Carmen Scherer, Mainz

Diachronic word formation: -er nominals in German

1. Introduction

The formation of -er nominals such as Lehrer and Töpfer in German and teacher and potter in English is a very productive word formation process in both languages (Fleischer/Barz 1995, Ryder 1991). Because the -er nominalization process shows a broad range of referents realized by the output and no stringent constraints on its input, it has often been discussed under different linguistic angles, for example complexity (Bauer 1979), polysemy (Booij 1986), argument structure (Rappaport Hovav/Levin 1992, Heyvaert 1998, Szigeti 2002) and incorporation (Rivet 1999). In this paper I will focus on the influence the word category of the base exerts in word formation change. My aim is to demonstrate that the productivity of the word formation pattern changes diachronically and that productivity varies dependent on the word category of the base. Finally I will argue that both the deverbal as well as the denominal -er subclass gain productivity, even though prior studies show increasing proportions of deverbal -er nominals and decreasing proportions of denominal ones.

In the next chapter an outline of -er nominals in present-day English and German will be presented. Chapter 3 provides a brief sketch of the history of the word formation pattern and discusses data from former studies. Subsequently the tendencies detected in chapter 3 will be evaluated in contrast to the data of the Mainz Newspaper Corpus in terms of type frequency (chapter 4) and productivity in the narrow sense (chapter 5). Chapter 6, finally, summarizes the findings on diachronic changes of -er nominals in German.

2. Outline of -er nominals in present-day German

-er nominals in present-day English show a wide scope of semantic, morphological and syntactic features. They may denote human beings (teacher), objects (tanker) and even immaterial things (reminder). They may be formed out of verbs (lover), nouns (gardener), geographic names (Londoner) and other word categories (downer, foreigner) which themselves may be morphologically simple as in washer or complex as in left-hander, cf. Marchand (1969), Ryder (1999).

Turning to -er nominals in German, by and large the same range of semantic, morphological and syntactic variation of input and output can be perceived, except for adjectival bases (Barfüsser) which are very rare in German. We find -er nominals denoting persons such as Lehrer 'teacher', objects such as Wasserkocher 'electric kettle (water cooker)' and abstract nouns such as Seufzer 'sigh (sigher)'. There are nominals based on verbs (Drucker 'printer'), nouns (Fleischer 'butcher (meater)'), geographic names (Norweger 'person from Norway
(Norwayer)’ and other bases (Zehner 'coin/bill with the value of ten (tenner))\(^1\). As in English, the bases may be simple as in Lehrer 'teacher' which is based on the simplex verb lehren 'to teach' or complex as in Handballer 'handball player' which is derived from the compound Handball 'handball'. Further examples are provided in Table 1.

Table 1. Classification of -er nominals in present-day German

<table>
<thead>
<tr>
<th>-er nominals</th>
<th>verbal base</th>
<th>nominal base</th>
<th>geographic name</th>
<th>other bases</th>
</tr>
</thead>
<tbody>
<tr>
<td>denoting persons</td>
<td>Lehrer 'teacher', Hutmacher 'hatter (hat maker)'</td>
<td>Gärtner 'gardener', Handballer 'handball player (handballer)'</td>
<td>Norweger 'person from Norway (Norwayer)'</td>
<td>Gegner 'opponent (againster)', Barfüßer 'barefooted monk (barefooter)'</td>
</tr>
<tr>
<td>denoting objects</td>
<td>Drucker 'printer', Wasserkocher 'electric kettle (water cooker)'</td>
<td>Benziner 'car that runs on gasoline (gasoliner)', Zweisitzer 'two-seater'</td>
<td>Champagner 'Champagne (Champagner)'</td>
<td>Zehner 'coin/bill with the value of ten (tenner)'</td>
</tr>
<tr>
<td>denoting abstract things</td>
<td>Seufzer 'sigh (sigher)'</td>
<td>Einakter 'play with only one act (one acter)'</td>
<td>Einser 'best grade in school (oner)'</td>
<td></td>
</tr>
</tbody>
</table>

According to Wellmann (1975), in present-day German almost three out of four -er nominals (72.4%) are based on verbs and a good 20% are denominal. -Er nominals formed out of geographic names (Norweger) and numerals (Zehner) are rare (<5%) but not uncommon, whereas -er nominals based on other word categories such as adjectives (Barfüsser) and prepositions (Gegner) are isolated cases.

3. -Er nominals from a historical point of view

The formation of -er nominals is a very old word formation option in the Germanic languages. The first -er nominals can be found as early as Gothic. More than 600 -er nominals are attested for Old High German (OHG) (Weinreich 1971) and according to Sütterlin (1887) about 2,200 -er nominals can be found in Middle High German (MHG). According to the data given in Wellmann (1975), today the German -er suffix is one of the three most productive suffixes (-ung, -heit/-keit, -er) in the formation of nouns, cf. Scherer (2003). Even though the -er suffix has a long tradition in the word formation morphology of several Germanic languages (German, English, Dutch) it cannot be traced back to a Germanic origin, but was rather borrowed from Latin in the second or third century AC (Brinkmann 1965).

\(^1\) Please note that deonomastic and denominal formations differ in that common nouns and proper nouns (names) are usually treated differently in word formation, i.e. most word formation processes operate either on common nouns or on names. In German, for example, -lich and -ig both form adjectives and take only common nouns as a base but not proper nouns.
In Latin, the suffix -arius first formed adjectives out of nominal bases but was soon used to create person-denoting nouns as well. In analogy to Latin -arius nominals, Gothic -areis nominals have long been supposed to be based solely on nouns (Sütterlin 1887, Maurer 1973), even though only verbal bases are attested for some of them (Wilmanns 1899). As far as OHG -ari nominals are concerned, some authors (Braune 1987, Henzen 1965) assumed that the majority of them were based on nouns as well. This assertion has been disproved by empirical data and has since been revised (Splett 2000). The studies of Weinreich (1971) and Maurer (1973) on agentive nouns in OHG provide empirical data for OHG -ari nominals. Their results show that -ari nominals in OHG were mostly based on verbs (56.9%-78.3%). More recent corpus studies reveal even higher proportions of deverbal -er nominals for later periods of German, namely Early New High German (ENHG) and NHG (66.7%-80.1%), cf. Table 2.

Table 2. Word category of the base of -er nominals: proportion of types

<table>
<thead>
<tr>
<th></th>
<th>OHG 1362-1490</th>
<th>OHG 1490-1528</th>
<th>OHG 1521-1572</th>
<th>OHG 18-19th c.</th>
<th>OHG 20th c.</th>
</tr>
</thead>
<tbody>
<tr>
<td>types</td>
<td>473</td>
<td>263</td>
<td>542</td>
<td>95</td>
<td>1,190</td>
</tr>
<tr>
<td>verb</td>
<td>56.9%</td>
<td>78.3%</td>
<td>80.1%</td>
<td>68.4%</td>
<td>77.4%</td>
</tr>
<tr>
<td>noun</td>
<td>43.1%</td>
<td>21.7%</td>
<td>17.5%</td>
<td>28.4%</td>
<td>16.0%</td>
</tr>
<tr>
<td>geogr. name</td>
<td>0.0%</td>
<td>0.0%</td>
<td>2.4%</td>
<td>2.1%</td>
<td>4.7%</td>
</tr>
<tr>
<td>others</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>1.1%</td>
<td>1.9%</td>
</tr>
</tbody>
</table>

Corresponding to the higher proportion of deverbal formations, the percentage of denominal -er nominals in ENHG and NHG is lower than in OHG – even if proper names were included. Even though the data given in Table 2 show no continuous development, a tendency towards higher proportions of deverbal and lower proportions of denominal -er nominals in ENHG and NHG may be perceived (Meibauer et al. to appear).

However, it has to be pointed out that the proportions given in Table 2 are based on a total type frequency varying between 95 (Müller 1993) and 1,467 (Wellmann 1975) -er nominals. Furthermore, for three of the studies (Maurer 1973, Weinreich 1971, Wellmann 1975) the population of word forms analyzed is unknown and only Brendel et al. (1997), Müller (1993) and Stricker (2000) provide information about the token frequency. The data provided in Table 2 therefore do not allow for more sophisticated analyses. Based on the data collected from former studies it is neither possible to judge whether the total number of -er nominals has increased or decreased during the last centuries, nor can one ascertain whether the -er nominalization pattern has become more productive. Finally, no claim can be made as to whether

---

All percentages given are based on counts by Scherer (2003).
the -er suffix has become more productive with one of the base categories than with the others.

To find answers to these questions a systematic analysis of a homogeneous diachronic corpus is necessary. The Mainz Newspaper Corpus (MNC) is such a corpus. It consists of nine subcorpora comprising 98,000 to 150,000 word forms each and covers the NHG period from the 17th century on. The earliest data were collected for the year 1609 when the first newspapers worldwide appeared, namely the Wolfenbüttel Aviso and the Straßburg Relation. The second subcorpus is dated 1650 with the other measuring points following every 50 years, the last one in 2000. Table 3 provides information about the size of the subcorpora and the number of -er nominals each of them contained. A detailed list of all newspapers included in the MNC can be found in chapter 7.

Table 3. Mainz Newspaper Corpus (MNC): overview

<table>
<thead>
<tr>
<th></th>
<th>1609</th>
<th>1650</th>
<th>1700</th>
<th>1750</th>
<th>1800</th>
<th>1850</th>
<th>1900</th>
<th>1950</th>
<th>2000</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>word forms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(in thousands)</td>
<td>98.9</td>
<td>98.3</td>
<td>97.7</td>
<td>102.8</td>
<td>101.1</td>
<td>108.6</td>
<td>149.5</td>
<td>136.5</td>
<td>137.4</td>
<td>1.031.1</td>
</tr>
<tr>
<td>-er types</td>
<td>159</td>
<td>96</td>
<td>172</td>
<td>231</td>
<td>238</td>
<td>339</td>
<td>532</td>
<td>581</td>
<td>767</td>
<td>2,084</td>
</tr>
<tr>
<td>-er tokens</td>
<td>693</td>
<td>352</td>
<td>630</td>
<td>751</td>
<td>646</td>
<td>911</td>
<td>1,553</td>
<td>1,581</td>
<td>1,890</td>
<td>9,007</td>
</tr>
</tbody>
</table>

Because corpora typically vary regarding their size, it is necessary to normalize corpus data in order to compare data of different origin (Cowie/Dalton-Puffer 2002). As a consequence, all MNC data given in the following have been normalized on the basis of 100,000 word forms.

4. Diachronic changes in the word category of the base

According to Stricker (2000) and Wellmann (1975) an overwhelming majority of -er nominals in NHG have verbal bases, cf. Table 2. Their findings are confirmed by the data of the MNC even though the proportion of deverbal -er nominals is slightly lower (65.4%), whereas denominal (21.8%) and deonomastic formations (10.7%) figure more prominently in the MNC, cf. Table 4. -Er nominals with bases which are neither verbal nor nominal nor onomastic will be neglected in the following.
Table 4. -Er nominals: proportion of types per base in the MNC

<table>
<thead>
<tr>
<th></th>
<th>1609</th>
<th>1650</th>
<th>1700</th>
<th>1750</th>
<th>1800</th>
<th>1850</th>
<th>1900</th>
<th>1950</th>
<th>2000</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>verbal base</td>
<td>42.8%</td>
<td>43.8%</td>
<td>44.8%</td>
<td>62.8%</td>
<td>57.1%</td>
<td>69.9%</td>
<td>66.7%</td>
<td>66.8%</td>
<td>69.6%</td>
<td>65.4%</td>
</tr>
<tr>
<td>nominal base</td>
<td>23.3%</td>
<td>26.0%</td>
<td>30.2%</td>
<td>21.6%</td>
<td>23.1%</td>
<td>20.1%</td>
<td>24.4%</td>
<td>23.1%</td>
<td>21.6%</td>
<td>21.8%</td>
</tr>
<tr>
<td>geogr. name</td>
<td>28.9%</td>
<td>27.1%</td>
<td>20.9%</td>
<td>10.8%</td>
<td>18.1%</td>
<td>8.3%</td>
<td>7.0%</td>
<td>9.3%</td>
<td>7.3%</td>
<td>10.7%</td>
</tr>
<tr>
<td>other bases</td>
<td>5.0%</td>
<td>3.1%</td>
<td>4.1%</td>
<td>4.8%</td>
<td>1.7%</td>
<td>1.8%</td>
<td>1.9%</td>
<td>0.9%</td>
<td>1.4%</td>
<td>2.1%</td>
</tr>
</tbody>
</table>

Based on the MNC data illustrated in Table 4, a number of diachronic developments of the subclasses can be detected. The proportion of deverbal -er nominals rises continually from 42.8% in the first to 69.6% in the last subcorpus. Yet the percentage of denominal -er nominals in the same time proves to be comparatively stable. With the exception of 1700, the proportion of nominal bases shows little variation (20.1%-26.0%). On the other hand, the proportion of deonomastic nominals decreases dramatically. In the 17th century more -er nominals were formed out of geographic names (27.1%-28.9%) than out of nouns (23.3%-26.0%), but the number of deonomastic nominals has fallen considerably since. From 1850 on they constitute less than 10% of all -er nominals in the corpus. This important decrease of deonomastic -er nominals in the 18th century may be connected to the profound historical changes that Europe and especially Germany underwent after the Thirty Years' War. Throughout the whole Middle Ages and right through the 17th century, the social and geographical provenance of a person has been particularly important which matches with high proportions of deonomastic -er nominals. With the rise of mercantilism and enlightenment in the 18th century, the individual became more important than his or her origin. Therefore, the need for deonomastic person denotations diminished. As person-denoting nouns constitute the large majority of -er nominals (Scherer 2003) the decline of deonomastic -er nominals may therefore be a by-product of the historical change from medieval to modern societies in central Europe.

However, if common nouns and proper nouns were treated as subclasses of nouns, as is usually done in word formation, the picture would be rather different. The proportion of denominal -er nominals would decrease even though this would mainly be due to onomastic bases. The distinction between nominal bases (common nouns) and onomastic bases (proper nouns) is therefore appropriate and may provide helpful insights in deonomastic word formation.

Overall, obvious differences in the diachronic development of the deverbal, denominal and deonomastic subclasses of the -er word formation pattern become manifest in the MNC. The

---

A linear regression was conducted for all MNC data to test the statistical significance of the data. Due to this test the proportional increase of the deverbal and the decrease of the deonomastic subclass proved highly significant (deverbal: p=0.001, R²=0.811; deonomastic: p=0.001, R²=0.798). No significant tendency could be detected for denominal -er nominals (p=0.231, R²=0.197).
data indicate that, diachronically speaking, the formation of -er nominals out of verbs has become more productive during the NHG period. They further suggest that there has been no significant change in the productivity of the denominal subclass and a remarkable loss of productivity in the formation of -er nominals based on geographic names.

However, it has to be kept in mind that diachronic comparisons of proportions are only of limited explanatory power. They do not allow a reliable evaluation of diachronic changes of a word formation pattern or its subclasses as percentages do not make reference to the population of types and/or tokens analyzed. This point is made more explicit by the data presented in Table 5. It shows that the total number of -er nominals more than tripled during the last four centuries. This means that percentages are based on a population of 161 types in 1609 but that 400 years later they refer to a population of 558 which is more than three times as many types as in 1609. As an example, the 23.3%-proportion of denominal -er nominals corresponds to 37 types in 1609 whereas in 1950 the almost identical proportion of 23.1% is made up by 98 denominal types.

Table 5. -Er nominals: types in the MNC per 100,000 word forms

<table>
<thead>
<tr>
<th></th>
<th>1609</th>
<th>1650</th>
<th>1700</th>
<th>1750</th>
<th>1800</th>
<th>1850</th>
<th>1900</th>
<th>1950</th>
<th>2000</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>types total</td>
<td>161</td>
<td>98</td>
<td>176</td>
<td>225</td>
<td>236</td>
<td>312</td>
<td>356</td>
<td>426</td>
<td>558</td>
<td>1,819</td>
</tr>
<tr>
<td>verbal base</td>
<td>69</td>
<td>43</td>
<td>79</td>
<td>141</td>
<td>135</td>
<td>218</td>
<td>237</td>
<td>284</td>
<td>389</td>
<td>1,190</td>
</tr>
<tr>
<td>nominal base</td>
<td>37</td>
<td>25</td>
<td>53</td>
<td>49</td>
<td>54</td>
<td>63</td>
<td>87</td>
<td>98</td>
<td>121</td>
<td>396</td>
</tr>
<tr>
<td>geogr. name</td>
<td>46</td>
<td>26</td>
<td>37</td>
<td>24</td>
<td>43</td>
<td>26</td>
<td>25</td>
<td>40</td>
<td>41</td>
<td>196</td>
</tr>
<tr>
<td>other bases</td>
<td>8</td>
<td>3</td>
<td>7</td>
<td>11</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>8</td>
<td>38</td>
</tr>
</tbody>
</table>

When evaluating word formation from a diachronic point of view it is therefore essential not to rely on the development of proportional allotments only but to consider other figures, i.e. type and token frequency and productivity indices, as well.

Table 6 shows the diachronic development of type frequency. To allow the comparison between the word formation pattern in general and its different subclasses, the development of types has been modeled on the basis of the first measuring point. This means that the type frequency found for the whole pattern and its subclasses in the first subcorpus equals 100%, with the other data referring to this number. For example, in case of the whole pattern, the number of types in 1609, i.e. 161 types, equals 100%, and the 558 -er nominals attested in 2000 correspond to 347%, i.e. in the last subcorpus we find three and a half times as many -er nominals as in the first one.

---

4 The data proved statistically significant on a very high level for the -er nominalization in general (p=0.000, $R^2=0.910$), the deverbal (p=0.000, $R^2=0.928$) and the denominal subclass (p=0.00, $R^2=0.893$), but not for the deonomastic subclass (p=0.958, $R^2=0.000$).
Table 6 substantiates the different diachronic patterns of type frequency of the -er nominalization in general and of its word-category based subclasses. Two of the subclasses, the deverbal and the denominal one, and the general pattern exhibit a significant increase of types. Strikingly, the frequency of deverbal -er nominals increases faster than the total number of -er nominals. In 2000 the number of deverbal -er nominals is about five and a half times as high as in 1609, whereas the total type frequency and the number of denominal types achieve 347% and 323%, respectively. In contrast, the number of deonomastic -er nominals is highest in the first subcorpus and no clear diachronic tendency can be detected throughout the period investigated.

As in case of the decreasing proportion of deonomastic nominals, the overall rise of -er nominals in German between 1609 and 2000 may be associated with the changeover between baroque and modern times. Due to fundamental changes in European politics, economies and societies, more designations for human beings and objects were needed and could easily be provided by the polysemous -er word formation pattern. On the other hand, it is not easy to explain why deverbal -er nominals increased by such a greater extent than denominal ones. A possible starting point is to analyze formal restrictions according to the referents realized by the -er nominals. A survey of the MNC data and of the noun forming word formation patterns as listed by Motsch (1999), for example, clearly indicates that the formation of object denotations out of nouns is heavily constricted whereas the formation of deverbal object nouns is
very productive. Further research should therefore act on the assumption that restrictions on the bases are influenced by semantic properties of the output.

The mismatch between the rise of the total number of -er nominals and the increase of the deverbal subclass results in the diachronic growth of the proportion occupied by deverbal -er nominals, cf. Table 4. Accordingly, the comparatively similar development of the -er nominalization pattern and the denominal subclass correlates with the rather stable proportion of denominal -er nominals in the MNC. Finally, the heterogeneous but stagnating development of the deonomastic subclass is reflected by the fluctuating but clear decline of the proportion of deonomastic -er nominals.

This means that if the type frequency of a subclass increases faster than the total number of types derived by a certain word formation pattern, the proportion of this subclass will increase (deverbal -er nominals). If, on the other hand, the number of types of a certain subclass rises very much the same way the whole pattern does, the proportion of the subclass will be quite stable (denominal -er nominals). If, finally, the type frequency of a subclass remains stable or even diminishes with the total type frequency rising at the same time, the proportion of this subclass will decrease (deonomastic -er nominals). Therefore, changes in the proportion of types of a certain subclass can only provide useful information when compared to the development of the total number of types analyzed.

5. The interdependency of productivity and the word category of the base

Productivity is one of the most difficult notions in morphology (Aronoff 1976, Mayerthaler 1981). Even though within the last 15 years important progress in this field has been made (Baayen 1989, Plag 1999, Bauer 2001) many questions still remain open. The basic definition of productivity to which most linguists seem to adhere claims that productivity "deals with the number of new words that can be coined using a particular morphological process" (Bauer 2001:211).

The productivity of a word formation process may vary diachronically (Bauer 2001). It may also vary according to certain features of its input or output (Kastovsky 1986, Baayen 1992, Bauer 2001). In chapter 4, for example, we have seen that the total number of -er nominals rose continually during the last four centuries. However, when comparing nominals with different base categories it turned out that the extent to which a subclass increased (or decreased) was closely linked to the word category of its base. In this chapter we will see that the same holds true for productivity.

A quantitative notion of productivity provided, productivity can be measured. Productivity has often been equated with the number of types or tokens derived by a given word formation pattern. In fact, the discussion on diachronic tendencies in chapter 4 was mainly based on type frequency. However, measuring productivity via type or token frequency ignores the core idea
of productivity, namely the generation of new coinages. A large number of types and tokens in a corpus may simply be due to the fact that a pattern was once used productively. But because productivity may vary diachronically, former productivity does not ensure that a pattern will still be used later on. This is the case e.g. for \(-e\) affixation in present-day German ('mattock', 'donation') which ranks as number 7 in Wellmann's (1975) corpus in type frequency (out of a total 40 nominal suffixes) but is considered to be only slightly productive in some contexts and completely unproductive in others (Motsch 1999). Therefore, as the number of types or tokens attested does not indicate to what extent a word formation process is used to produce new coinages, neither frequency should be used exclusively when judging the productivity of a word formation process (Plag 1999).

In their paper on productivity in English derivation Baayen/Lieber (1991) discuss several different methods to quantify productivity. In the following I will concentrate on one of those, namely the productivity index \(P\), 'productivity in the narrow sense'. The calculation of \(P\) is based on the close link between the number of new coinages and the number of hapax legomena in a corpus as exemplified e.g. by Baayen/Lieber (1991) and Baayen (1992). \(P\) is defined as the proportion of hapax legomena with a given affix \(n_1\) in a corpus within the total number of tokens \(N\) with that same affix \((P = n_1/N)\). Thus, \(P\) is supposed to measure "the probability of coming across new, unobserved types" (Baayen/Lieber 1991:809) which are supposed to be indicators for the number of new coinages in a corpus.

Productivity in the narrow sense has been calculated for the \(-er\) nominalization pattern and each of its subclasses for each subcorpus separately. The results are shown in Tables 7 and 8. For further productivity data in the MNC cf. Scherer (2003).

### Table 7. \(-Er\) nominals: productivity in the narrow sense \((P)\) in the MNC

<table>
<thead>
<tr>
<th>Year</th>
<th>1609</th>
<th>1650</th>
<th>1700</th>
<th>1750</th>
<th>1800</th>
<th>1850</th>
<th>1900</th>
<th>1950</th>
<th>2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>types total</td>
<td>0.076</td>
<td>0.065</td>
<td>0.078</td>
<td>0.079</td>
<td>0.121</td>
<td>0.123</td>
<td>0.138</td>
<td>0.153</td>
<td>0.223</td>
</tr>
<tr>
<td>verbal base</td>
<td>0.128</td>
<td>0.085</td>
<td>0.142</td>
<td>0.089</td>
<td>0.130</td>
<td>0.151</td>
<td>0.134</td>
<td>0.154</td>
<td>0.231</td>
</tr>
<tr>
<td>nominal base</td>
<td>0.017</td>
<td>0.038</td>
<td>0.039</td>
<td>0.056</td>
<td>0.107</td>
<td>0.057</td>
<td>0.149</td>
<td>0.140</td>
<td>0.209</td>
</tr>
<tr>
<td>geogr. name</td>
<td>0.093</td>
<td>0.072</td>
<td>0.075</td>
<td>0.078</td>
<td>0.119</td>
<td>0.231</td>
<td>0.175</td>
<td>0.193</td>
<td>0.182</td>
</tr>
</tbody>
</table>

6 The diachronic change in the productivity \(P\) proved statistically significant on a high level for the \(-er\) nominalization in general \((p=0.000, R^2=0.851)\) and any of the subclasses \(\text{deverbal: } p=0.029, R^2=0.518, \text{denominal: } p=0.000, R^2=0.841, \text{deonomastic: } p=0.008, R^2=0.661)\).
Independent of the word category of the base, the productivity of the -er nominalization pattern and its subclasses rises significantly. Still, some differences in the diachronic development of the whole pattern and the word category-based subclasses are evident. Regarding the pattern in general, $P$ has been rather constant in the 17th and 18th centuries (0.065-0.079) but increased considerably from 1800 on. Finally, in 2000 the -er nominalization process is almost three times as productive as it used to be in 1609, cf. Table 7. Compared with the general pattern, the deverbal and deonomastic formation processes have been more productive throughout the entire period investigated, with only one exception each. By contrast, the deonominal pattern proved less productive than the -er nominalization in general in any of the periods investigated with the exception of 1900. However, whereas the productivity $P$ of the deverbal pattern constantly oscillated between 0.085 and 0.154 until 1950, it finally climbed up to 0.231 in 2000, i.e. about twice the value it had in 1609, cf. Table 7. This rise in productivity is rather small compared to the rise of deverbal types which, in 2000, is five and a half times higher than in 1609, cf. Tables 5-6. This mismatch is not due to a smaller number of new coinages, in fact there are eight times as many hapax legomena in 2000 as in 1609. It is rather the rise of token frequency (2000=443%) that exceeds the rise of the total population of tokens (2000=196%) which is responsible for the comparatively small increase of $P$ in the deverbal subclass. It is therefore evident that the diachronic study of type frequency and productivity in the narrow sense may lead to different results. Therefore, when evaluating pro-
ductivity, both, type frequency and productivity in the narrow sense, should be considered.

Compared with the deverbal subclass and the \(-er\) nominalization in general, the denominal subclass gains strikingly in productivity \(P\) over the last four centuries. Rather low values for \(P\) between 1609 and 1850 (0.017-0.057) are followed by considerable rises in productivity in 1800 and 1900 with a constant increase in productivity from 1900 on. As a result, in the last period investigated, \(P\) (0.209) is five to ten times higher than for the first four periods, cf. Table 7. This is mainly due to important rises in the number of hapax legomena with the overall amount of tokens in the subcorpora remaining rather stable. It should also be stressed that the increase in frequency and productivity of the subclasses is closely linked to the rising complexity of the \(-er\) nominals, cf. Scherer (2003).

As for the third subclass investigated, the productivity of deonomastic \(-er\) nominals rises significantly during the last four centuries. \(P\) shows only little variance in the first four subcorpora (0.72-0.93), it then skips to a maximum in 1850 (0.231) before it finally oscillates on a level twice as high as in the 17th and 18th century (0.175-0.193). This is a rather unexpected result given the fact that no significant increase in the frequency of types, tokens and hapax legomena could be detected. But the fact that the token frequency in the subcorpora decreases faster then the number of types gives rise to a larger number of hapax legomena and as a consequence to higher values of \(P\). As a result, productivity in the narrow sense grows significantly, even though the frequency of types and tokens is decreasing. This result proves, once again, that it is essential to compare different measures when talking about productivity.

6. Conclusion

Word formation processes are subject to diachronic change that can be measured in terms of productivity. As for the extent and manner of change, they may vary according to different subclasses of one and the same word formation pattern.

In this paper diachronic change has been discussed on the basis of the \(-er\) nominalization in German. The study of the Mainz Newspaper corpus showed that in order to judge diachronic changes in word formation, a more detailed empirical analysis has to be applied than is the norm. Neither the survey of proportions of subclasses in different (sub)corpora nor the examination of type (or token) frequency allow for reliable results. For e.g. the proportion of denominal \(-er\) nominals in NHG as manifested in the MNC data remained rather stable. Never-

7 Actually, both indices, type frequency \(V\) and productivity in the narrow sense \(P\), have been combined in global productivity \(P^*\), another productivity index discussed in Baayen/Lieber (1991) and Baayen (1992, 1993).

8 Apparently speakers tend to judge complex deverbal \(-er\) nominals such as Klavierspieler 'piano player' as derivation, whereas for complex denominal \(-er\) nominals such as Innenpolitiker 'politician dealing with domestic policy (domestic policy+er)' compound-reading seems to be preferred. Thus, to avoid disparate treatment of deverbal and denominal \(-er\) nominals in the MNC no distinction has been made between derived and compounded \(-er\) nominals following Bauer (1979).
theless, the number of -er nominals rose diachronically from 1609 to 2000, as did productivity in the narrow sense. Overall, the MNC indicates significant gains in productivity as substantiated in type frequency and productivity $P$ for the word formation pattern investigated. The MNC data also confirmed the claim that productivity may vary according to different features of the nominals derived by one single word formation pattern. Of the three subclasses distinguished (deverbal, denominal, deonomastic) two, the deverbal and the denominal one, showed a clear diachronic rise of productivity, but differ in the extent of terms of type frequency and productivity in the narrow sense. For the third subclass, the deonomastic -er nominals, no clear diachronic tendency can be ascertained. Even though the frequency of types, tokens and hapax legomena exhibits no significant tendency, the productivity $P$ rises in a statistically significant way. The overall rise of -er nominals was possibly initiated by the fundamental historical changes that took place between the 17th and the 20th centuries, whereas the development of the individual subclasses appears to be linked to other features of the input and output of the word formation pattern.

Two main results have to be emphasized, the first one being theoretical, the second one methodological. First, productivity is not an invariant attribute of word formation processes. It varies diachronically as well as according to features of the input and output of the process concerned. Second, because morphological productivity is a highly complex phenomenon, it is essential to consider different measures of productivity when evaluating the productivity of a certain word formation pattern.

7. Corpus

1609 = Aviso Relation oder Zeitung (1609). Wolfenbüttel 1609-1624.
   Including: Wochentliche Donnerstags Zeitung, Ordinari Diengstags Zeitung and APPENDIX Der Wochentlichen Zeitung.
   Including: Mercurii Relation, Oder Wochentliche Reichs Ordinari Zeitungen / von vnderschiedlichen Orthen and Extra=Zeitungen.
1900 = Berliner Tageblatt (1900). Berlin 1872-1939.
8. References


Carmen Scherer
Johannes Gutenberg-Universität
FB 13, Deutsches Institut
D - 55099 Mainz
cscherer@mail.uni-mainz.de