i>	257,216	3.4
<i>-nza</i>	208,365	2.8
<i>in-</i>	146,982	2.0
<i>-bile</i>	102,904	1.4
<i>-ezza</i>	69,090	0.9
<i>-(t)ura</i>	63,800	0.9
<i>-issimo</i>	51,636	0.7
<i>-trice</i>	23,780	0.3
<i>-aggio</i>	22,019	0.3
<i>-evole</i>	19,076	0.3

inner cycles would considerably increase the token frequency for some affixes (particularly for *-bile* and *ri-*), while for others (like *-mento*, *-issimo*, and *-mente*) it would be irrelevant.

We come now to the main point, namely ranking the productivity values for these affixes. In Table 3 the affixes are ordered according to their values of  $P$ , calculated for three different values of  $N$ : 19,000, 50,000, 100,000. This is done by two different methods. In the left columns, the values are calculated applying the method sketched in section 2, namely, resorting to real data taken from differently sized subcorpora for each affix. In the right columns (printed in italics), the values of  $P$ , for the same values of  $N$  as above, are calculated via the binomial interpolation procedure discussed in Baayen (2001: 63–65), which takes as input the whole frequency spectrum of each affix calculated at full corpus. The computational procedure to extract the expectation values for  $h(N)$  is provided in the CD-ROM enclosed with Baayen's book (cf. Baayen 2001: 223–225). It can be seen that the two rankings align substantially, which can be considered as an empirical validation of Baayen's interpolation procedure. On this basis, the latter can be safely employed in future research to calculate  $P(N)$  following the variable-corpus approach, with a substantial spare of time.

Due to the sharp differences in token frequency, not all the affixes can be compared directly together: the blanks in Table 3 correspond to values of  $N$  which are too high for the least frequent affixes, with no available data. The value  $N = 50,000$  is the most suitable to embrace the greatest number of affixes, but the lowest value of 19,000 is necessary to include the three less frequent ones, namely *-trice*, *-aggio*, and *-evole* respectively. On the other hand, with the most frequent affixes, data cannot be fully reliable if calculated for too low values of  $N$ . Take, for instance, the paramount case of *-(z)ione*. For such a frequent suffix, the value  $N = 19,000$  is reached after about 1,300,000 corpus tokens, corresponding to a subcorpus size of just twenty days. Clearly, hapax legomena in a corpus of only 1,300,000 tokens can hardly be taken a priori as instances of very rare words, let alone new formations: many of them will simply be words with an average frequency of 1:1,000,000, which is manifestly still too high a value to consider them not to be stored in the mental lexicon (cf. Baayen 1994: 453). The distorting effect of these "spurious" hapaxes will reduce itself progressively as long as  $N$  increases: this means that the comparison between values for  $P$  becomes more reliable — whenever possible — in the rightmost columns of Table 3. To emphasize the affixes mainly subject to this effect, in Table 3 we put the less reliable values in brackets, that is, those extracted from subcorpora under a threshold of 6,000,000 tokens (about three months of our newspaper).

Table 3. Italian derivational affixes ranked by productivity at different values of  $N$ 

Affixes	$P(N) \cdot 10^3$					
	$N = 19,000$		$N = 50,000$		$N = 100,000$	
	real data	<i>binomial interpolation</i>	real data	<i>binomial interpolation</i>	real data	<i>binomial interpolation</i>
<i>-issimo</i>	25.8	24.3	12.9	12.8		
<i>-mente</i>	(19.0)	(20.6)	10.1	9.9	6.4	6.0
<i>-bile</i>	11.3	11.9	6.3	6.4	4.1	4.1
<i>-ità/-età</i>	(13.4)	(13.9)	6.3	6.8	3.7	4.1
<i>-trice</i>	10.8	11.1				
<i>-(t)ore</i>	(9.4)	(10.6)	5.0	5.3	3.2	3.4
<i>-mento</i>	(10.4)	(9.8)	4.9	5.0	3.1	3.1
<i>-(z)ione</i>	(13.4)	(13.7)	(5.1)	(5.4)	2.7	3.0
<i>ri-</i>	(6.5)	(6.0)	3.8	3.3	2.3	2.2
<i>-(t)ura</i>	6.6	6.8	3.5	3.5		
<i>in-</i>	4.1	4.6	2.1	2.2	1.3	1.3
<i>-ezza</i>	2.7	2.5	1.3	1.2		
<i>-aggio</i>	1.5	1.5				
<i>-nza</i>	0.7	0.9	0.3	0.4	0.2	0.2
<i>-evole</i>	0.3	0.3				

The distortion discussed above explains why a couple of affixes do not appear to be ranked uniformly in the three columns of Table 3: the crossing between  $P(N)$  curves concern *-ità/-età* with respect to *-bile* and *-(z)ione* with respect to *-(t)ore* and *-mento*. Accordingly, we ranked them following the rightmost column. Expectedly, in both cases it is the suffix with the highest token and type frequency (*-(z)ione* and *-ità* respectively) whose  $P(N)$  appears to be overestimated for low  $N$ . It is also possible that in such cases the differences in productivity between the given affixes are too slight to allow for a linguistically significant ranking.

Let us give some general commentary to the ranking in Table 3 and discuss its linguistic plausibility. A first consideration concerns the two top scorers, *-issimo* and *-mente*. Their place at the top of the list makes perfect sense linguistically, since every linguist would agree that they represent borderline cases between inflection and derivation. There is no full consensus about which side they should be placed on: *-mente* is more often seen as derivational (but cf. Haspelmath [1996: 49–50] on its English equivalent *-ly*; for a discussion, see Ricca 1998) and *-issimo* as inflectional (at least within the Italian tradition; but cf. Rainer 2003). Therefore, no wonder that they exhibit a higher productivity than any “typically derivational” affix, with *-issimo*, arguably the more inflectional of the two, displaying a still higher value than *-mente*.

Results are equally reasonable at the other end of the list. The suffixes *-nza* and *-evole* are the only two instances within the group under consideration that would be labelled as nearly unproductive even from a qualitative point of view: although both of them are still well established in the lexicon (as their token and type frequencies in our corpus also confirm, see Table 4 below), few neologisms are registered in dictionaries in the last half century,<sup>10</sup> and this is reflected by the very low number of hapax legomena found in our corpus. True, some hapaxes occur, while for an ideally unproductive affix one would plainly hypothesize no neologisms at all; but, when scrutinized, these scarce instances turn out mainly to be very rare words and not really new coinages (e.g. *perdonanza* ‘forgiveness’, *afferenza* ‘membership’, *bisognevole* ‘needy’, *ammaestrevole* ‘exemplary’).<sup>11</sup> In other words, they represent a sort of unavoidable “ground noise” due to the double nature of hapax legomena, being either neologisms or simply rare words: as the ranking itself shows, this noise is not loud enough to distort the picture.

Looking now at the rest of Table 3, two interesting subgroups come into consideration, namely those of (nearly)-synonymous affixes competing in the same domain. The larger one is given by the five action noun suffixes. We have already mentioned *-nza* as being at most very marginally productive. All the remaining four (*-mento*, *-(z)ione*, *-(t)ura*, *-aggio*) have to be considered productive from a qualitative point of view (cf. Gaeta 2004: 322). This is an interesting situation of multiple near-synonymy within a single derivational category, and therefore, a case in which a reliable quantitative ranking would be particularly welcome. Our results partially achieve this task, as they clearly indicate a much lower productivity rate for *-aggio* with respect to the remaining three, even if in Table 3 *-aggio* can be safely compared only with *-(t)ura*, due to its very low token frequency. The ordering for the latter, namely *-mento* > *-(z)ione* > *-(t)ura*, looks plausible but it is not so neatly clear-cut to draw absolutely sure conclusions, especially for the first two suffixes. At any rate, it would be in agreement with most linguists’ expectations, since the lexical relevance of *-mento* with respect to *-(z)ione* has continuously grown up from Latin times onwards (see Thornton 1990–1991). However, nowadays *-(z)ione* enjoys the advantage of being the only choice for the very numerous neologisms in *-izzare* ‘-ize’ (see also section 5). The suffix *-(t)ura* is undoubtedly productive as well, but tends to privilege specialistic domains.

The other synonymous affixes are the pair *-ità*/*-ezza*, the two main devices in Italian to form quality nouns from adjectives. In this case, the ranking is clear, with *-ità* well above *-ezza*. This is again in full agreement with our expectations, since, as Rainer (1989: 269) has shown, *-ezza*

attaches productively only to adjectival formations in *-to* and *-evole*, and to derived bisyllabic bases (moreover, the latter two kinds of formations comprise very few new entries), while *-ità* is the dominant or unique choice for many adjectival derivatives, included the most productive types in *-oso*, *-ale* and *-bile* (Rainer 1989: 299).

Concerning the prefixes *ri-* and *in-*, Table 3 clearly places both of them within the productive segment of Italian derivation, although their relevance in productivity is less high than in token frequency, especially for *in-*. The values for *ri-* locate the prefix not far from the highly productive suffixes for action nouns, while *in-* falls clearly below the main deadjectival formations, though doubling the productivity of a still productive suffix like *-ezza*. As for the comparison between the two prefixes, the lower value for *in-* with respect to *ri-* matches linguists' expectations, since the former has a learned flavor and undergoes relevant semantic restrictions (cf. Gaeta and Ricca 2003b: 107).<sup>12</sup>

It is interesting to compare the ranking given in Table 3 with the one resulting from applying the original Baayen's procedure, namely calculating all  $P$ 's referring to the whole corpus, that is,  $P(N_{max})$ , and consequently for values of  $N$  different for each affix and directly proportional to their token frequency. This is shown in Table 4, which also reports the values for  $N$ ,  $V$ , and  $h$  for each affix when the full 36-months corpus has been sampled.

The differences between the two orderings are immediately apparent. If  $P$  is calculated referring to the whole corpus for all affixes, the less frequent

Table 4. *Italian derivational affixes ranked by productivity calculated for the full 36-months corpus (Baayen's procedure)*

Affixes	$P(N_{max}) \cdot 10^3$	$N_{max}$	$V(N_{max})$	$h(N_{max})$
<i>-issimo</i>	12.5	51,636	1697	643
<i>-trice</i>	9.4	23,780	645	224
<i>-bile</i>	4.0	102,904	1117	409
<i>-(t)ura</i>	3.0	63,800	561	189
<i>-mente</i>	2.6	317,725	2767	825
<i>-(t)ore</i>	1.7	273,706	1480	461
<i>-mento</i>	1.6	257,193	1405	402
<i>-ità/-età</i>	1.5	356,852	1962	544
<i>-aggio</i>	1.3	22,019	115	29
<i>ri-</i>	1.2	270,066	935	312
<i>in-</i>	1.0	146,982	767	148
<i>-ezza</i>	1.0	69,094	324	70
<i>-(z)ione</i>	0.5	1,043,979	2363	486
<i>-evole</i>	0.3	19,076	61	6
<i>-nza</i>	0.1	208,362	225	29

ones are heavily favored, due to the presence of *N* in the denominator of (1). Thus, for instance, the productivity for *-aggio* would turn out to be nearly equal to that of *-ità/-età*, a plainly counterintuitive result if one just looks at the great difference between the hapaxes of the two suffixes, mostly new coinages.

We will mention three arguably implausible results in Table 4. The first one concerns the dramatic splitting in productivity among the three action noun suffixes *-mento*, *-(z)ione*, and *-(t)ura*. Due to their great differences in token frequencies, the relatively unfrequent *-(t)ura* climbs up to a value even higher than that of the nearly-inflectional *-mente*, while conversely the huge suffix *-(z)ione*, burdened by its over one million tokens, sinks down, not far from the unproductive suffix *-evole*, which is manifestly absurd.

The second undesirable result concerns precisely *-mente*, since by this procedure its recognized borderline status between inflection and derivation loses any quantitative support. This suffix does not form anymore a pair with *-issimo* as in Table 3, where it displays only a slightly lower productivity (about 30% less); instead, it is outranked by the latter by a factor of five. On the other hand, three clearly derivational affixes sufficiently small in token frequency now rank above *-mente*. The most striking example is *-trice*, which even threatens the top position of *-issimo*. The case of *-mente* finds a close parallel in English *-ly*, included by Baayen and Renouf (1996) in their counts based on the *Times* corpus, which is remarkably similar to ours both in size and textual character. The *P* value for *-ly* is much lower than the ones for *-ness* and *un-* and even falls slightly below the values for the comparatively almost unproductive *-ity* and *in-*.<sup>13</sup>

A third questionable result concerns the comparison between the two agentive suffixes *-(t)ore* and *-trice*, which are roughly parallel to the already mentioned case of the Dutch pair *-er/-ster*. Like the Dutch case, *-trice* can be viewed as the marked member of the pair, since *-(t)ore* covers also the common gender meaning. With respect to Dutch *-ster*, *-trice* probably has a greater expansion potential, since it is also common for denoting instruments. Moreover, differently from Dutch, both *-(t)ore* and *-trice* can be widely used as adjectives, displaying gender agreement (examples have been already given in (2d); for more details, cf. Lo Duca [2004: 352–356] and Ricca [2004: 442–444]). At any rate, *-trice* is definitely much less frequent than *-(t)ore*, and again this fact has a dramatic impact on the productivity value calculated along Baayen's lines, which for *-trice* is more than five times the value for *-(t)ore*. Much similarly to the case of *-er/-ster*, this result does not look linguistically plausible to us, as argued in section 1. The variable-corpus procedure, on the contrary,



Table 5. Ordering affixes by the number of *h* in the whole corpus

Affixes	$h(N_{max})$	$N_{max}$
<i>-mente</i>	825	317,725
<i>-issimo</i>	643	51,636
<i>-ità/-età</i>	544	356,852
<i>-(z)ione</i>	486	1,043,979
<i>-(t)ore</i>	461	273,706
<i>-bile</i>	409	102,904
<i>-mento</i>	402	257,193
<i>ri-</i>	312	270,066
<i>-trice</i>	224	23,780
<i>-(t)ura</i>	189	63,803
<i>in-</i>	148	146,982
<i>-ezza</i>	70	69,094
<i>-nza</i>	29	208,362
<i>-aggio</i>	29	22,019
<i>-evole</i>	6	19,076

sets the two suffixes about on the same level of productivity, with just a slight preference for *-trice*.

A useful comparison can also be made with  $P^*$ , the hapax-conditioned measure of productivity discussed in section 1, which reduces to  $h(N_{max})$ , that is, the absolute number of hapax legomena found in the corpus. This is done in Table 5, where only the value of  $N_{max}$  is reported for reference besides the relevant ordering parameter  $h(N_{max})$ .

No doubt, this ordering looks linguistically much more significant than the one in Table 4. Following a suggestion by Baayen (pers. comm.), it is possible to evaluate the overall correlation between the ranking in Table 5 and those given in Table 3. We took into consideration four sets of data: the values for  $P(50,000)$ , and the only partially reliable values for  $P(19,000)$ , both as real data and as calculated according to the binomial interpolation. The results are as follows:

- (4) Correlation values between  $h(N_{max})$  and:
- $P(19,000)$  (real data):  $r = 0.9119$  (15 affixes)
  - $P(19,000)$  (binomial interpolation):  $r = 0.9363$  (15 affixes)
  - $P(50,000)$  (real data):  $r = 0.9033$  (12 affixes)
  - $P(50,000)$  (binomial interpolation):  $r = 0.9105$  (12 affixes)

The high correlation values in (4) imply that the rankings obtained by applying Baayen's hapax-conditioned measure of productivity do not differ substantially from ours for the greatest part of the affixes considered. However, they do not exclude the occurrence of some discrepancies. On

the whole, it may be said that  $h(N_{max})$ , as ordering parameter, has the effect of pushing up the most frequent affixes. Consequently, *-mente* surpasses *-issimo* and the splitting between *-(t)ura*, *-mento*, and *-(z)ione* occurs again, but in the opposite order, with respect to Table 4 (even if to a lesser extent). This effect acquires very big proportions at least in one case, namely, the comparison between the two agentive suffixes, the masculine *-(t)ore* and the much less frequent feminine *-trice*. As discussed above, these two suffixes display approximatively the same productivity according to our procedure, while according to the original Baayen's procedure *-trice* results as about five times more productive than *-(t)ore*. The opposite effect is now found if the calculation is based on  $h(N_{max})$ : in this case the value for *-trice* is more than twice lower than the one for *-(t)ore*.<sup>14</sup>

Perhaps more importantly, as already noticed in section 1, by adopting  $h(N_{max})$  as the ordering parameter, there is no more direct connection with the main intuition of (1), namely, the very convincing concept of productivity as the probability of encountering a new type of a given affix after  $N$  tokens of *that affix* have been sampled. Our approach, on the contrary, entirely saves this idea, and at the same time it makes inter-affix comparison feasible and meaningful.

#### 4. Allomorphies and lexicalizations

In section 3 we discussed a number of Italian affixes without declaring explicitly what we really counted as types/tokens of an affix. In the first works by Baayen and collaborators, these aspects had been rather neglected. More recently, Plag et al. (1999) have been more explicit about the questions connected with affix allomorphy and especially with lexicalizations. However, to our knowledge, nobody has attempted to evaluate quantitatively the effect that these questions may have on the measure of productivity. This is also probably due to the fact that most English affixes investigated do not display too serious problems of allomorphy. In this respect, Italian qualifies as a harder test since it displays heavy allomorphies, which have even given rise to different formulations of the same word formation rule.

Besides allomorphy proper, to be dealt with below, there are two further problematic issues deserving discussion, which will be illustrated referring to the action noun suffixes.

The first one concerns what we label "baseless derivatives". This group comprises forms like *detrimento* 'detriment', *ovazione* 'ovation', *massaggio* 'massage', *cesura* 'interruption'. Synchronically, they are simplexes,

since they cannot be related to any extant base. However, their ending and their semantics clearly point to an interpretation in terms of action nouns. There might be good theoretical (and/or psycholinguistic) reasons to include these lexemes in our sample, since they might induce the activation of the respective suffixes, thus influencing their availability in the mental lexicon.<sup>15</sup> On the other hand, in some cases it is rather difficult to discriminate between the examples mentioned above, where an eventive semantics is clearly present, and other instances, where at most the ending can be identified. This is the case of items like *elemento* ‘element’, *dimensione* ‘dimension’, *equipaggio* ‘crew’, *figura* ‘figure’. In view of these difficulties we preferred to exclude all baseless derivatives from the count.

A second issue, somehow specular to the preceding one, regards the heavily lexicalized items, like for instance *sedimento* ‘sediment’ vs. *sedere* ‘sit’, *stazione* ‘station’ vs. *stare* ‘stay’, *temperatura* ‘temperature’ vs. *temperare* ‘temper’, *sentenza* ‘verdict’ vs. *sentire* ‘hear, feel’. In these examples, a morphotactically transparent lexeme is completely unrelated to the verb base from a semantic point of view, at least synchronically. Given their idiosyncratic meaning, it is questionable whether the occurring suffix is really being activated when such words are used. Besides these fully lexicalized items, there are of course several cases where we can speak of regular polysemy in the sense of Apresjan’s (1974) (cf. also Rainer 1993: 136; Nikiforidou 1999; Gaeta 1999, 2004: 316–318): take, for instance, *abitazione* ‘house’ vs. *abitare* ‘inhabit’, *accampamento* ‘(military) camp’ vs. *accamparsi* ‘(to) camp’, *ingranaggio* ‘gear’ vs. *ingranare* ‘put into gear’, *creatura* ‘creature’ vs. *creare* ‘create’. Although these latter items cannot be used at all as action nouns (*abitazione* cannot mean ‘the fact of inhabiting’ and *creatura* cannot mean ‘creation’), the meaning shift from action to place or result has a systematic character both within and across languages; furthermore, it is widely attested in derivatives which keep the action meaning, like *redazione* (both ‘act of compiling’ and ‘editorial office’) and *trasmissione* (both ‘act of broadcasting’ and ‘TV program’). Admittedly, the boundaries in this matter are again rather fuzzy; however, we preferred to keep fully lexicalized items like *sentenza* distinct from all instances of more or less regular polysemic drift, and to exclude only the former from the count.

Coming now to illustrate the questions related to allomorphy, we will discuss in detail the paramount case of the suffix *-(z)ione*. For this suffix even the format of the word formation rule is problematic, since it cannot be safely determined what is the form of the base or of the suffix. Basically, two approaches have been defended in the literature: Thornton (1990–1991) assumes that the base is given by the verbal theme VT (i.e. the root plus the thematic vowel) to which a suffix *-zione* is added,

whereas Scalise (1984: 67) assumes that the base is the past participle of the verb and the suffix takes the form *-ione*:<sup>16</sup>

(5)		Thornton (1990–1991)	Scalise (1984)
a.	<i>fondazione</i> ‘foundation’	[[fonda] <sub>VT</sub> -zione] <sub>N</sub>	[[fondat] <sub>PastPtc</sub> -ione] <sub>N</sub>
	<i>spedizione</i> ‘shipment’	[[spedi] <sub>VT</sub> -zione] <sub>N</sub>	[[spedit] <sub>PastPtc</sub> -ione] <sub>N</sub>
	<i>apparizione</i> ‘apparition’	[[appari] <sub>VT</sub> -zione] <sub>N</sub>	*[[appars] <sub>PastPtc</sub> -ione] <sub>N</sub>
b.	<i>delusione</i> ‘disappointment’	*[[deludi] <sub>VT</sub> -zione] <sub>N</sub>	[[delus] <sub>PastPtc</sub> -ione] <sub>N</sub>
	<i>assunzione</i> ‘employment’	*[[assumi] <sub>VT</sub> -zione] <sub>N</sub>	[[assunt] <sub>PastPtc</sub> -ione] <sub>N</sub>

Notice that Scalise’s approach fails to derive a form like *apparizione*, where the past participle is *apparso*, incorrectly generating a form like *\*apparsione*. On the other hand, Thornton’s approach cannot capture the cases in (5b) containing an irregular past participle, which are correctly generated by Scalise’s hypothesis: for them, Thornton is forced to assume a lexically governed allomorphy. At any rate, both approaches must reckon with a certain amount of lexically governed allomorphy, since the following forms are generated neither by Scalise’s nor by Thornton’s approach:

(6)		Thornton (1990–1991)	Scalise (1984)
	<i>adesione</i> ‘adhesion’	*[[aderi] <sub>VT</sub> -zione] <sub>N</sub>	*[[aderit] <sub>PastPtc</sub> -ione] <sub>N</sub>
	[[ades] <sub>LatPtc</sub> -ione] <sub>N</sub>		
	<i>emissione</i> ‘emission’	*[[emetti] <sub>VT</sub> -zione] <sub>N</sub>	*[[emess] <sub>PastPtc</sub> -ione] <sub>N</sub>
	[[emiss] <sub>LatPtc</sub> -ione] <sub>N</sub>		

The actual forms in (6) go back to the form of the Latinate perfect participle, which must therefore be conceived as a sort of morpheme (cf. Aronoff 1994). We are thus confronted with three main strata of formations.<sup>17</sup> The same holds for some other important Italian suffixes, in particular *-(t)ore* and *-(t)ura* among those dealt with here (for a convincing discussion of the Latinate past participle allomorphy in Italian morphology, see Rainer 2001). Our corpus allows us to estimate quantitatively — probably for the first time with data of comparable size — the relevance of each stratum, both at the tokens and at the type level. The data are reported in Table 6 for *-(z)ione*, *-(t)ura*, *-(t)ore*.

It can be seen from Table 6 that the contribution of formations from irregular past participle or Latinate morphemes is very relevant in tokens, much less so in types, and even less in hapaxes, which shows that the formations which can be related to the verbal theme constitute the only really productive type.

However, from the point of view of *P* calculations, we cannot in principle exclude the other kinds of formations from the count, since they are presumably analyzable as derivatives by native speakers. Therefore, it

Table 6. *Types, tokens and hapaxes for the three strata of -(z)ione, -(t)ura and -(t)ore derivatives*

	<i>N</i>		<i>V</i>		<i>h</i>	
		(%)		(%)		(%)
<i>-(z)ione</i>						
VT + <i>-zione</i>	552,818	53.0	1930	81.7	440	88.5
Irreg. It. PP + <i>-ione</i>	153,530	14.7	176	7.5	20	4.0
Latinate PP + <i>-ione</i>	269,861	25.8	204	8.6	17	3.4
Other allomorphies	67,770	6.5	53	2.2	9	1.8
Total	1,043,979	100.0	2363	100.0	486	100.0
<i>-(t)ura</i>						
VT + <i>-tura</i>	23,198	36.4	512	90.6	181	95.8
Irreg. It. PP + <i>-ura</i>	29,494	46.2	38	6.7	5	2.6
Latinate PP + <i>-ura</i>	5466	8.6	7	1.2	0	0.0
Other allomorphies	5642	8.8	8	1.4	3	1.6
Total	63,800	100.0	565	100.0	189	100.0
<i>-(t)ore</i>						
VT + <i>-tore</i>	160,142	58.5	1307	88.3	436	94.6
Irreg. It. PP + <i>-ore</i>	50,213	18.3	56	3.8	7	1.5
Latinate PP + <i>-ore</i>	50,031	18.3	85	5.7	12	2.6
Other allomorphies	13,320	4.9	32	2.2	6	1.3
Total	273,706	100.0	1480	100.0	461	100.0

is important to test what impact their inclusion has on the measure of productivity. For convenience and brevity, the exemplification will be provided on the basis of the suffix *-(z)ione*; similar results also occur for the other suffixes discussed above.

First, we calculate this impact when Baayen's procedure is applied, as reported in Table 7. We made three counts corresponding to the possible allomorphic types seen above: the first one includes only the formations generated by Thornton's analysis; the second one adds to the former all formations generated on the basis of an irregular past participle following Scalise's analysis; the third one is the maximal choice adding all further allomorphies. In addition, a further count including also the baseless and lexicalized types is provided for reference, although these items

Table 7. *Productivity of -(z)ione including or excluding allomorphies (Baayen's procedure)*

Allomorphic types	$N_{max}$	$V(N_{max})$	$h(N_{max})$	$P(N_{max}) \cdot 10^3$
VT + <i>-zione</i>	552,818	1930	440	0.80
Incl. all PP's + <i>-ione</i>	706,348	2106	460	0.65
Incl. all allomorphies	1,043,979	2363	486	0.47
Baseless and lexicalized types	1,255,281	2483	497	0.40

Table 8. *Productivity of -(z)ione including or excluding allomorphies (our procedure)*

Allomorphic types	Subcorpus size (months)	$P(N_0 = 552,818) \cdot 10^3$
VT + <i>-zione</i>	36	0.80
Incl. all PP's + <i>-ione</i>	28	0.77
Incl. all allomorphies	19	0.73
Baseless and lexicalized types	16	0.70

have not been taken into account in our calculations. It can be seen that proceeding this way, the inclusion of the different allomorphic and lexicalized types substantially modifies the productivity of the suffix, by a factor of two in the worst case.

However, a different picture comes out when adopting our procedure of measuring the productivity rates for equal values of  $N$ , with subcorpora of different sizes, as can be seen in Table 8. In this case, we chose  $N_0 = 552,818$ , which is the highest possible value available for all counts, namely  $N_{max}$  for the most restrictive type VT + *-zione*. To get the corresponding  $P(N_0)$  values for the other set of formations, we had to refer to subcorpora of suitable size, adding a tiny correction by linear interpolation.

The productivity rates now vary to a much more limited extent when the different allomorphic types are included or not. The small variation in favor of the most restrictive choice (the one including derivations from the verbal theme only) is clearly to be expected, given that it comprises nearly all the new coinages. Adding all the allomorphic types, however, does not change the picture radically.

For the sake of clarity, it has to be stressed that the results in Table 8 do not mean at all that the processes involving stem allomorphy display the same order of productivity of the formations built via a transparent rule (namely VT + *-zione*). Rather, the contrary holds. From Table 6 we can see that the items involving stem allomorphies in the whole corpus amount to 491,161 tokens and only 46 hapaxes. Considering these items *alone*, we get  $P(491,161) = 0.094 \cdot 10^{-3}$ . This is an extremely low value if compared to  $P(491,161)$  calculated for transparent formations only: the latter is not retrievable from the tables above, but is easily computable, giving the result  $0.84 \cdot 10^{-3}$ , which is higher by about a factor of nine.<sup>18</sup>

Such a neat contrast is of course most welcome: indeed it is quite an obvious result from a linguistic point of view, meaning that only the VT + *-zione* rule is really productive. What Table 8 shows, however, is that the productivity value for the overall formation process is not very sensitive to the inclusion of allomorphic items, despite the fact that their contribution in tokens is around 50%. It is this relative stability, even in a

particularly complex case like *-(z)ione*, which induced us to adopt the maximal count quite generally for all the data reported in Table 3, given that this is the preferable option on (psycho-)linguistic grounds.

## 5. The inner derivational cycles

In this section we address another methodological question which could undermine — at least in principle — the reliability of quantitative methods to calculate productivity, namely, how to deal coherently with multiply affixed words, like *conventionalize* or *reprintable* (cf. Dal 2003: 20). The standard choice in these cases (defended explicitly in Plag 1999: 29) has always been to select as affix tokens only those words in which the affix has been attached last (i.e. *-ize* and *-able* in the examples given above: in the following we will refer to them as the “outmost derivational cycle,” opposed to all others as “inner cycles”). There are good grounds for this choice, which is also operationally easier, at least for suffixes. However, there are problems both (i) in applying the criterion and (ii) in justifying it from a psycholinguistic point of view.

As for (i), when prefixation and suffixation co-occur, it is not always easy to identify the outmost cycle. To recall a previously mentioned example, should *underdevelopment* be considered a token of *under-* (which implies that it has to be analyzed as [*under*-[*development*]]) or of *-ment*, following the conceivable alternative analysis [[*underdevelop*]-*ment*]? Such uncertainties can be more than sporadic: for instance, since Baayen and Renouf (1996; according to Plag 1999: 108) only included the outmost cycle when computing the data for *un-* and *-ly* in their *Times* corpus, it is legitimate to wonder to which affix they assigned the numerous words like *unwillingly*. An analysis as [[*unwilling*]-*ly*] seems preferable to keep *un-* exclusively as a deadjectival prefix, and we did so in our counts as well for similar formations like *in-util-mente* ‘uselessly’ from *utile* ‘useful’. Nevertheless, the alternative options are conceivable, especially from a semantic perspective.

As for (ii), as Plag (1999: 108) rightly observes, affixes in inner cycles are presumably parsed (at least when the base word is not strongly lexicalized) much like the outmost ones, and therefore they contribute to the morphological competence of the speakers; thus excluding them is little justified psycholinguistically. This holds especially in case of co-occurrence of prefixation and suffixation, when both the outmost and the inner affix are in an equally salient position at the two edges of the word. When comparing, for instance, *repayable* with *unpayable*, the former occurrence of *-able* seems unlikely to be more salient for the speaker than





Table 9. *The relevance of inner cycles for some Italian derivational affixes*

Affixes	$N_{max}$ including outmost cycle only	$N_{max}$ including inner cycles	% of tokens from inner cycles
<i>-mente</i>	317,725	317,725	0
<i>-ezza</i>	69,094	69,236	0.2
<i>-issimo</i>	51,636	51,894	0.5
<i>-mento</i>	257,193	276,856	7.1
<i>in-</i>	146,982	202,744	27.5
<i>-izza-</i>	96,491	149,061	35.3
<i>ri-</i>	270,066	500,912	46.1
<i>-bile</i>	102,904	247,547	58.4

type of *re-* among inner cycle derivations would mean that the corpus contains some occurrences of a word like *reseller* but no occurrence of the verb *resell*. This should not be often the case, since derivatives are normally less frequent than their bases. As for hapaxes, the contribution of inner cycles could even be negative, since the corpus may contain a derived word from a base which is attested only once, which would result in canceling what counted as one hapax in the “outmost-cycle only” procedure.

Similarly to the case of allomorphies in section 4, we compared the values for productivity obtained when inner cycles are included or not, by following Baayen’s procedure and ours. As Baayen (pers. comm.) points out, the inclusion of inner cycles is problematic from a statistical point of view. When inner cycles are included, there is no more one-to-one correspondence between word tokens and affix tokens in the sample, since a word like *unplayable* counts both as a token of *un-* and a token of *-able*. This has the undesirable consequence that the events involved in the two affix distributions are no more fully independent from each other, which impairs the reliability of statistical testing. This is the main reason which has induced Baayen to straightforwardly avoid inner cycles in his work (so did Plag 1999 as well). However, Baayen himself (pers. comm.) recognizes the linguistic interest in investigating this facet of productivity; and with the caution imposed by the statistical caveat discussed above, it may be said that the results give further support to the variable-corpus approach proposed here.

As expected, Baayen’s productivities, reported in Table 10, become much lower when inner cycles are included, due to the substantial increase in tokens, not sufficiently compensated by a possible increase in hapaxes. Notice that this holds not only for *-bile* and *ri-*, whose hapaxes increase minimally, and for *in-*, where they are even reduced, but also for *-izza-*, which interestingly shows on the contrary a relevant increase in

Table 10. *Baayen's productivity with and without inner cycles*

Affixes	$N_{max}$		$V(N_{max})$		$h(N_{max})$		$P(N_{max}) \cdot 10^3$	
	all cycles	outmost cycle only	all cycles	outmost cycle only	all cycles	outmost cycle only	all cycles	outmost cycle only
<i>-bile</i>	247,547	102,904	1203	1117	417	409	1.7	4.0
<i>in-</i>	202,744	146,982	779	767	140	148	0.7	1.0
<i>-izza-</i>	149,061	96,491	882	717	346	280	2.3	2.9
<i>ri-</i>	500,912	270,066	989	935	325	312	0.6	1.2

Table 11. *Productivity with and without inner cycles following the variable-corpus procedure*

Affixes	$N_0$	Subcorpus size for the all-cycle count (months)	$P(N_0) \cdot 10^3$	
			outmost cycle	all cycles
<i>-bile</i>	102,871	15	4.0	2.9
<i>in-</i>	146,982	26	1.0	0.9
<i>-izza-</i>	96,491	23	2.9	3.0
<i>ri-</i>	270,066	19	1.2	1.1

hapaxes, mainly due to the high productivity of the sequence *-izzazione* ‘-ization’.<sup>19</sup> This difference in behavior between *-izza-* and the other affixes is particularly welcome, since it allows our approach to be tested in very dissimilar conditions.

The results for  $P$  according to the variable-corpus procedure are reported in Table 11. As in the case of allomorphies dealt with in section 4, we chose for  $N_0$  the highest possible value, namely  $N_{max}$  for the outmost cycle count. To get the corresponding  $P(N_0)$  values for the all-cycle count, we had to refer to subcorpora of suitable size for each affix, adding the usual correction by linear interpolation.

In this case, the data for the two counts show a substantial alignment for three affixes out of four. For the divergent behavior of *-bile*, the following explanation can be suggested.

The data for the  $V$ s and  $h$ s of *-bile*, reported in Table 10, show that the inner cycles provide a significant increase in types, but — differently from the case of *-izza-* mentioned above — this is not matched by the number of hapaxes, which remains quite stable. In other words, when inner cycles are included, new types of *-bile* are found indeed; however, they are not rare types at all. This state of affairs looks intriguing, as one would not expect in general further derivations to be ordinary words when their bases are not so. The point lies probably in the peculiar interaction

displayed by *-bile* with one of its main further derivations, namely the negative prefixation with *in-*. It is a matter of fact that, on semantic and pragmatic grounds, many common negative adjectives in *-bile* (*introvabile* ‘unfindable’, *instancabile* ‘tireless’, *imperturbabile* ‘imperturbable’) correspond to positive formations (*?trovabile*, *?stancabile*, *?perturbabile*) which are marginal at most. In this case, then, it seems that the outer cycle distribution of types does not really mirror the general profile of *-bile* derivation. Considering only the outmost cycle results into including a sizeable quantity of rare words in *-bile*, like *trovabile*, without including very common formations from the same base, like *introvabile*. In such a situation, the exclusion of the inner cycles has the effect of introducing many “spurious” hapaxes, and therefore considerably enhances the productivity value. Under this perspective, the mismatch between the two counts for *-bile* acquires a precise linguistic motivation. Therefore, far from pointing out a shortcoming of our method, it rather detects a linguistic reality specific for that suffix.

To sum up, in the absence of peculiar semantic restrictions like those occurring for *-bile*, our variable-corpus procedure gives basically identical results regardless of whether inner derivations are included or not, even in cases where they provide a massive contribution in tokens. Therefore, in most cases the ranking obtained will be reliable also when inner derivations are not taken into account. This achievement has pleasant practical consequences, since computing inner cycles is always much time-consuming and not rarely it becomes quite unfeasible: when the affix has not sufficient phonetic/graphic substance, its occurrences within a complex word get confused with identical meaningless sequences much more often than at the word’s edge.

Finally, it could be asked what is the behavior of the alternative Baayen’s measure, namely the full-corpus number of hapaxes,  $h(N_{max})$ , that is,  $P^*$ , with respect to the inner-cycle problem. From Table 10 it can be seen that it performs better than Baayen’s  $P(N_{max})$ , since for the reasons already discussed, inner derivations do not often contribute much to hapax legomena. But when this happens, as is the case for *-izza-*, also  $P^*$  as ordering parameter fatally gets in troubles.

## 6. Conclusion

In this article we proposed a new procedure for calculating productivity values for derivational affixes on the basis of large textual corpora. In particular, this procedure gives stable results with respect to two serious difficulties arising in this sort of approaches, namely (i) the treatment of

allomorphic, not fully segmentable, or strongly lexicalized items and (ii) the impact of inner derivations on the counts. At the same time, our approach keeps intact Baayen's original concept of productivity as a probability measure. This is not to say, of course, that the method can be applied to any affix without restraint. Much investigation on the limits of applicability still needs to be done and probably requires some more theoretical understanding, which we have to leave for further research. However, here we would like to point out some issues still in need of clarification.

A first trivial limitation concerns those affixes whose phonetic shape is too slight to be isolated against bare meaningless sequences. Think of Italian affixes like *-ìa* (*allegro* → *allegria* 'cheerfulness'), *-io* (*frusciare* → *fruscio* 'rustle[N]') or even worse, one-phoneme prefixes like *a-* (*a-normale* 'abnormal') or *s-* (*s-fortuna* 'misfortune', *s-legare* 'untie') which would require manual scrutinizing of all words beginning with the same letter. The extreme case in this sense is given by conversion, a fairly productive process in Italian. Similar inconveniences arise in case of widespread homonymy (for instance, action nouns in *-ata* with identical feminine past participles, relation adjectives in *-are* with infinitives, and so on). For some — but not all — of these cases a (automatically) tagged corpus would help, but apart from their level of reliability (cf. several observations in Plag [1999: 109]), there are not many languages other than English with tagged corpora of suitable size.

A more interesting point regards the limits of applicability of the variable-corpus procedure to compare affixes with extremely divergent token frequencies. In Italian, it is not difficult to find derivational morphemes, which everyone would consider without doubt as qualitatively productive, totalling about one thousand tokens in a 75-million corpus like ours, with a ratio around 1:1,000 with respect to *-(z)ione*, the frequency topscorer. One instance is *-aggine*, forming quality nouns from adjectives with derogatory semantics (as in *sbadataggine* 'carelessness', *cocciutaggine* 'stubbornness'), which totals only 914 tokens in the whole corpus. Perhaps a more significant group is given by several "international" prefixes of learned origin, but today currently used in everyday speech mostly with evaluative function, like *iper-*, *macro-*, *maxi-*, *mega-*, *micro-*, *mini-*, *ultra-*, and so on (for some discussion on these, cf. Iacobini 1999; Gaeta and Ricca 2003b: 108–110, and fn. 14). Trying to compare directly, by our method, these items with the "big" affixes investigated thus far would yield nonsense. A possibility — already partially exploited here — could be having recourse to indirect comparison, by means of "bridging" affixes of intermediate token frequency which could satisfactorily be compared with both the very big and the very small ones; and a

still safer attitude would obviously be just limiting oneself to compare the small affixes with each other. However, it remains unclear whether, with such small frequencies at play, the methodology itself would remain linguistically reliable. Many such items — like the evaluative prefixes mentioned above — display very low token frequencies because they are so recent that practically no entrenched derived words with stable reference exist in the lexicon. Obviously this entails a very high value in productivity, but it is not entirely clear if we are dealing here with exactly the same notion of productivity that can be applied to items already firmly established in the lexicon.

A final unsolved question has a somehow complementary character with the preceding one. What is the upper threshold for  $N$ , above which the proposed calculation of productivity ceases to give linguistically meaningful results? Such a threshold must exist, since any measure of productivity with  $N$  in the denominator should go to zero as the number of affix tokens approaches infinity. It is possible — as Baayen and Lieber (1991: 837) cursorily suggest — that for very big corpora the ratio  $h:V$  could turn out to be a more significant ranking parameter. For extremely large corpora, there is also an exhaustion effect to take into account: the number of types attested in the corpus could approach the total of possible outputs of the derivational rule, which, although in principle unlimited, should be practically finite even for the most productive affixes. The hapax-based statistic measures assume, on the contrary, as a workable approximation that the number of possible formations is infinite (cf. Baayen 1989: 97 for a discussion). Clearly, this exhaustion effect should manifest itself at different thresholds depending on the domains investigated, presumably much earlier with verbal than with nominal derivatives. Needless to say, we have to leave this topic for further research.

*Received 2 December 2002*  
*Revised version received*  
*31 March 2004*

*University of Naples*  
*University of Turin*

## Notes

- \* This work, developed within the FIRB-project “L’italiano nella varietà dei testi,” coordinated by Carla Marengo, has also been partially funded by the Italian Ministry of Education, University and Research (MIUR). The whole article, as well as the computational work, is the result of the close collaboration of both authors; however, for academic purposes, L.G. is responsible for sections 1 and 4 and D.R. for sections 2, 3, and 5. We want to express our gratitude to Marco Tomatis for technical support, and to Harald Baayen for valuable comments and constructive criticism on a previous version

of this article. Correspondence address: Livio Gaeta, Dipartimento di Filologia Moderna “S. Battaglia”, Università di Napoli “Federico II”, Via Porta di Massa 1, I-80133 Napoli, Italy. E-mail: livio.gaeta@unina.it.

1. For the curves  $V(N)$ , Baayen (1989: 120, 2001: 121) states that a statistical significance test can be easily obtained by applying the following approximated formula for the variance:

$$(i) \quad \sigma^2 \approx E^{(2N)}[V] - E^{(N)}[V] - (E[h])^2/N < E^{(2N)}[V] - E^{(N)}[V]$$

where  $E^{(N)}[V]$  is the expectation value of  $V(N)$ . The procedure makes it possible to do so only for  $N$  below  $\frac{1}{2}N_{max}$ , where  $N_{max}$  is the total token number of the given affix in the corpus. The confidence intervals thus calculated are all very narrow, at most amounting to a few percent of the value of  $V(N)$  (therefore, they are not represented in Figure 1). Unfortunately, an equally simple procedure is not available for the productivity curves  $P(N)$  dealt with in the article.

2. Another open question concerns the size of the sample in terms of affix types. As observed by Bauer (2001: 151), “there is not enough information available to be able to give a precise estimate of the size of the sample that would be required to give a reliable statistic.” Thus, it is unclear if it is meaningful to compare the productivity index of an affix occurring in, say, 50 types with the productivity of another one occurring in 1000 types.
3. Plag (1999: 31, fn. 29) has made it clear that the CELEX corpus used by Baayen and Lieber (1991) is not entirely reliable for their purposes, since it does not include a substantial fraction of hapaxes, thus underestimating all productivity values. At any rate, the contrast between *-ee* and *-er*, if not quantitatively sure, should still be valid as an illustration.
4. Baayen himself briefly considers the possibility of evaluating  $P(N)$  at equal  $N$ , but discards it arguing that “what is being studied by comparing affixes for identical sample size is [...] not the competence core of morphological productivity but pragmatic usefulness, a concept that, to our mind, is a component of the pretheoretical notion of morphological productivity” (Baayen 1989: 117). Baayen does not seem to take this alternative procedure into consideration in any of his later works.
5. Some more details on the technical aspects can be found in Gaeta and Ricca (2002).
6. On the unreliability of a full automatic processing in quantitative morphological studies, cf. Evert and Lüdeling (2001).
7. The data reported in Figure 3 are obtained by fitting the data with a power regression curve. Although this choice is not fully adequate theoretically (for a discussion, see Baayen 1989: 105–106), from a practical point of view it gives satisfactory results as long as interpolations and not extrapolations are involved (the coefficients of determination  $R^2$  are around 0.99). For our purposes, we verified that in most instances, even a linear interpolation between the values of  $P(N)$  taken from two contiguous sub-corpora (say, of 25 and 26 months) gives nearly identical results.
8. Taking, for instance, De Mauro (2000) — the largest Italian dictionary available on CD — and selecting only the items of frequent use (the ones which presumably serve as bases for the great majority of further derivations), 4583 nouns, 1381 adjectives, and 1516 verbs are found. Similar proportions hold by selecting the “Base Vocabulary” items of another recent dictionary available on CD, Sabatini and Coletti (1997).
9. Moreover, problems of comparability would certainly arise within the denominal domain itself, given the great differences in size of the relevant subdomains for different denominal derivations. Compare, for instance, the suffixes forming all-purpose relational adjectives (e.g. *-ale* in *nazionale* ‘national’, *rettorale* ‘chancellor’s’) with some

extremely specialized suffixes such as *-eto* ‘plantation of N’, which is nevertheless productive in its domain (cf. *bananeto* ‘banana field’). Further complications in denominal derivation are to be expected due to its interaction with proper names both as input and output. As for the former, should we put deanthroponimic and/or ethnic derivations together with derivations from common nouns? Notice that several suffixes, like *-esco* or *-ino*, perform both functions. And concerning the output, how to deal with morphologically and semantically transparent family and place names?

10. De Mauro (2000) lists only seven neologisms dated after 1950 and analyzable as deverbal for *-evole*, and 26 for *-nza*. Most of the former are of marginal use, while the latter mainly belong to the scientific terminology.
11. The situation is not exactly the same for *-evole* and *-nza*. For the latter suffix, some hapaxes, like *costringenza* ‘constraint’ or *piagnucolenza* ‘whimpering’, which are not attested in De Mauro (2000) or in any other recent dictionary available on CD, could really be new coinages, reflecting that kind of minimal productivity of jocular and/or analogical nature (often labelled “creativity,” cf. Bauer 2001: 62–71), which can scarcely be ruled out for *any* affix with some relevance in the lexicon, but is completely outweighed by the bulk of smoothly accepted, unnoticed new formations whenever truly productive affixes come into play.
12. Intuitively, one should perhaps expect still a higher value for *ri-*, on a par with the other most productive derivational processes listed in Table 3. A factor limiting its productivity may be the fact that *ri-* is a verbal affix: verbs are on the whole less easy to form than nouns and adjectives, as pointed out by the size of the respective type inventories in any large dictionary.
13. The values for *P* are not given in Baayen and Renouf (1996), but can be found in Plag (1999: 113), who also relates this surprising result to the overwhelming frequency of *-ly* (around 1,000,000 tokens). A similar point is made in Bauer (2001: 153).
14. More generally, there is no logical necessity that the results obtained by  $h(N_{max})$  and by *P* at equal *N* have to display strong correlation. Outside the data discussed in this article, a clear counterexample comes from the data on the low frequency evaluative prefixes *super-*, *iper-*, *mega-*, *ultra-*, and *maxi-* discussed in Gaeta and Ricca (2003b: 108–110). While our procedure assigns similar values to all five prefixes, the count of hapaxes heavily favors the most frequent one, namely *super-*, resulting in a complete lack of correlation between the two measures ( $r = 0.27$ ).
15. For instance, Corbin (1987: 188) labels such items as *mots complexes non construits*, and comments: “Linguistiquement, ce sont des mots non construits, qui ont néanmoins une structure interne.” But she also concludes listing them as lexical nonderived units (Corbin 1987: 463).
16. Notice that to correctly derive the forms *fondazione*, *spedizione*, etc., a further morphological rule of affrication /t/ → /ts/ in front of the suffix *-ione* must be assumed in Scalise’s approach. From a theoretical point of view, it does not seem fully advisable to posit a productive word formation rule involving such a heavy morphological alternation. Although independent evidence for the same rule can be found in Italian morphology (e.g. derivations like *Egitto* → *egiziano* ‘Egyptian’, *inerte* → *inerzia* ‘inertia’, cf. Scalise 1994: 153), these instances do not belong to productive processes, in contrast with the case in point.
17. For the sake of completeness, there are further minor cases of allomorphy, namely: (i) root-based derivations, possibly also requiring the already mentioned affrication rule: *gest-ire* → *gestione* ‘management’, *adott-are* → *adozione* ‘adoption’; (ii) a handful of derivatives from the verbal theme with a further allomorph *-gione*: *impiccare* → *impiccagione* ‘hanging’, *guari-re* → *guarigione* ‘recovery’.

18. As a further check, calculating  $P(100,000)$  for the same set including only allomorphic items in *-ione* gives the value  $0.37 \cdot 10^{-3}$ , which places the productivity of this formation process, taken alone, at the bottom of Table 3, slightly above *-nza*.
19. Action nouns in *-izzazione* can apparently be coined also in absence of a well established corresponding verb in *-izzare*, and the well-known preference of newspaper language for nominal style may further enhance the process. Two among the many instances from our corpus are: *aggressivizzazione* 'aggressiv-ization', *angelizzazione* 'angel-ization', etc.

## References

- Anshen, Frank; and Aronoff, Mark (1988). Producing morphologically complex words. *Linguistics* 26, 641–655.
- Apresjan, Julius D. (1974). Regular polysemy. *Linguistics* 12, 5–32.
- Aronoff, Mark (1994). *Morphology by Itself*. Cambridge, MA: MIT Press.
- Baayen, Harald (1989). A corpus-based approach to morphological productivity. Statistical analysis and psycholinguistic interpretation. Unpublished doctoral dissertation, Vrije Universiteit, Amsterdam.
- (1992). Quantitative aspects of morphological productivity. In *Yearbook of Morphology 1991*, Geert Booij and Jaap van Marle (eds.), 109–149. Dordrecht: Kluwer.
- (1993). On frequency, transparency and productivity. In *Yearbook of Morphology 1992*, Geert Booij and Jaap van Marle (eds.), 181–208. Dordrecht: Kluwer.
- (1994). Productivity in language production. *Language and Cognitive Processes* 9(3), 447–469.
- (2001). *Word-Frequency Distributions*. Dordrecht: Kluwer.
- ; and Lieber, Rochelle (1991). Productivity and English word-formations: a corpus-based study. *Linguistics* 29, 801–843.
- ; and Renouf, Antoinette (1996). Chronicing the Times: Productive lexical innovations in an English newspaper. *Language* 72, 69–96.
- Bauer, Laurie (2001). *Morphological Productivity*. Cambridge: Cambridge University Press.
- Corbin, Danielle (1987). *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Niemeyer.
- Dal, Georgette (2003). Productivité morphologique: définitions et notions connexes. *Langue Française* 140 (special issue: La productivité morphologique en questions et en expérimentations), 3–23.
- De Mauro, Tullio (2000). *Grande Dizionario Italiano dell'Uso*. Turin: UTET.
- Evert, Stefan; and Lüdeling, Anke (2001). Measuring morphological productivity: is automatic preprocessing sufficient? In *Proceedings of the Corpus Linguistics 2001 Conference*, Paul Rayson, Andrew Wilson, Tong McEnery, Andrew Hardie, and Shereen Khoja (eds.), 167–175. Lancaster: UCREL.
- Gaeta, Livio (1999). Polisemia e lessicalizzazione: un approccio naturalista. *Italianische Studien* 20, 7–27.
- (2004). Nomi d'azione. In *La formazione delle parole in italiano*, Maria Grossmann and Franz Rainer (eds.), 314–351. Tübingen: Niemeyer.
- ; and Ricca, Davide (2002). *Corpora* testuali e produttività morfologica: i nomi d'azione italiani in due annate della *Stampa* (1996–1997). In *Parallela IX. Testo-variazione-informatica/Text-Variation-Informatik. Atti del IX Incontro italoaustriaco dei linguisti*, Roland Bauer and Hans Goebel (eds.), 223–249. Wilhelmsfeld: Egert.



- ; and Ricca, Davide (2003a). Frequency and productivity in Italian derivation: a comparison between corpus-based and lexicographical data. *Italian Journal of Linguistics/Rivista di Linguistica* 15(1) (special issue: Morphological productivity), 63–98.
- ; and Ricca, Davide (2003b). Italian prefixes and productivity: a quantitative approach. *Acta Linguistica Hungarica* 50, 93–112.
- Haspelmath, Martin (1996). Word-class-changing inflection and morphological theory. In *Yearbook of Morphology 1995*, Geert Booij and Jaap van Marle (eds.), 43–66. Dordrecht: Kluwer.
- Iacobini, Claudio (1999). I prefissi dell'italiano. In *Fonologia e morfologia dell'italiano e dei dialetti d'Italia. Atti del XXXI Congresso della Società di Linguistica Italiana*, Paola Benincà, Alberto Mioni, and Laura Vanelli (eds.), 369–399. Rome: Bulzoni.
- Lo Duca, Maria Giuseppa (2004). Nomi d'agente. In *La formazione delle parole in italiano*, Maria Grossmann and Franz Rainer (eds.), 191–218. Tübingen: Niemeyer.
- Nikiforidou, Kiki (1999). Nominalizations, metonymy and lexicographic practice. In *Issues in Cognitive Linguistics. 1993 Proceedings of the International Cognitive Linguistics Conference*, Leon de Stadler and Christoph Eyrych (eds.), 141–163. Berlin and New York: Mouton de Gruyter.
- Plag, Ingo (1999). *Morphological Productivity*. Berlin and New York: Mouton de Gruyter.
- ; Dalton-Puffer, Christiane; and Baayen, Harald (1999). Morphological productivity across speech and writing. *English Language and Linguistics* 3, 209–228.
- Rainer, Franz (1987). Produktivitätsbegriffe in der Wortbildungslehre. In *Grammatik und Wortbildung romanischer Sprachen*, Wolf Dietrich (ed.), 187–202. Tübingen: Narr.
- (1989). *I nomi di qualità nell'italiano contemporaneo*. Vienna: Braunnmüller.
- (1993). *Spanische Wortbildungslehre*. Tübingen: Niemeyer.
- (2001). Compositionality and paradigmatically determined allomorphy in Italian word-formation. In *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler Presented on the Occasion of His 60th Birthday*, Chris Schaner-Wolles, John R. Rennison, and Friedrich Neubarth (eds.), 383–392. Turin: Rosenberg & Sellier.
- (2003). Studying restrictions on patterns of word-formation by means of the Internet. *Italian Journal of Linguistics/Rivista di Linguistica* 15(1) (special issue: Morphological productivity), 131–140.
- Ricca, Davide (1998). La morfologia avverbale tra flessione e derivazione. In *Ars Linguistica. Studi offerti da colleghi ed allievi a Paolo Ramat in occasione del suo 60° compleanno*, Giuliano Bernini, Pierluigi Cuzzolin, and Piera Molinelli (eds.), 447–466. Rome: Bulzoni.
- (2004). Aggettivi deverbali. In *La formazione delle parole in italiano*, Maria Grossmann and Franz Rainer (eds.), 419–444. Tübingen: Niemeyer.
- Sabatini, Francesco; and Coletti, Vincenzo (1997). *DISC—Dizionario Italiano Sabatini Coletti. Edizione in CD-Rom*. Florence: Giusti.
- Scalise, Sergio (1984). *Generative Morphology*. Dordrecht: Foris.
- (1994). *Morfologia*. Bologna: Il Mulino.
- Thornton, Anna M. (1990–1991). Sui deverbali italiani in *-mento* e *-zione* (I–II). *Archivio Glottologico Italiano* 75, 169–207 and 76, 79–102.
- van Marle, Jaap (1992). The relationship between morphological productivity and frequency: a comment on Baayen's performance-oriented conception of morphological productivity. In *Yearbook of Morphology 1991*, Geert Booij and Jaap van Marle (eds.), 151–163. Dordrecht: Kluwer.

