PRODUCTIVITY OF THE SUFFIXES *-NESS* AND *-ITY* IN 17TH-CENTURY ENGLISH LETTERS: A SOCIOLINGUISTIC APPROACH

Pro Gradu Thesis
Department of English
University of Helsinki
29 April 2008
Tanja Säily

Contents

1. Introduction

1.1. Purpose of the study

The linguistic case I study is as follows. There are two roughly synonymous suffixes, *-ness* and *-ity*, which are typically used for forming abstract nouns from adjectives, as in example (1) below.

(1) *generous* [ˈdʒɛnərəs] + *-ness* → generousness [ˈdʒɛnərəsnɪs]
generous [ˈdʒɛnərəs] + *-ity* → *generosity* [dʒɛnəˈrɒsɪtɪ]

The first suffix, *-ness*, is etymologically native, while *-ity* entered the language as a result of contact with French during the Middle English period. The foreignness of *-ity* can be readily discerned from the above example: it changes the form of its base from [ˈdʒɛnərəs] to [dʒɛnəˈrɒs], whereas with *-ness* there is no change. In addition, the meaning of words in *-ity* is often not entirely compositional, i.e., not deductible from the meanings of the base and the suffix. Thus, it is both phonologically and semantically more opaque than *-ness* (cf. Riddle 1985: 443–444; Aronoff and Anshen 1998: 246).

What I am interested in doing with the suffixes is to compare their morphological productivity, a concept famously defined by Bolinger (1948: 18) as "the statistically determinable readiness with which an element enters into new combinations". More specifically, I wish to examine whether the productivity of each suffix varies between different social groups, as defined by Labovian sociolinguistic categories such as gender and social status. Many linguistic features show sociolinguistic variation, but to date this has been studied little in the case of morphological productivity, and not at all with the otherwise closely scrutinised pair of *-ness* and *-ity*.

My data come from the 17th-century part of the *Corpus of Early English Correspondence* (1998; henceforth known as the CEEC). I have chosen personal letters as my material because they are one of the closest registers to speech, which is the primary medium of language and the most fertile ground for linguistic change (Nevalainen and Raumolin-Brunberg 2003: 28).[1] This time period is interesting because it is to be expected that *-ity* would by this time have spread to wider use from the more literate registers in which it entered the language. Furthermore, a pilot study (Säily 2005) using the smaller *Corpus of Early English Correspondence Sampler* (1998; henceforth known as the CEECS) showed a gender difference in the use of *-ity* in letters of the 17th century.

My hypothesis is that *-ity*, as a learned and etymologically foreign suffix, is (1) less productive than *-ness* in this material; and (2) less productive with poorly educated social groups, such as women and the lower ranks, than with well-educated groups, such as men and the higher ranks. As to the productivity of *-ness*, I do not expect to find significant differences between social groups.

The main method used in this study for measuring productivity is comparing type counts, i.e., how many different words in *-ity* and *-ness* are used by the different social groups. Type counts are by no means a perfect measure of productivity, but they can be made more useful by restricting the kinds of words that are counted. One restriction employed in this study is that the suffixed word must have had an extant base at the time when the letters were written; another is that the word must not have been in the language for much more than a century, as evidenced by its first attestation

---

[1] I use the term **register** in the sense of Biber, Conrad and Reppen (1998: 135), i.e., "as a cover term for varieties defined by their situational characteristics", which include "purpose, topic, setting, interactiveness, mode, etc."

date in the *Oxford English Dictionary* (henceforth the OED). These restrictions increase the probability that the word in question was formed productively from suffix and base rather than retrieved as a whole from the mental lexicon of the writer.

There is, however, a further problem in comparing type counts: as there are different amounts of data from the different social groups, their type counts cannot be compared directly. They also cannot be normalised, because normalisation assumes that the measurement grows linearly with the amount of data, which is not the case with type counts, as will be shown in Section 6.4 below. Samples of equal size could be taken from each group, but this would needlessly discard valuable data. The little-known method used in this study facilitates both comparing data obtained from corpora of varying sizes and establishing the statistical significance of the results. While based on the standard statistical technique of permutation testing, the method has never been used widely in corpus linguistics; furthermore, researchers have mostly used it to verify results from more sophisticated methods involving inter- or extrapolation, not recognising the robustness of this method on its own.

Thus, the merits of this study are twofold. Firstly, it contributes to linguistic knowledge in the fields of morphology and historical socio-linguistics. Secondly, it adapts the statistical method of permutation testing for the corpus-linguistic problem of comparing type counts.

The thesis proceeds in the following manner. After a preliminary discussion of productivity and why it is worth studying (Section 1.2), the theoretical background to the study is presented in two parts, one on morphology (Section 2) and the other on sociolinguistics (Section 3). Next comes an extensive survey of the state of the art: Section 4.1 reviews previous research on *-ness* and *-ity* in present-day English, while Section 4.2 surveys historical studies of the suffixes and productivity in general.

This is followed by a statement of the research question (Section 5) and a description of the material and methods used in the study (Section 6). Section 7 presents the results, which are further explained and evaluated in Section 8. Finally, Section 9 concludes the thesis with a summary of the main points of the study and some implications for future research.

1.2. On productivity

The concept of morphological productivity is a problematic one. Linguists cannot seem to agree on quite what it comprises (cf. Bauer 2001: 1) and how it might be assessed in each individual case or in general. Nevertheless, I have decided to choose it as one of my bases for analysis; in this section, I shall explain the reasons for my decision.

Firstly, morphological productivity is connected to something that has been recognised as a fundamental property of language at least since Humboldt (Robins 1990 [1967]: 192–193): speakers' ability to create infinitely many new combinations out of the finite linguistic resources they have at their disposal. In word-formation, this is manifested in the way speakers can make new words based on existing words or word-forming elements (Plag 2006: 537) with the help of a few guidelines. If a certain morphological process (such as suffixation with *-ness*) can be used by speakers to coin new words, then that process may be called productive.

Of course, it is debatable how big a proportion of everyday speech, whether at the level of sentences, phrases or words, actually is 'new' in the sense 'never heard before'. Furthermore, one might guess that the lower the level, the smaller the probability of newness; thus, the coining of new words might be seen as a marginal phenomenon in terms of frequency. Baayen and Renouf (1996: 75), for example, found in their *Times* newspaper corpus of roughly 80 million words only 348 new words formed with the suffix *-ness* and 143 with *-ity*. However, they defined 'new' in a very

narrow way: a word was considered new if it occurred only once in the corpus and was not listed in a major dictionary (1996: 76).

What then would be a more pertinent definition of newness? Baayen and Renouf (1996: 76) use the above definition because they approach the issue from the point of view of the language community or language as a whole: a word is only new if it has not appeared elsewhere in the language community, whose language is represented by corpora and dictionaries. However, as Baayen and Renouf themselves point out (1996: 77), there is another conceivable viewpoint: that of the individual user of the language.

Even if a word has appeared somewhere in the language community, it may be new to individual users of the language — in fact, Baayen and Renouf (1996: 77–78) claim that rare words may be new to, or not listed in the mental lexicons of, *most* of the users of the language. Psycho-linguistically speaking, words that have a frequency of 1 per million are already considered very rare indeed and are probably not listed in the lexicons of the users; therefore, words that occur once in an 80-million-word corpus could well be considered new even without a dictionary check (Baayen and Renouf 1996: 78). There are 739 such *-ness* words and 280 *-ity* words in their corpus, which are in my opinion fairly large numbers considering that the total number of different words in the corpus is 2,027 for *-ness* and 1,020 for *-ity* (1996: 83–84). Of course, the numbers would have been even bigger if Baayen and Renouf had taken into account a wider range of low-frequency words than just those occurring once (1996: 78).

It seems, then, that the production of words that are new to the individual user is not a marginal phenomenon, at least not in newspaper English. Baayen and Renouf go so far as to propose that "at least in reading, productive word-formation rules are put to use on a regular daily

basis" (1996: 94). Therefore, I would see morphological productivity as worth studying.

2. Morphology

2.1. Basic concepts

Before embarking on a study of suffixation, some basic concepts need to be introduced and defined. The concept of a **word** I shall take *a priori*, but I shall split from it some more specific concepts. A **lexeme** comprises all the possible shapes that a word can have, such as *shoot*, *shoots*, *shooting* and *shot* for the verbal lexeme *shoot*; the individual shapes are called **word-forms** (Bauer 1983: 11). A **lemma** is the word-form conventionally used to represent a lexeme, e.g., in a standard dictionary (Bauer 1983: 12, who calls it a citation form). Especially in older material, there may be variation in the spelling of the word-forms; these variants, which are what we actually see in the text, I shall call **orthographic forms**.

According to Bauer (1983: 13), **morphology** is the study of the internal structure of word-forms. As noted by Plag (2003: 10), a **complex word** like *unfaithfulness* (my example) can be broken down into its smallest meaningful units, **morphemes**: *un*-, *faith*, *-ful* and *-ness*. Plag (2003: 10) classifies morphemes into two kinds: **free morphemes** such as *faith* that can occur by themselves, and **bound morphemes** such as *un*-, *-ful* and *-ness* that can only occur with other morphemes. A free morpheme occurring by itself is called a **simplex** (Bauer 1983: 30) or **mono-morphemic word** (Plag 2003: 25).

According to Plag (2003: 10–11), the central meaningful element of a word can be called the root, base or stem. Bound morphemes that attach to the central element are called **affixes**; these can be divided into **prefixes** (such as *un*-), which occur before the central element, **suffixes** (such as

-ness), which occur after it, and **infixes** (such as *-bloody-* in *abso-bloody-lutely*), which occur inside it. Plag (2003: 10–11) explains the different terms for the central element as follows. The **root** consists of a single morpheme that can be either free like *faith* or bound such as the Latinate *simul-* (as in *simulant*, *simulate*, *simulation*). The **base** is a wider concept: it is used for any central element, whether an indivisible root or a complex word, to which an affix can be added. The **stem** has various meanings in the literature, the most common of which is 'the base of an inflection'; following Plag's (2003: 11) lead, I shall avoid using this ambiguous term.

Plag (2003: 20–21) sees the morpheme as a linguistic sign that has two sides: form and meaning. For example, the morpheme *un-* consists of the form, or **morph**, [ʌn] and the meaning 'not'. The form of a morpheme can vary; these variants are called **allomorphs** (Plag 2003: 27–28). For instance, the form of the base *eccentric* [ɛkˈsɛntrɪk] changes when the suffix *-ity* is attached to it: [ɛksɛnˈtrɪs]+[ɪtɪ] (Romaine 1985: 451). Plag (2003: 21) says that when two morphemes are combined, the meaning of the resulting complex word is often **compositional** and hence transparent — e.g., *un-* 'not' + *happy* 'happy' = *unhappy* 'not happy'. He notes (2003: 22), however, that this is not always the case — for example, *late* 'after the due time' + *-ly* 'in an X manner' = *lately* 'recently', not 'in a late manner' (see the discussion on lexicalisation below).

The **morphological process** of adding an affix to a base is called affixation (more specifically, prefixation, suffixation or infixation); this can be either inflectional or derivational. **Inflectional affixation** is used to create the different word-forms of a lexeme (Bauer 1983: 29); it encodes grammatical categories such as plural, person, tense or case (Plag 2003: 14). **Derivational affixation** is used to create new lexemes (Bauer 1983:

29), and it is a subtype of this process that I am concerned with here: creating new words by using *-ness* and *-ity* suffixation.

How do speakers form new words? According to Plag, word-formation is not an arbitrary process but seems to be **rule-governed**: for example, most adjectives can take the suffix *-ness*, and the resulting noun will regularly have the meaning 'the property of being X', where X denotes the meaning of the base (2006: 537). Or, given the words *unhappy*, *unkind*, *unfaithful*, *untrue*, *uncommon* and *analysable*, a speaker can easily decipher the meaning of *unanalysable*, even if she has not encountered that word before (Plag 2003: 30). There must be some kind of system in speakers' minds that makes this possible; according to Plag (2003: 37–38), some say it is the general mechanism of **analogy** that is at work, while others claim that when there are multiple instances of the same pattern, there must be a rule by which they are formed.

A typical word-formation rule might look like the one presented in (2), adapted from Plag (2003: 35):

(2) Word-formation rule *un*-$_1$

phonology:	/ʌn/-X
base:	X = adjective
semantics:	'not X'
constraints:	– derivatives with simplex bases must be interpretable as contraries – further restrictions on possible base words …

Analogy, on the other hand, is simply "a proportional relation between words", as exemplified in (3) below (Plag 2003: 37). In the first example, the relationship between items a and b is the same as the relationship between items c and d. Item d has been formed from c on the pattern of a : b. Concrete examples are provided in ii–iv.

(3) i. a : b :: c : d
ii. eye : eyewitness :: ear : earwitness
iii. ham : hamburger :: cheese : cheeseburger
iv. sea : sea-sick :: air : air-sick

In Plag's opinion (2003: 38), the advantage to a rule-based approach is that it explains the existence of systematic structural constraints on morphological processes as well as why some processes are more frequently utilised than others: the constraints are explicitly listed in the rule, and processes that are never or seldom used just do not have a rule associated with them (Bauer 2001: 77). However, as both Plag (2003: 38) and Bauer (2001: 96) admit, analogy is certainly employed to some extent; furthermore, I do not think that rules as clear as the one in (2) really exist in speakers' heads — the reality must be much fuzzier than that, with analogy playing a large part and interacting with other factors such as speakers' knowledge about how other speakers use the forms in question. The fuzziness hypothesis is supported by the considerable number of exceptions to the strict rules proposed by linguists (cf. Bauer 1983: 293–294).

In addition to fuzzy word-formation rules, speakers must have some words stored in their minds to which the rules can be applied. This storage space is called the **mental lexicon** (Plag 2003: 4). Words listed in the mental lexicons of speakers are called **existing words**, while words that are not listed there but could be formed by a rule are called **potential words** (Plag 2003: 46–47). Existing words can develop idiosyncratic meanings or pronunciations by a process known as **lexicalisation** (Bauer 2001: 44–45); the above-mentioned *lately* 'recently' is a case in point. Another good example is the word *business* ['bɪznɪs] 'the production of goods and services for profit', which has diverged in both form and meaning from the original ['bɪzɪnɪs] 'the state or property of being busy'.

Armed with these concepts, we may now tackle the issue of defining productivity.

2.2. Morphological productivity

Morphological productivity is a multi-faceted phenomenon; as Plag (2006: 547–549) shows, it is a derived notion instead of a theoretical primitive, but potentially useful in describing word-formation. Plag (2003: 44) defines productivity as the "property of an affix to be used to coin new complex words". Following the structure in Plag (2006), I shall discuss the qualitative, quantitative and psycholinguistic aspects of productivity, as well as some pragmatic and structural restrictions or constraints on it.

2.2.1. Productivity as a qualitative notion

Productivity can be conceived of as a qualitative, either–or notion: either an affix can be used to coin new words or it cannot. This view is advocated by, e.g., Bauer (1983: 99–100), who does not consider semi-productivity a useful construct. Plag (2006: 540), on the other hand, proposes three categories of morphological processes: those clearly unproductive, those clearly productive and those in between. I am not convinced of the usefulness of either of these views. It seems to me that an affix, or the process of forming words with it, can never be said to be clearly unproductive — there is always the possibility that somebody uses it to coin a new word. This one-off use can be called analogy instead of productivity, but where do we draw the line between the two; how many words must be coined for a process to be called productive?

This question is also posed by Plag (2006: 539–540), and it leads him to the three-way classification presented above, but that does not in my opinion really answer the question. Which would we classify as clearly unproductive and which as in-between? Besides, as Dalton-Puffer (1996: 222) points out, it is possible that analogy only differs from rule-based productivity in degree rather than in kind, so again there are no clear-cut boundaries (cf. Bauer 2001: 97 and the discussion in Section 2.1 above).

Furthermore, just like the distinction between clearly unproductive and in-between processes, the distinction between in-between and clearly productive ones is far from being straightforward. Again, how many new words must be coined for a process to be called clearly productive rather than in between; or are there some other criteria by which the classification can be made? Plag's exact definition of the in-between category is "those processes that are not easily classified as either productive or un-productive" (2006: 540) — I think most, if not all, processes would fall into this category, which would make the categorisation somewhat pointless.

Therefore, it seems to me that rather than asking *whether* a process is productive or unproductive or semi-productive, a better research question would be to ask *how* productive it is along some scale (or several), perhaps in comparison with another process or among different groups of people. This is, in fact, precisely what I aim to do in the present work. This scalar view of productivity will be discussed in the following section.

2.2.2. Productivity as a quantitative notion

Productivity can also be conceived of as a quantitative notion: an affix can be used to coin new words to some degree. Several ways of measuring this degree have been proposed in the literature. Baayen (1993) presents three measures, which he calls the category-conditioned degree of productivity (P), the hapax-conditioned degree of productivity (P*) and the activation level (A). All of these are based on counting **tokens** (N) and **types** (V) of words belonging to a certain morphological category — for example, how many instances of *-ness* words and how many different *-ness* words a corpus contains, respectively. Of special interest are the so-called hapax legomena or **hapaxes** (n_1), words that occur only once in the corpus, because these are seen to predict the number of new words.

According to Baayen and Lieber (1991: 810), a large proportion of the types of productive affixes are hapaxes, and the frequency distribution of the types is asymmetrical in general: there are more types that occur once than those that occur twice, more types that occur twice than those that occur three times, and so on. Overall, there are many types that occur only a few times in the corpus, and few types that occur many times. With less productive categories, the number of hapaxes is lower (there may be more **dis legomena**, types that occur twice, than hapaxes), and the frequency distribution is less skewed.

The **category-conditioned degree of productivity** P is defined as the ratio between the number of hapaxes with a given affix and the total number of tokens with that affix in the corpus: $P = n_1/N$. According to Baayen and Lieber (1991: 809–810), it expresses the probability of observing new types with the relevant affix when N tokens with the affix have been sampled. If the size of the corpus is increased, N will increase, and so will the number of types V, but at a different rate from N (Baayen and Lieber 1991: 811). This can be illustrated by drawing a graph with the values of V at different points of sampling on the *y* axis and the values of N on the *x* axis; in other words, V may be plotted as a function of N, V(N). See Figure 1 for a schematic example using not affix types and tokens but all of the different words and the number of running words in a text.

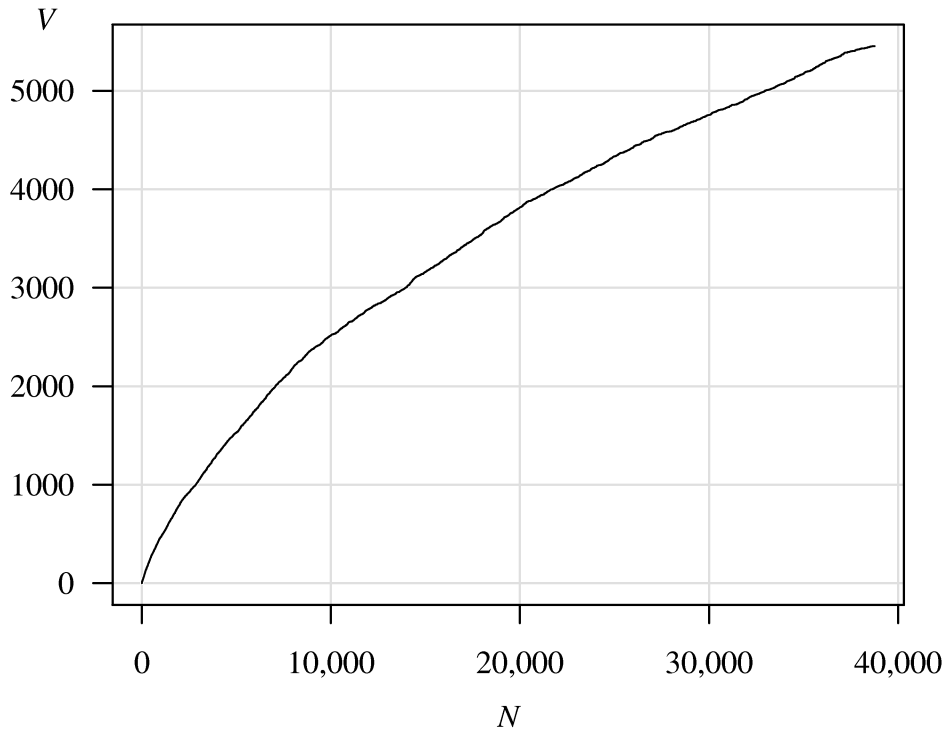

Figure 1. The growth curve of all types V as a function of all tokens N in the Project Gutenberg e-text of Joseph Conrad's *Heart of Darkness*, <http://www.gutenberg.org/2/1/219/>.

When only a few tokens have been observed, i.e., when N is small, new types will be found quite frequently, i.e., V will grow rapidly. As more and more types are found, the rate of growth will slow down. It is precisely this rate of growth that is expressed by the category-conditioned degree of productivity P. If the number of tokens observed is M, then P(M) is the slope of the tangent to the growth curve of V in the point (M, V(M)). A large value of P(M) indicates that there are many types yet to be sampled, which would suggest that the affix is productive. If, on the other hand, P(M) is small, the growth curve is about to even out, and the number of new types to be expected is small, which would suggest that the affix is unproductive (their term) or less productive (my term). (Baayen and Lieber 1991: 811–812.)

Baayen and Lieber (1991: 817) point out that once we have calculated the P for all of the tokens N in our sample, we know little about when the

growth curve of V would flatten out if the sample size were increased; we cannot use P to predict the absolute number of types that would be found in a larger sample. In other words, P is dependent on the size of the corpus. Furthermore, because P is based on the number of tokens of a particular morphological category, it cannot be directly compared with a P calculated from the number of tokens of another morphological category, unless the numbers of tokens are of a similar magnitude (Baayen 1993: 191). For example, we cannot compare the degrees of productivity calculated for *-ness* and *-ity* in the same corpus if there are many more tokens of *-ness* than of *-ity*.

Baayen's (1993) second measure of productivity is the **hapax-conditioned degree of productivity** P*. It is defined as the ratio between the number of hapaxes with a given affix and the total number of all hapaxes in the corpus: $P^* = n_1/h$. This measure indicates how much the affix contributes to the overall vocabulary growth of the corpus (Baayen 1993: 193). According to Hay and Baayen (2003: 101), the hapax-conditioned degrees of productivity of different affixes within the same corpus can be compared by using n_1 counts alone, h being constant.

2.2.3. Productivity as a psycholinguistic notion

Baayen also has a third measure of productivity, the **activation level** A (1993: 195–196). It is defined as the number of tokens representing those types of a given affix whose frequency of occurrence is smaller than a threshold θ. The measure is motivated by the idea that people process complex words through two competing routes simultaneously: by **parsing** and by retrieving the whole word directly from their mental lexicon. Which route is faster depends on the frequency of the word: common words are readily available in the lexicon, while words rarer than the threshold value are processed by identifying and combining the affix and base. This

process of parsing maintains the activation level of the affix; the level indicates how quickly the affix will be recognised and combined with the base.

Baayen (1993: 203) admits that choosing a suitable threshold value is problematic. He assumes that only words that are semantically transparent will maintain the activation level. The higher θ is, the more semantically opaque words it will include: words are usually the more opaque the more frequent they are, because frequent words have accumulated more meanings or their meaning may have changed. Therefore, the threshold is best kept fairly small.

Hay (2001) improves on this measure by introducing the concept of relative frequency; she shows that it is not the frequency of the affixed word alone that affects parsability, but rather the relation between the frequencies of the affixed word and its base. Her work will not be discussed further here because it is not directly applicable to unlemmatised historical corpora, in which it would be an all too time-consuming task to track down all of the different spelling variants of the bases and search for each of them.

In addition to the semantic transparency mentioned above, phonological transparency has also been shown to affect productivity. According to Hay and Baayen (2003: 105), there is a wealth of evidence that English-speaking people use phonotactic probabilities when parsing words they hear. For example (2003: 105), as the /pf/ transition in *pipeful* is unlikely to occur within a mono-morphemic word in English, people hearing this combination in running speech will immediately posit a morpheme boundary between *pipe* and *ful*, which will help them to decompose the complex word. These unlikely combinations facilitate the parsing route rather than the whole-word route of processing; therefore, complex words containing unlikely transitions at morpheme boundaries are more likely to

be stored separately as affix and base rather than as a whole word in the mental lexicon; this in turn means that the affixes contained in complex words like these are more likely to be used productively (2003: 105–106).

Frequency and transparency are by no means the only factors determining the productivity of an affix. There are a number of other constraints relating either to language use (pragmatic constraints) or to language structure (structural constraints), and it is to these that we now turn.

2.2.4. Pragmatic constraints

One productivity constraint that I think could be classified as pragmatic is the usefulness of the potential word for members of the speech community. Kastovsky (1986: 594–595) names two functions of word-formation: labelling and syntactic recategorisation. The former is used to refer to a new concept or object, while the latter replaces a phrase or a clause with a single complex word in order to condense information, create stylistic variation or facilitate text cohesion (Plag 2003: 59–60). The two functions are exemplified in (4) and (5), respectively.

(4) The Time Patrol also had to unmurder Capistrano's great-grandmother, unmarry him from the pasha's daughter in 1600, and uncreate those three kids he had fathered. (Kastovsky 1986: 594)

(5) If that's not civil, civilize it and tell me. (Kastovsky 1986: 595)

Plag (2003: 59–60) adds a third function, namely that of expressing an attitude, as in (6):

(6) Come here, sweetie, let me kiss you. / Did you bring your wonderful doggie, my darling? (Plag 2003: 59)

According to Plag (2006: 550), one of the most important pragmatic factors affecting productivity is fashion. As noted by Renouf (2006), there is a kind of ebb and flow in what is in vogue and what is not. For example, extralinguistic developments can be seen to have influenced the extent of use of the affixes *mega-*, *giga-*, *mini-* and *-nik* (Plag 2006: 550). The

second pragmatic constraint mentioned by Plag is that new words must "denote something nameable" (2006: 550). By this he means that the new concept cannot be overly complex — typical derivative affixes only add a very simple and general meaning to that denoted by the base (e.g., adjectival *un-* 'not X').

Plag (2003: 61) warns that we should not automatically assume a pragmatic reason for why some new formation is impossible — there may well be structural constraints involved, and the existence of these should in fact be checked before entering into any usage-based speculations.

2.2.5. Structural constraints

According to Plag (2006: 550–551), structural factors influencing productivity can be divided into phonological, morphological, syntactic and semantic constraints. Furthermore, they can be either general or process-specific; the latter may relate to what the base or the derived word must be like. As Plag (1999: 43–44) states, however, the boundaries between these divisions can be fuzzy. Let us first have a look at examples of process-specific constraints (Table 1 below). Here the first three constraints apply to the base and the last one to the derived word.

Type	**Example**	**Constraint**
Phonological	Suffixation of verbal *-en* (as in *blacken*)	Only attaches to base-final obstruents, does not take bases that have more than one syllable
Morphological	Suffix combination *-ize-ation*	Words ending in the suffix *-ize* can be turned into a noun only by adding *-ation*
Syntactic	Adjectival suffix *-able*	Normally attaches to verbs
Semantic	Suffix *-ee* (as in *employee*)	Derivatives with the suffix must denote sentient entities

Table 1. Process-specific structural constraints on productivity (Plag 2006: 551).

As for general structural constraints, Plag (1999: 45) lists ten of these but swiftly dismisses most of them. Among the more tenable ones is the **unitary output hypothesis**, which states that derivatives from a certain

word-formation process form a group uniquely distinguishable from others through its phonological, syntactic and semantic properties (Plag 1999: 49). While Bauer (2001: 127–128) points out that semantic unity or disunity is in the eye of the analyst, he admits that the hypothesis is relatively uncontroversial.

Another general constraint is **blocking**, which is defined by Aronoff (1976: 43) as "the nonoccurrence of one form due to the simple existence of another". Van Marle (1986: 607) distinguishes between two special cases, which Rainer (1988: 159) calls token-blocking and type-blocking. According to Rainer (1988: 161), **token-blocking** occurs when the creation of a morphologically complex word, such as *stealer*, is blocked by an existing synonymous word, such as *thief*. Rainer shows that it does not matter whether the blocking word is idiosyncratic (like *thief*) or regularly derived, as long as it is stored in the lexicon (1988: 164–167).

According to Rainer (1988: 161–164), token-blocking may only occur under the three conditions of synonymy, productivity and frequency. Firstly, the blocked word and the blocking word must be truly synonymous; secondly, the blocked word must be a potential word in the sense that there is a productive morphological process by which it could be formed; and thirdly, the blocking word must be frequent enough to be retrieved from the lexicon faster than the blocked word can be formed.

Van Marle (1986: 613–617) notes that some blocking words radiate a stronger "blocking force" than others; Rainer (1988: 163–173) shows that the strength of the force depends on the frequency of the blocking word. The less frequent the stored word is, the greater the likelihood that the speaker will fail to activate it, which according to Plag (2006: 552) explains the occasional failure of blocking and the occurrence of synonymous doublets. However, Rainer (1988: 164) remarks that the blocking force is resisted by another force dependent on the productivity of the

morphological process that would produce the potential word; it is the interplay of these two forces that actually determines whether blocking succeeds or fails.

To predict which words are more likely to have doublets, I think Baayen's activation level (see 2.2.3 above) could be used, so that the words falling within a certain minimum frequency range would be the most likely candidates. Conversely, perhaps the token frequencies of the doublets occurring in a given corpus could be used to select a suitable threshold for Baayen's measure.

According to Plag (2003: 66), **type-blocking** "has been said to occur when a certain affix blocks the application of another affix". Van Marle's (1986: 608) domain hypothesis suggests that affixes can be divided into two groups: special cases, which can only be applied within a restricted domain subject to specific constraints, and general cases, whose domains are unrestricted except for the paradigmatic limitation that they do not include bases belonging to the domains of rival special cases. For example, *-ness* suffixation could be seen as a general case that is blocked by the special case of *-ity* suffixation (Plag 2006: 552).

Plag (2006: 552–553) shows, however, that there are at least three problems with this kind of analysis: Firstly, even though one of Rainer's (1988: 173) preconditions for type-blocking is synonymy, *-ness* and *-ity* are not always synonymous, as shown by Riddle (1985). Secondly, there are plenty of attested doublets, which means that the domains do not completely exclude one another. Thirdly, it is unclear how putative cases of type-blocking can be distinguished from token-blocking — if some form in *-ness* is avoided, how can we rule out the possibility that it is because the equivalent form in *-ity* exists in the lexicon? The first two problems may not apply to all cases, but I think the third problem is a crucial one. Plag (2003: 67–68) in fact suggests that we should reject the notion of type-

blocking altogether, as we can only verify the existence of token-blocking. I agree; however, I think this should not be taken as an indication that affix rivalry does not exist but simply that it cannot be adequately described with the concept of type-blocking as defined above.

3. Sociolinguistics

3.1. Basic concepts

This study is a sociolinguistic one: it deals with language in the context of society. Therefore, a brief introduction to the theory and terminology of sociolinguistics is in order. A key idea of sociolinguistics is that language variation can be socially significant and that language change can be socially motivated. One of the first people to recognise the importance of this fact to linguistics was Labov, whose *Sociolinguistic Patterns* (1978 [1972]) has become a classic in the field. For example, Labov (1978 [1972]: 4–42) shows that the centralisation of certain diphthongs in the speech of the inhabitants of Martha's Vineyard has the social meaning of indicating that the speaker is a native Vineyarder. The observed increase in centralisation over time he shows to have originated with the fishermen, who were the most opposed to the flood of summer visitors to the island.

Labov concentrates on macro-level sociolinguistics, i.e., "large-scale social factors, and their mutual interaction with languages and dialects" (1978 [1972]: 183). This is also my concern here. The factors most commonly studied include social class, gender, age and region. According to Milroy and Gordon (2003: 95), the category of **social class** comes from sociology, where it is used in two different models. The first model goes back to functionalist sociology and describes social classes as a flexible continuum of shared values and *consensus*, while the second model developed by both Marx and Weber treats class divisions as discrete and

based on *conflict*. The consensus model sees occupations as the main way of distinguishing between different classes, whereas the conflict model gets its divisions from people's different relations to the market.

Milroy and Gordon (2003: 96–97) argue that Labovian socio-linguistics, with its peaceful and harmonious concept of speech community, overwhelmingly follows the consensus model. There are nevertheless some sociolinguists who question it, pointing out that the very fact that there are so many vigorous non-standard vernacular communities could be interpreted as evidence of conflict. Indeed, if there were no conflict, where would language change come from? However, Milroy and Gordon note that both models are potentially useful in sociolinguistics and that different kinds of data require different approaches.

Another social category that has proved to be a strong factor in language variation and change is **gender**. Like Nevalainen and Raumolin-Brunberg (2003: 110), I use the term gender instead of sex, because gender is a social construct, the characteristics of which can change over time. According to Milroy and Gordon (2003: 103), one generalisation to be made from previous work on present-day sociolinguistics is that women seem to prefer supralocal forms, i.e., ones that are fairly widely distributed, whereas men prefer local forms, which are often stigmatised. Thus, women are often the leaders of supralocal language change, also known as supralocalisation, in which a linguistic feature spreads from one region to neighbouring areas (Nevalainen and Raumolin-Brunberg 2003: 112).

According to Labov (1994: 46, 83–84), linguistic differences between different **age** groups can be due to either *age-grading*, which is a regular change of linguistic behaviour with age repeated in each generation, or actual *generational change*. As it seems that the latter option is more likely (Chambers 1995: 8), sociolinguists have conducted so-called *apparent-time* studies, in which they study the language of members of different

generations in order to track language change. There can also be change in which all age groups behave in the same way; this Labov (1994: 83–84) calls *communal change*.

The study of **regional** variation and change comes to sociolinguistics from dialectology (Milroy and Gordon 2003: 11–12). According to Nevalainen and Raumolin-Brunberg (2003: 165), the process of regional change can follow the *wave model* of spreading gradually outwards from a centre, such as from an urban area to the surrounding rural areas. An alternative mechanism of diffusion is *dialect hopping*, in which a change hops from one urban centre to another, skipping the rural areas in between (Nevalainen and Raumolin-Brunberg 2003: 165).

In sociolinguistics, the factors described above are studied in relation to the **linguistic variable**, which Milroy and Gordon (2003: 88) define as a linguistic item with variant realisations that refer to the same thing but covary with different items or social categories. According to Milroy and Gordon (2003: 88), the use of a variant can be described quantitatively, in terms of percentages, rather than as an either/or situation. The methodology for investigating such variation was largely developed by Labov (1978 [1972], etc.). In my study, I regard *-ness* and *-ity* as variants between which speakers can choose when they wish to form an abstract noun meaning something like 'the property of being X'. This is not an unproblematic point of view to take, as the suffixes are not completely synonymous, but it can in my opinion be justified (cf. Hudson 1996 [1980]: 171).

In addition to contemporary studies, sociolinguistics has recently begun to be applied to historical material. According to Nevalainen and Raumolin-Brunberg (2003: 2), the first systematic attempt at this was made by Romaine (1982). Nevalainen and Raumolin-Brunberg themselves are pioneers in this field, which is now called **historical sociolinguistics**. Besides their mother disciplines of sociolinguistics and historical

linguistics, historical sociolinguists draw on social history to ensure the social and historical validity of their work (Nevalainen and Raumolin-Brunberg 2003: 8–11). To ensure the validity of my work, I need to reconstruct the sociohistorical situation of the period of my study, 17th-century England.

3.2. English society in the 17th century

3.2.1. Overview

According to Wrightson (1993: 112–132), the population of England doubled between the 1520s and 1680, from circa 2.5 million to circa 5 million people. This caused migration to economic opportunities, which were to be found among pastoral agriculture, centres of manufacture and urban centres, especially London, whose population increased tenfold between the 1520s and 1700. Population growth and urbanisation increased the demand for foodstuffs and other merchandise, which in turn led to the integration of English economy and the expansion of internal trade. Landlords were able to extract higher rents for their lands; some of them also farmed and developed new techniques.

Wrightson (1993: 140–148) describes the period 1580–1630 as one of gathering crisis, at least for the poor: there were harvest failures, wars, famines and attacks of the plague. After that, 1630–1680, came recovery: population growth stabilised, fertility was reduced, there was less subsistence migration, crops were larger, prices were lower, wages were higher, and there was more demand for manufacture. The polarisation of society remained, however, and there were large numbers of poor people.

Nevalainen and Raumolin-Brunberg (2003: 31) see two developments as the most important ones in Tudor and Stuart England: the reformation in the sixteenth century and the revolution in the seventeenth. The impact of

the reformation, viz. the encouragement of lay literacy and the transferring of church lands into private hands, can be seen in the seventeenth century as well. As for the revolution, Nevalainen and Raumolin-Brunberg (2003: 31–32) note that while it did not have a lasting effect on social structures, the decades of upheaval certainly affected the minds of those who lived through them, and the increased geographical mobility which caused more encounters between people from different backgrounds seems to have facilitated language change.

Even though Wrightson (1993: 222–228) downplays the effect of the civil war on English society, he does take the year of the restoration, 1660, as a sort of landmark: by then he sees a more closely integrated national society as well as more local social stratification, which led to conflicts. The middling sort became closer to its superiors, participating in national affairs such as the revolution, and was divided from the poor multitudes, who were not participants but objects.

3.2.2. Social structure

In this period, we talk about social **rank** rather than class. As Laslett (1965: 22) explains, the class system had not yet arisen — if class is defined as "a number of people banded together in the exercise of a collective power, political and economic", in this pre-industrial society there was only one class, and most people would not have belonged to it. People did have different status levels, however, and for those levels I (following Nevalainen and Raumolin-Brunberg 2003: 33) choose to use the contemporary term rank.

There are many ways of dividing 17th-century society into ranks. A rough dichotomy would be gentry vs. non-gentry (Laslett 1965: 26); a tripartite division into the better sort, the middling sort and the poorest sort has also been suggested by Nevalainen and Raumolin-Brunberg (2003:

136, Model 4).[2] In Nevalainen and Raumolin-Brunberg's (2003: 136) Model 3, these contemporary labels are exchanged for the more neutral upper, middle and lower ranks, with the additional category of social aspirers above the middle ranks. Social aspirers were middle-ranking people who advanced to the upper ranks — for instance, merchants who became gentlemen or members of the upper clergy — and who would have wished to show their learning and gentility even in their language use.

Nevalainen and Raumolin-Brunberg's (2003: 136) Model 2 is an even more fine-grained division, with royalty at the very top, followed by nobility, gentry, clergy and social aspirers; next professionals (e.g., army officers, lawyers, medical doctors and teachers) and merchants; and, finally, other non-gentry (such as yeomen, husbandmen, craftsmen, labourers, cottagers and paupers). In Nevalainen and Raumolin-Brunberg's (2003: 136) Model 1, their most fine-grained model, the basic distinctions are the same as in Model 2, but the gentry is further subdivided into upper gentry (consisting of knights and baronets) and lower gentry (including esquires and gentlemen); furthermore, both upper and lower gentry are divided into a non-professional and a professional section, the latter of whose members held high government offices (2003: 137). Clergy, too, is divided into upper (bishops) and lower (the rest).

According to Wrightson (1993: 27–30), the rural gentry was as a rank preferred to the urban merchants and professionals; nevertheless, they were closely related, as merchants and professionals were often the younger sons of gentry, and a successful merchant or professional could acquire land and retire to the country, thus becoming a part of the gentry. The ownership of

---

[2] The better sort consists of royalty, nobility, gentry and clergy, while the middling sort are professionals and merchants, and the poorest sort other non-gentry (see Models 1–2 below).

land and freedom from manual labour were the crucial criteria in deciding who was a gentleman (1993: 25). There were other avenues open to urban aspirers besides the rank of a country gentleman, however: professionals could advance to powerful governmental positions, and merchants could be active in guilds and the government of the city.

The category of **gender** is a pertinent one in 17th-century society as well. Women had an inferior status in comparison with men, and their rank came from the rank of their father or husband. The husband was the head of the household, and wife-beating was allowed by law, though frowned upon by people. The letters of gentlewomen to their husbands in this period show an anxiety to please (Wrightson 1993: 94), and women were expected to speak modestly in mixed society; in all-female contexts they could speak more freely (Mendelson and Crawford 1998: 212–213). Lower down on the social scale women could be more assertive in their speech (Wrightson 1993: 96), and rhetoric (including scolding, gossip, storytelling and folk-lore) was indeed one form of empowerment for women (Mendelson and Crawford 1998: 215–218).

Whereas men could freely move in both public and private spheres, women were mostly confined to the private sphere, with little opportunity for higher education or participation in the running of the society. According to Mendelson and Crawford, the situation changed somewhat during the civil war, but after the restoration women were forced to retreat from the public sphere at least formally (1998: 401–2, 419). As for employment, in addition to taking care of household chores in their own home, women could work as servants (young ladies of the gentry could be in the service of a noble kinswoman or the queen), and the poorer sort could assist their husbands in the shop or even work in the field (Laslett 1965: 2–3, 11–12). Mendelson and Crawford say that women were on the whole less mobile than men (1998: 301), which could mean that in this

period it was men rather than women who were the leaders of supralocal language change.

In this society, **age** mattered much as it does today. Children were counted in their fathers' ranks and elders were to be respected. Child-rearing was different in that many children were sent away to be raised in other people's households when they reached adolescence, often as servants or apprentices (Wrightson 1993: 41–42, 112). Even gentlemen's children were sent to the households of relatives or friends; this was seen as an important part of their education. Thus, the young did different things and were more mobile than older people, which could have affected their language as well. People would also have adapted their language depending on whether their audience was older or younger than they themselves were.

As for **regional** differences, perhaps the most important thing to be considered is migration. During the period of rapid population growth, 1580–1630, there was a great deal of migration to better economic opportunities, as noted in Section 3.2.1 above. This meant that especially London became a melting pot of different dialects, and social mobility was easier there (Nevalainen and Raumolin-Brunberg 2003: 38–40). During the civil war, both soldiers and civilians were forced to move around a lot, and some royalists had to go into exile on the continent for the interregnum. Outbreaks of the plague, the last of which occurred in 1665 (Wrightson 1993: 146), caused those who were able to move away from the infected places.

In addition, of course, young people (except for the eldest son and heir of each family) moved out of their childhood homes. Gentlemen's sons would go away to university, which would unify their customs (Wrightson 1993: 191–192), and many would take a "grand tour" of Europe as a finishing touch to their education. Country gentlemen made frequent visits to London as members of parliament and in legal matters, sometimes

bringing their families along to enjoy the season's entertainment. The royal court, comprising the members of the royal family and high-ranking government officials, formed an entity of its own, even though it was physically located in the London area (Nevalainen and Raumolin-Brunberg 2003: 51).

3.3. English language in the 17th century

3.3.1. Overview

This section widens the focus from society to sociolinguistics. According to Nevalainen and Tieken-Boon van Ostade (2006), the linguistic situation in 17th-century England came to be as follows. After the Norman Conquest, French had replaced English as the language of government and administration, with Latin retaining its position as the language of scholarship and the church. The society was thus triglossic: French and Latin were used in high-prestige contexts, while English became a low-prestige language spoken in everyday situations. French, however, had lost much of its former prestige by the time of Henry V's reign (1413–22), and English began to take its place, becoming once again a literary language. The next century saw English taking on even more functions, for it was now sometimes used instead of Latin in academic discourse and even in the church, thanks to ideas brought about by the Renaissance and the Reformation. By the 17th century, conscious efforts were being made towards the standardisation of written English, while spoken English was left more or less to its own devices.

What kind of English was it, then, that they used in the 17th century? In the aftermath of the Norman Conquest, French loanwords had flooded the language; the erosion of its inflectional morphology had become increasingly rapid, while derivational morphology had lost some affixes and

gained many others; syntactically, word order had become more rigid. With the Renaissance, Latin became a more important source of borrowing. New vocabulary was needed for the widening range of functions of English, and the prestigious Latin style was imitated in syntax as well; furthermore, the spelling (and sometimes the pronunciation) of many French loans was 'etymologised' so that it followed more closely its ultimate Latin origin.

The influence of Latin was not all-pervasive, however. Initially, only people with a classical education — most often men from the upper ranks (see 3.3.2 below) — were able to fully understand and adopt the Latinate vocabulary and style. Therefore, their use was restricted mostly to those situations in which the participants were learned men. In other words, their use was both socially stratified and context-specific.

Even among learned men, the influx into English of Latinate words and affixes, often nearly synonymous to existing native ones, was not universally embraced. Purists preferred to coin new words using native processes of word-formation, and Latinate words were ridiculed as "inkhorn terms". Nevertheless, some of them stayed in the language, to the extent that the earlier triglossia with English, Latin and French could be said to have been replaced by a diglossia within English between the native 'low' variety and the Latinate/Romance 'high' variety. Latinate specialist terms were collected in hard-word dictionaries such as Cawdry's (1604), which was geared at "ladies, gentlewomen, or any other unskilfull persons".

3.3.2. Education and literacy

Around 1600, only about a third of English males was literate in the sense of being able to both read and write (reading was taught before writing, and they were seen as two different skills). Women's level of literacy was much lower throughout the Early Modern English period. There were also

regional and social differences in the ability to write and spell: there was more full literacy in towns, especially London, and among the highest ranks. Between 1580–1700, Wrightson (1993: 190–191) estimates that the level of illiteracy was 0% for professionals, 35% for yeomen, 79% for husbandmen, 85% for labourers and servants, and 89% for women!

In the 16th century, most of the male gentry (including nobility) had been educated by domestic tutors, but in the 17th century this was replaced by formal education in grammar schools teaching humanist rhetoric and religious knowledge. Similarly, at the age when they would have formerly gone into service in noble households, it was now fashionable to send young gentlemen to universities and the inns of court. Tutors at Oxford and Cambridge gave instruction in classics, logic, rhetoric, history, theology and modern language, while the inns of court prepared young gentry for local judicial offices. Wrightson (1993: 192–193) claims that the spread of university education unified the minds and manners of gentry, so that by 1660 they formed an independent intelligentsia.

The sort of classical education given at universities was restricted to male gentry and professionals; apprentices only needed to read, write and account, and for countrymen the ability to read and write was useful but not a necessity, anything beyond that being a luxury. Poor people received little or no formal education; for the slightly more prosperous, grammar school was an option, but they would not go on to university unless aiming at a clerical career. Girls, of course, could not go to either grammar school or university; gentry females may well have been taught by domestic tutors, but the rest would only have learned to read and write in petty school, if that. There was thus a hierarchy of education reflecting rank, wealth and gender. (Wrightson 1993: 188–190.)

4. State of the art

From the preceding general overview of the aspects of morphology and historical sociolinguistics relevant to this thesis, it is time to turn to more specific concerns. Section 4.1 presents methods used in and results gained from studies of *-ness* and *-ity* in present-day English (PDE). Section 4.2 moves on to historical studies of productivity, discussing not only *-ness* and *-ity* but also some theoretical and methodological problems arising from a diachronic approach to productivity. For the convenience of the reader, there is a summary of the main points at the end of each section (pp. 41 and 54 below).

4.1. PDE studies

According to Marchand (1969: 312–315, 334–336), both *-ness* and *-ity* form abstract nouns with the meaning 'state, quality, condition of –', although he presents the last two senses in the reverse order for *-ness*. Apparently working on data from the OED as well as anecdotal evidence, he finds that *-ity* is only attached to adjectives, whereas *-ness* can be used with various other bases except for verbs. Furthermore, he says that *-ity* is synchronically only productive with a few Latinate bases such as *-able*, *-ible* and *-ic*; with *-able* bases, there is a further restriction that they must be deverbal adjectives with a passive meaning, otherwise the corresponding noun will be in *-ness*. He also notes that *-ness* is constrained insofar as some bases are more often used by other noun-forming suffixes, such as *-al*, *-an*, *-ar* and *-able* by *-ity*.

Similarly, Aronoff (1976: 36), working on PDE from within the generative framework, suggests that while the general productivity of *-ness* may be higher than that of *-ity* in that *-ness* has fewer morphological constraints operating on it (cf. Section 2.2.5), one should also consider the productivity of the suffixes with particular classes of bases. Thus, by

comparing lists of words in *-ity* and *-ness* obtained from a rhyming dictionary, he finds that *-ness* is more productive (has a longer list) than *-ity* with adjectives ending in *-ive* (e.g., *perceptive*), but that *-ity* is more productive than *-ness* with adjectives ending in *-ile* (*servile*).

Studying the base *Xous* (*monstrous*), Aronoff (1976: 38–43) finds, firstly, that the semantics of the *-ity* derivatives is less coherent than that of the *-ness* derivatives. Secondly, he finds that the phonology of *-ity* is more complex than that of *-ness*: *-ity* shifts the stress of the word to the syllable preceding the suffix, and this syllable is always lax, as in *mendacious*/*mendacity*; furthermore, *-ity* "sometimes triggers the loss of the *ous* which precedes it", as in *simultaneous*/*simultaneity*/**simultaneosity* (Aronoff 1976: 40). This truncation is lexically governed in bases of the class *Xulous* (*bibulous*), which makes the suffix less productive with them, but phonologically conditioned in bases of the class *XVcious*: if the vowel preceding *ci* is *e*, there is no truncation (*speciosity*), but if it is *a* or *o*, the truncation happens (*mordacity*, *precocity*).

Continuing in a similar vein, Anshen and Aronoff (1989) study the productivity of *-ness* and *-ity* on *-ive* and *-ible* bases. Their method of measuring productivity is somewhat different from Aronoff's (1976): in addition to comparing lengths of word lists, they compare the mean frequencies of the bases and the derived words for *-ness* vs. *-ity* in a million-word corpus of present-day American English; furthermore, they use an experiment in which productivity judgments are elicited from subjects. They find that *-ness* is somewhat preferred over *-ity* with *-ive*, and *-ity* is clearly preferred over *-ness* with *-ible*.

Using an 18-million-word subcorpus of the Cobuild corpus (PDE, mostly British English) contained in the CELEX database, Baayen and Lieber (1991: 804–805) first successfully replicate Anshen and Aronoff's (1989) result that *Xivity* has a higher mean token frequency than *Xiveness*,

which implies that *Xivity* is more often lexicalised and thus has a lower productivity (recall the discussion in Section 2.2.2 above). Moreover, they are able to establish that the result is statistically significant in the subcorpus. To do this, they (unlike Anshen and Aronoff) leave out dictionary data from the calculations: "the mixing of frequency data from a corpus with data from a dictionary is highly questionable, since it is entirely unclear on what kind of sample space our probability measure has to be defined" (Baayen and Lieber 1991: 840).

Baayen and Lieber (1991) then turn to the measure called the category-conditioned degree of productivity P (see Section 2.2.2), which they argue is better than mean token frequency for analysing type richness and hence productivity (1991: 815–816). Comparing the value of P for *-ness* with that for simplex nouns, which are on the bottom line of productivity, they conclude that the value for *-ness* is so much higher that *-ness* has to be productive; furthermore, they show that even though the number of types (= the extent of use V) is almost equal for *-ness* and *-ity*, the value of P is larger for *-ness* than for *-ity*, which corresponds to the common intuition that *-ness* is more productive than *-ity* in PDE (1991: 820–822). They do not present a way of establishing the statistical significance of the results, however; nor do they take into account the fact that P is based on the number of tokens of a particular morphological category, and the numbers of tokens for simplex nouns (2,781,258), *-ness* (17,481) and *-ity* (42,252) are so different as to render comparisons between them unreliable at best.

Finally, Baayen and Lieber (1991: 824–826) examine the frequency distributions of certain derivational subdomains of *-ness* and *-ity* using histograms (bar graphs) that show the number of types occurring once, twice, 3–10, 11–100 and 101+ times. They find that for simplex words, *-ness* has larger numbers of types at the low frequencies, which implies that

it is more productive than *-ity* in this subdomain. Only *-ness* can attach to complex words containing native suffixes, while only *-ity* occurs with the Latinate *-ic* and *-al* in their corpus. With *-able*/*-ible*, *-ity* is more productive than *-ness*; with *-ive*, the reverse is true. With *-ous*, *-ity* has more types than *-ness*, but an inspection of the frequency spectrums and P values suggests that it is *-ness* which is more productive in this subdomain. Baayen and Lieber (1991: 826) note that "to some extent the rivals divide up the range of possible bases and show productivity in disjoint segments of this range".

Using the same corpus as Baayen and Lieber (1991), Baayen (1993: 193) calculates the hapax-conditioned degree of productivity P* for *-ness* (77) and *-ity* (29), which again corresponds to the intuition that *-ness* is more productive than *-ity*. Baayen (1993: 202) also gives the activation levels A for *-ness* (791) and *-ity* (337) at θ=8 (see Section 2.2.3); these too point to *-ness* being more productive than *-ity*. According to Baayen (1993: 205), the measures P* and A are especially well suited to ranking productive affixes, while the category-conditioned degree of productivity P is best used for distinguishing between productive and unproductive processes (but recall the problem with comparing P values discussed above).

In an 80-million-word corpus consisting of issues of the *Times* (a British newspaper) from September 1989 to July 1993, Baayen and Renouf (1996) find that the number of new types sampled after 40 months is 0.97 per day for *-ness* and only 0.37 per day for *-ity* — yet another piece of evidence in favour of *-ness* being the more productive of the two in PDE. The number of hapax legomena (= nonnormalised P* value) in the last month of sampling is 739 for *-ness* and 280 for *-ity*; out of these hapaxes, 348 and 143 (respectively) are neologisms in that they are not listed in a comprehensive dictionary. Measured as a function of sampling time, however, the number of neologisms is very low for *-ity*, while *-ness* shows a

higher degree of innovation; the productivity of *-ness* even seems to increase towards the end of the sampling period.

Looking at derivational subdomains of *-ness*, Baayen and Renouf (1996: 83–85) find the largest number of neologisms among derivations from simplex bases (such as *left*), adjectives in *-y* (*crabby*) and compound N + A adjectives (*fashion-conscious*); other fairly common bases include those in *-ish*, *-less*, *un-*, *-ed* and *over-*. By contrast, *-ness* fares poorly with adjectives in *-ing*, which Baayen and Renouf attribute to a semantic mismatch. For *-ity*, the most productive subdomains are bases in *-able* and simplex bases, while the middle ground is occupied by *un-*, *in-*, and *-al*. Affix generalisation, i.e., use with other than adjectival bases, is common with both *-ness* and *-ity* (*next-to-nothingness*, *terrority*), contrary to Marchand's (1969) findings; Baayen and Renouf suggest that this phenomenon has less to do with productivity than with the semantics of the affix in question.

Measured in a way that takes into account the other affixes in the study (Baayen and Renouf 1996: 88–89), *-ness* is unexpectedly productive with adjectives in *-y* and not particularly productive with *-ed*, *-able*, *-al* and *-ing*; out of these, at least *-able* seems to be 'reserved for' *-ity*, which is consistent with Baayen and Lieber's (1991: 826) comment on affix rivalry quoted above. Baayen and Renouf (1996: 83) do, however, find some *-ness* and *-ity* doublets, e.g., *curiousness* / *curiosity*, which indicate a failure in (token-)blocking (see Section 2.2.5). Finally, to find out if their results would generalise to other than newspaper English, Baayen and Renouf (1996: 91) compare them with results obtained from the 18-million-word subcorpus of the Cobuild corpus used by Baayen and Lieber (1991) and find that for *-ness* and *-ity* the results do seem to generalise.

According to Aronoff and Anshen (1998: 244–245), *-ity* seems to be more productive than *-ness* in 20th-century English, which goes against the

results obtained by Baayen and his associates. As Aronoff and Anshen themselves admit, however, this result could be skewed: it is based on entries in the OED, and it may well be that *-ity* words are more likely to be listed in the dictionary simply because they are more unusual and thus more memorable than *-ness* words. Furthermore, as Baayen and Renouf (1996: 69) point out, it would be commercially unappealing for dictionary-makers to print all of the productively formed *-ness* words whose meaning is completely transparent to everybody anyway.

Reminiscent of Van Marle's (1986) domain hypothesis discussed in Section 2.2.5, Aronoff and Anshen (1998: 243–244) claim that *-ness* is the default, qualitatively least restricted case, and *-ity* a special case for forming de-adjectival nouns in English. Unlike Van Marle, however, they say that the default case is not necessarily blocked from domains used by the special cases; for example, *-ness* can attach to most of the bases used by *-ity* except for those in *-ible*. Aronoff and Anshen (1998: 246) further claim that *-ity*, as a special, less productive case, is less predictable when it comes to the meaning of the derived word; therefore, words in *-ity* are easily coined for technical terms with specialised senses, such as *productivity*.

Taking Kastovsky's (1986) two functions of word-formation (discussed in Section 2.2.4) as their starting point, Baayen and Neijt (1997) compare *-ness* with the Dutch suffix *-heid*, which is also a very productive suffix forming abstract nouns from adjectives. While *-heid* is their main object of study, they find that *-ness* is very similar to it in that both can be used in two different functions — in Kastovsky's terms, labelling and syntactic recategorisation.

Baayen and Neijt (1997: 566–567) give examples of words in *-ness* that they claim are used for labelling concepts in English (*business*, *illness*, *consciousness*, *happiness*), and of those that are used for syntactic re-categorisation (*wrongheadedness*, *tenderheartedness*, *stand-offishness*, *dis-*

orderliness). They note that in both English and Dutch, those derivatives that are used for labelling often translate into simplex words or words formed by less productive suffixes in the other language, while the derivatives used for syntactic recategorisation can simply be translated by using the other suffix under consideration. The words used for both functions are formally quite regular, but those used for labelling are more complex semantically; in addition, they are higher in frequency. Conversely, Baayen and Neijt's (1997) data from a Dutch newspaper corpus suggest that the recategorisation function is typical for hapax legomena, which are also more context-dependent than high-frequency words.

Plag, Dalton-Puffer and Baayen (1999) take up a thus far neglected point of view by studying derivation in conjunction with register variation. Their material consists of the 100-million-word *British National Corpus* divided into three subcorpora: written language (90 Mw), context-governed spoken language (6 Mw) and everyday conversations (4 Mw). To be able to compare the measures of productivity obtained from the subcorpora, which vary widely in size, they use vocabulary growth curves estimated through binomial interpolation; these kinds of methods will be discussed further in Section 6.4. Plag, Dalton-Puffer and Baayen (1999: 219) find that the extent of use for both *-ness* and *-ity* varies significantly between the three corpora.

For both *-ness* and *-ity*, the average extent of use V and category-conditioned degree of productivity P are the highest for written language and the lowest for everyday conversations, with context-governed spoken language occupying the middle ground. The average extent of use seems to be a little higher for *-ity* than for *-ness* in all three subcorpora, which in fact corresponds to Aronoff and Anshen's (1998: 244–245) result that *-ity* is more productive than *-ness*, obtained through counting types listed in the OED. However, the average P values seem to be larger for *-ness* than for

-ity, suggesting that *-ness* is more productive than *-ity* in all three registers. (Plag, Dalton-Puffer and Baayen 1999: 222–223, my observations from their figures.)

Plag, Dalton-Puffer and Baayen (1999: 225) explain the higher degree of productivity of abstract nouns in written language by referring to Kastovsky's (1986) functions of word-formation: the recategorisation function is needed more often in written than spoken language, because spoken language also has other means of maintaining reference (e.g., prosody and deixis). Interestingly, they note that while the suffix *-heid* (which is similar to *-ness*, as shown by Baayen and Neijt 1997) seems to thrive in the recategorisation function, it may be that *-ity* is more readily used in the labelling function. This ties into Aronoff and Anshen's (1998: 246) observation that *-ity* is often used to coin technical terms; also note Kastovsky's (1986: 597) remark that "types producing labels exclusively or predominantly are as a rule much less productive".

Investigating affixed words with simplex bases in the same corpus as Baayen and Lieber (1991) and Baayen (1993), Hay and Baayen (2003: 115–119) apply Principal Components Analysis to several interrelated measures of affix behaviour and come up with two dimensions along which the affixes vary: "parsability" and "usefulness". The "parsability" dimension includes high token and type-parsing ratios, high mean base frequency, high productivity (P), low mean derived frequency, and low-probability junctural phonotactics (recall Section 2.2.3). The "usefulness" dimension, on the other hand, consists of a high number of parsed types and tokens, a high number of hapaxes and different types, high entropy, and a low probability of sampling the same word twice.

According to Hay and Baayen (2003: 122), an affix is robustly productive if it is moderately parsable and at least moderately useful; otherwise its status is tenuous. They find that *-ness* is on average

moderately parsable and highly useful, while -*ity* is not parsable but moderately useful (Hay and Baayen 2003: 121). I think this could be taken to mean that the productivity of -*ity*, at least with simplex bases, is both lower than that of -*ness* and more susceptible to change.

In a study of what constrains possible suffix combinations, Hay and Plag (2004) examine the attested combinations of fifteen suffixes in the *British National Corpus*, the CELEX database, the OED and the Internet. Hay and Plag find -*ness* to be among those suffixes that have the fewest selectional and parsing constraints operating on them — in their Figure 1 (2004: 580), it in fact emerges as the least constrained of all in the hierarchy of attested suffix combinations. Even though Hay and Plag (2004) do not study all of the de-adjectival noun-forming suffixes in English, this result seems to lend support to earlier claims by, e.g., Aronoff (1976) and Aronoff & Anshen (1998) that -*ness* could be a sort of default, least constrained case among these, or at least less constrained than -*ity*, with complex bases in general.

Summary

In sum, most scholars seem to agree that -*ness* is more productive than -*ity* in present-day English. This varies by the type of base, however: -*ity* is more productive than -*ness* with bases in -*al*, -*able*, -*ible*, -*ic* and -*ile*, among others, while -*ness* beats -*ity* with bases such as -*ive*, -*ous* and, of course, native bases. The productivity of both suffixes also varies by register: both are more productive in written than spoken language in general, and least productive of all in everyday conversations. The methods used for obtaining these results are increasingly sophisticated and include statistical as well as experimental approaches.

The results from the above studies also indicate that the overall lower productivity of -*ity* is due to factors at several levels of language. The

phonology of *-ity* is more complex than that of *-ness*, and its junctural phonotactics is not sufficiently different from morpheme-internal transitions to facilitate parsing; morphologically, there are fewer bases to which it can attach, as native bases are usually out of the question; semantically, its meaning is less coherent than that of *-ness*; and therefore, it is pragmatically better suited to the labelling function, which is not needed as often as recategorisation, at least not in written language.

4.2. Historical studies

After gaining an overall view of the present-day situation, we go back in time to see how it came about. How does productivity change, and how has the productivity of *-ness* and *-ity* changed? How can we study both the change and snapshots of the situation at certain points in the past? These are some of the questions that the following studies have tried to answer.

According to Marchand (1969: 312–314, 334), while *-ness* is a native suffix, *-ity* entered the language in loans from French during the 14th and 15th centuries — hence its Middle English spelling *-ite*, *-itee*. Marchand also claims that words in *-ity* were first derived "on a Latin basis of coining", i.e., imitating the actual or potential Latin noun in *-itās*. He says this is why the phonology of *-ity* is so complex and why it so rarely attaches to native bases and never to Latin participles such as *absolute*, *complete*.

Marchand (1969: 313–314) claims that words in *-ableness* are usually older than corresponding words in *-ability*, the latter pattern having originated in 16th-century borrowings of *-able*/*-ability* word pairs from Latin; Dalton-Puffer (1996: 107), however, antedates such pairs to Middle English. Marchand also notes that the pool of acceptable bases for *-ability* has widened from Latin-coined adjectives to those in which *-able* has been

combined with a native base; he gives examples of this from the 18th century onwards.

Marchand (1969: 334–336) finds that adjectives of French origin were commonplace as bases for *-ness* by 1300. He also notes that while potential bases for *-ness* have included adjectives and participial adjectives since Old English, in Modern English (i.e., from the year 1500 onwards?) it also accepts other word classes and, mainly since the 19th century, even phrases. The OED (s.v. *-ness*, suffix) adds to this that verbs as bases for *-ness* are rare and date mainly from Old and Middle English, while some nouns occur as bases in all periods. Marchand (1969: 336) observes an interesting difference between words in *-ness* and other abstract nouns: with *-ness*, semantic drift from an abstract state to a concrete instance is much rarer than with the others. This is also noticed by Romaine (1985: 455–456), who says that semantic drift is by contrast very common in *-ity*, by Dalton-Puffer (1996: 84) for Middle English, and by Nevalainen (1999: 398) for Early Modern English.

Examining new words in Early Modern English (EModE, 1500–1700) by means of a 2-percent sample of the OED, Barber (1976: 166–168) finds that roughly twice as many words come about through word-formation as through borrowing; in fact, suffixation alone contributes almost as many words as borrowing. This finding is contradicted by Wermser's (1976: 40, as cited in Nevalainen 1999: 350–351) data from the *Chronological English Dictionary*, which list all of the words in the *Shorter Oxford English Dictionary*: his results point to borrowing and word-formation having an almost equal status. Nevertheless, taking into account the fact discussed above that dictionaries tend not to list all productively formed words, word-formation may have exceeded borrowing in this period.

Barber (1976: 166–184) also finds many more neologisms in the EModE period than the Middle English (ME) period, almost as many as in

the next period, 1700–1900; furthermore, his data show a peak of new words from Latin, French and affixation around 1590–1660, which is in accordance with Wermser's (1976, as cited in Nevalainen 1999: 351–352) observation of a peak in both borrowing and word-formation around the year 1600. Barber notes that these results are not completely reliable, however, because the OED contains different amounts of data from different periods; this is shown clearly by Hoffmann (2004: 25). Nevertheless, the expansion of vocabulary may well have accelerated in the EModE period, as new words were needed for the new functions in which English was used (see Section 3.3.1).

In comparing loans with derived words, Barber (1976), like Marchand (1969), touches upon an important issue that needs to be taken into account when counting *-ity* types for the purpose of assessing productivity: some of them are borrowings and thus have originally little to do with productive formations. For example, while Barber gives *invincibility* as a word formed by suffixation (1976: 186), he classifies *immaturity* as a loanword from Latin (1976: 172), based on the etymologies provided by the OED.

As for productively formed neologisms in *-ness* and *-ity*, Barber (1976: 185–186) finds 70 and 9 of these, respectively; *-ness* is the most productive nominal suffix in his data. The situation thus seems to resemble that in PDE; but recall that Aronoff and Anshen (1998: 244–245) found more words in *-ity* than in *-ness* among 20th-century neologisms in the OED — if *-ness* is underrepresented in the OED for the 20th century, it may be so for the EModE period as well, which would make the difference in productivity between *-ness* and *-ity* much greater in EModE than in PDE.

Nevalainen (1999: 398), however, describes both *-ity* and *-ness* as "very productive" in EModE, and notes (1999: 352) that borrowed suffixes and prefixes constitute an increasing proportion of affixation as a whole: "from some twenty per cent at the beginning of the Early Modern English

period to seventy per cent at the end of it (Wermser 1976: 64)". She further notes that the EModE period was unusually tolerant towards formally related words with the same meaning (1999: 334). Barber does indeed observe some doublets among newly formed words in *-ness* and loanwords in *-ity*: *immatureness* / *immaturity* and *immenseness* / *immensity*.

Romaine (1985) is one of the first to explicitly consider the theoretical and methodological issues in productivity from a diachronic perspective. Even though a diachronic approach to productivity is in her view essential if we wish to understand the synchronic situation, Romaine (1985: 457–458) sees some problems with it. There is no access to informants' intuitions about possible words; furthermore, dictionary listings of actual words from earlier periods are even less comprehensive than listings of present-day words, and the first attestation dates in, e.g., the OED are unreliable. Starting from the assumption that change stems from variation, Romaine hypothesises that competing patterns of word-formation establish themselves through social or stylistic specialisation — in variationist terms, they become markers. She suggests that one method of investigating changes in the productivity of competing patterns would be to examine the available lexical resources in renderings of the same text in different periods of English, i.e., to see how they vary over time within a single topic and register.

Romaine (1985: 458–462) proceeds to present an analysis of the *-ness* and *-ity* formations in King Alfred's, Chaucer's and Queen Elizabeth I's translations of Boethius's *De Consolatione Philosophiae*, along with some other examples. She finds that while Alfred (a pre-conquest writer) uses *-ness* alone, both Chaucer and Elizabeth have a variety of *-ity* forms which they can use instead of *-ness*. With *-ness*, they use an increasing number of so-called **hybrid formations**. In these formations, the etymologically native suffix is combined with a borrowed base, as in the word *fragileness*.

Romaine (1985: 451) suggests that the most important extralinguistic factor in the increase in the productivity of *-ness* is register. In Middle English, neologisms in *-ness* and especially new hybrid formations are most prominent in the new literate registers, such as religious and philosophical texts (Romaine 1985: 464). These are also where *-ity* is introduced into the language.

Romaine (1985: 461–462) shows that, although some Old English words and affixes were lost, educated writers from the 14th or 15th century onwards "could draw on four distinct sub-systems of derivational morphology": native base + native suffix, foreign base + foreign suffix, and hybrid formations with native base + foreign suffix or foreign base + native suffix. Education was key here: those without a classical education were unable to use the foreign stock until much later, which led to the diglossic situation discussed in Section 3.3.1. In other words, borrowed patterns diffused through the language at different rates, depending on social and stylistic factors (Romaine 1985: 464).

Romaine (1985: 452) also gives some details on the early history of *-ness*, which is related to modern German *-nis* and Gothic *-assus*/*-inassus*. Apparently, it used to attach predominantly to *n*-stem verbs, from which it acquired the /n/. In Old English, it formed abstract feminine nouns, competing with at least *-hood*, *-dom* and *-ship*; *-ness* became the most productive of these, with *-ship* taking second place. Riddle (1985: 450–451) notes that *-ness* words, by Old English mostly based on adjectives, also began to compete with simplex nouns formed from adjectives through conversion. Then, just as *-ness* was becoming the strongest of the native processes forming abstract nouns, *-ity* arrived.

Most of the studies introduced above assume that even if the derivatives of *-ness* and *-ity* may vary in semantic coherence or undergo semantic drift through lexicalisation, the suffixes themselves are in

principle synonymous. Riddle (1985), however, argues that this is not quite true in PDE, and offers a historical explanation. She claims that "-*ness* tends to denote an embodied attribute or trait, while -*ity* tends to denote an abstract or concrete entity" (Riddle 1985: 437). Thus, she seems to anticipate Kastovsky's (1986) functions of word-formation, i.e., syntactic recategorisation and labelling, which have been shown to be typical of -*ness* and -*ity*, respectively (see Section 4.1). However, while Kastovsky sees the distinction as pragmatic and context-dependent (1986: 595–596), for Riddle it is a semantic one.

Furthermore, Riddle (1985: 444) suggests that the fact that -*ness* and -*ity* favour different bases may be due to the semantics of the base, which needs to be compatible with the semantics of the suffix used. This could explain the PDE pattern discovered by Marchand (1969: 313–314) that -*ity* only attaches to those -*able* bases that have a passive meaning (e.g., *adaptable*, *comparable*); these seem to go well with the objective 'entity' meaning of -*ity*, while the other -*able* bases (e.g., *charitable*, *conscionable*) seem to match the subjective 'embodied trait' meaning of -*ness*. According to Riddle (1985: 445–446), other factors influencing suffix choice and productivity include blocking, phonological transparency, lack of need for a word with a particular meaning (cf. Kastovsky 1986: 597–598), the historical factor of borrowed -*ity* words becoming fixed in usage early on and thus less prone to be replaced by -*ness*, and the greater social prestige of the Latinate -*ity*.

As for the historical explanation for why most but not all words in -*ness* vs. -*ity* are semantically distinct in PDE, Riddle (1985: 446–449) claims that this is because of an ongoing change spreading through lexical diffusion. Many older words in -*ness* that denoted an entity have been replaced by other words, have lost the entity meaning, or the meaning has changed to denote a trait. Conversely, when -*ity* appeared, it could initially

also denote a trait, but many of these words have now been replaced or lost, or the trait meaning has changed or been lost.

Riddle (1985: 451–455) suggests that the change could be due to the wholesale borrowing into Middle English of French and Latin theological terms with entity senses. This replacement of Old English religious words would have affected the language of the common people as well as the upper ranks, which would explain why the borrowing did not lead to only register variation. The change was later reinforced by entity loans in the field of scholarship, which were disseminated to common people by means of increased education after the reformation. Thus, in spite of the diglossia argued for, e.g., by Romaine (1985), Riddle claims that there was enough contact between *-ness* and *-ity* words for the semantic divergence to continue. This claim is supported by Adamson (1989: 214), who says that the 'high' and 'low' variety synonyms in English have indeed developed a systematic difference in meaning: "The H forms have connotations of conceptual clarity and emotional neutrality, while the L forms are associated with physical reality and subjective response."

Methodology-wise, Riddle (1985: 457) criticises the sort of experiments and analyses in which the possibility of a semantic distinction is not taken into account. Her own methods include classifying *-ness* and *-ity* words in dictionaries into different meaning categories and comparing the numbers or percentages of words belonging to each category (1985: 452–453).

Besides the problem of distinguishing between borrowings and actual derivatives already noted by Marchand (1969), Anshen and Aronoff (1989), in the diachronic portion of their study, make explicit a further problem with type counts as a measure of productivity: they do not distinguish between current and past productivity. A word that was coined hundreds of years ago is given equal weight as a type as a formation that is

genuinely new to the period under study. Anshen and Aronoff's (1989: 199–200) solution to this problem is to check their words' first attestation dates in the OED and count the number of new types for each century (although, as noted by Romaine (1985: 457), the OED is not a completely reliable source for dates). Thus, they are able to investigate diachronic changes in the productivity of *-ness* and *-ity* with *-ive* and *-ible* bases.

Anshen and Aronoff (1989: 200–201) find that the ratio between *-iveness* and *-ivity* formations has remained roughly the same since 1600, with *-ness* in the lead. With *-ibleness* and *-ibility*, however, *-ity* has gained more and more ground from *-ness*; Anshen and Aronoff do not attempt to explain why this should be the case. The overall number of nouns based on *-ible* decreases in the 19th century, suggesting a general loss of productivity with that base. As their data come from a rhyming dictionary, whose sample space is unknown, Anshen and Aronoff's results should be taken with a grain of salt.

The next study, by contrast, is an early example of historical corpus linguistics. Dalton-Puffer (1992: 477–478) investigates the productivity of four deverbal suffixes in Middle English and finds that their ranking is the same whether measured on a scale of morphotactic transparency or by the number of hybrid formations found in the Middle English part (0.5M words) of the *Helsinki Corpus of English Texts*. Among the Romance suffixes, she finds no hybrid formations in *-acioun* or *-aunce*, and only a couple in the most transparent of them, the consonant-initial *-ment*; in contrast, over 18% of the types in the native *-ung* are hybrid. I think this qualifies Romaine's (1985: 461–462) claim about educated writers having four systems of derivational morphology at their disposal: even for them, the systems may have been available in different degrees, with foreign base + native suffix the least accessible case. Dalton-Puffer (1992: 475) also analyses deverbal *-ness* by comparing the number of its instances between

the three roughly same-sized Middle English periods in the corpus (M1–M3); she finds it to be on a strong decline.

Using type counts from the same material as in the above study, Dalton-Puffer (1996: 81–85) finds *-ness* to be steadily and highly productive in Middle English, with 220 types in all, and among them an increasing number of hybrid formations. She speculates that the decrease in deverbal *-ness* could be due to its complex semantics and to the possibility that it was originally used in loan translations from Latin, which went out of fashion by Early Middle English. Interestingly, she calls these calques "labels for new concepts" and notes that they were replaced by actual loanwords; I think their demise could be related to the overall decline in the entity meaning (or in the labelling function) of *-ness* noticed by Riddle (1985).

Dalton-Puffer (1996: 106) finds that the number of different words in *-ity*, 74 altogether, increases rapidly in the third ME period (1350–1420) in the corpus, which she says is typical of the borrowed suffixes. This makes sense in the light of Culpeper and Clapham's (1996: 215) OED-based result that the peak period in loanwords from French was just before that, 1251–1375. Based on the low number of 'native + foreign' hybrid formations in the *Helsinki Corpus*, Dalton-Puffer nevertheless suggests that Romance suffixes, including *-ity*, were not productive in Middle English (1996: 220–222). She accounts for the few existing hybrids (*scantetee* for *-ity*) by assuming that they were created through analogy; however, as Dalton-Puffer herself admits, there are no strict boundaries between analogy and rule-based productivity, and hybrid formations are not the best or only measure of productivity.

On Dalton-Puffer's scale of morphotactic transparency, adapted from Natural Morphology, *-ity* occupies the middle ground, which predicts that it has a fair chance of survival or productivity in later stages of the language;

-ness, of course, is at the top of the scale, and is also the most frequent suffix forming abstract nouns (1996: 120, 128). Examining *-ness* and *-ity* doublets, the majority of which appear in the last period, Dalton-Puffer (1996: 127–130) finds no systematic meaning differences between them but admits that she has too little data to ascertain that they really are synonymous. She estimates that the divergence in meaning described by Riddle (1985) would not be observable until after 1420, which is the cut-off point in Dalton-Puffer's own data.

In a comprehensive literature survey of Early Modern English lexis and semantics, Nevalainen (1999: 357) writes that Early Modern English saw an increase in 'native + foreign' hybrid formations, which implies that the borrowed affixes were by now productive. For *-ity*, hybrid formations were rare; according to Nevalainen (1999: 398), it was mostly productive with bases such as *-able*/*-ible*, *-ic*, *-al* and *-ar* from the 15th century onwards. As for *-ness*, Nevalainen (1999: 398) reports that it prefers native bases but is not limited to them. Nevalainen (1999: 357) says that by the end of the period the productivity of the borrowed suffixes seems to have decreased again, as their use was beginning to be restricted to technical terms.

Cowie and Dalton-Puffer (2002) continue Romaine's (1985) and Anshen and Aronoff's (1989) work on theoretical and methodological issues in diachronic word-formation. They tentatively accept the statement that productivity can be measured as the number of new words between two points in time (2002: 418). After questioning the notion of a new word (cf. Section 1.2), they problematise the concept "a point in time" (2002: 420–421): while Plag (1999: 101) considers his period of 1900–1985 as "small enough to exclude major diachronic developments", Baayen and Renouf (1996: 89–92) find change over their considerably shorter period of 1989–1992. In historical corpora, the periods that count as points in time

typically need to be much longer than that, to ensure that there is enough data. Thus, Cowie and Dalton-Puffer (2002: 421) conclude that the distinction between a synchronic point in time and a diachronic stretch of time is "a matter of definition and methodological necessities".

Examining dictionary-based methods of studying changes in productivity, Cowie and Dalton-Puffer (2002: 423–424) note that in addition to comparing type counts based on their first attestation dates in a dictionary, they could be compared between dictionaries made at different times, or between different editions of the same dictionary. However, they also note the drawbacks of dictionaries, including the fact that the coverage of the OED varies widely between different periods in terms of both the number of citations and text types, which makes diachronic comparisons unreliable.

Moving from dictionaries to corpora, Cowie and Dalton-Puffer (2002: 424–425) suggest that the main problem with historical corpora is their small size but that their advantages of providing access to contextual and stylistic factors outweigh this limitation. On comparing type counts between (sub)corpora, they note (2002: 425–428) that sometimes, if the corpora are of approximately the same size, the difference in the number of types may be large enough to be noticeable in the raw figures. The trouble begins when the corpora vary in size: type counts cannot legitimately be normalised, as the number of types does not grow linearly with the size of the corpus.

Even if the type counts were comparable, there is a further problem in determining what it means if type frequency increases over time. It does not straightforwardly imply productivity, as the types in small corpora are not necessarily aggregated, i.e., the types of a productive affix in a later period may not consist of all of the types in the earlier periods plus new types. Baayen and Renouf's (1996) method of only counting the new types

in each period will only work if the corpus is large enough so that it does not take much sampling time to encounter all the types from the centuries before the time period represented by the corpus and thus to make the pattern within the universe of the corpus observable. Therefore, Cowie and Dalton-Puffer (2002: 430–431) suggest using a starting lexicon to eliminate part of the non-new types; nevertheless, it remains unclear how much each subperiod is affected by the remaining non-new types, beyond the fact that the effect decreases over time.

Finally, Cowie and Dalton-Puffer (2002: 431–432) assess the potential of using Baayen and Renouf's (1996) method of counting hapax legomena to track changes in productivity. They conclude that while it may be possible in large corpora, it will not work in the small historical corpora currently available, as hapaxes only catch real neologisms in large corpora. I would say that if a corpus, even a small one, is a representative enough sample of the (kind of) language under study, it could be possible to use hapax counts as an indicator of productivity even if none of the hapaxes is a real neologism. A similar stance is taken by Baayen (1993: 187), who defends the reliability of hapax-based productivity measures obtained from a corpus of only 600,000 running words.

Two recent MA theses which explore the morphological productivity of affixes in the history of English merit a mention here. Showing that methods developed for synchronic studies can also be applied to diachronic ones, Hlava (2005) uses Baayen's measures P and P* in addition to type and token counts and hybrid formations to investigate the productivity of five negative prefixes in the CEECS. Keeping in mind Baayen's (1993: 187) caveat that the P values cannot be relied upon in the case of affixes represented by an extremely small number of tokens, she is nevertheless able to extract some very interesting results even from this sampler corpus of 450,085 words.

Březina (2005) carries out both dictionary- and corpus-based research in a study of the prefixes *in-* and *un-* in Early Modern English. He notes that the approaches complement each other: the dictionary (the OED) provides significantly more instances than the corpus, while the corpus (the CEECS) also provides information on actual usage, such as frequencies, in a speech-like text type (2005: 158–159). However, what makes Březina's work remarkable is that he combines derivational morphology with historical sociolinguistics by investigating gender variation in the use of the prefixes in the CEECS. Apparently basing his conclusion on token counts only, Březina (2005: 147) finds that "women's use of the *un-* and *in-* forms is more progressive than men's use". An examination of the frequency of each type with men and women, however, reveals that "women repeatedly use a more limited number of *in-* forms" and that 78% of women's total use of *in-* comes from their 20 most frequently used types, while the men's figure is 63% (Březina 2005: 151).

Summary

The main results from the above studies concerning the productivity of *-ness* and *-ity* throughout the centuries can be summarised as follows. The native suffix *-ness* dates from before the Old English period; originally, it was mainly used with verbs, but since Old English it has most frequently been based on adjectives. Noun bases are infrequent but occur in all periods of English, and other parts of speech excepting verbs seem to have become increasingly possible as bases, which may indicate an increase in productivity, or possibly a change in the semantics of the suffix (cf. Baayen and Renouf 1996: 84–85). Competition between *-ness* and other native processes was resolved in favour of *-ness*, which became the most productive of the deadjectival suffixes forming abstract nouns by the end of the Old English period.

In Middle English, *-ity* entered the language through loanwords from French and later also through calques on Latin; while it did not reach the level of productivity of *-ness*, it began to compete with *-ness* as a more learned deadjectival suffix meaning approximately the same thing. The productivity of *-ness* also increased during this time, mostly in the new literate registers. By Early Modern English, both *-ness* and *-ity* were very productive, as new words were needed to facilitate the use of English in an ever-widening variety of functions; however, by the end of the EModE period, the productivity of *-ity* had decreased, its use being now limited to mostly technical terms. The functional/semantic differentiation between *-ity* and *-ness* is a process continuing to this day: while *-ity* is mainly used for labelling abstract or concrete entities, *-ness* is used more for syntactic recategorisation with the meaning 'embodied trait'.

As for variation in the productivity of *-ness* and *-ity* among different classes of adjectival bases, *-ness* has been commonly attached to adjectives of French origin since 1300, although native bases were still preferred in at least Early Modern English; by contrast, hybrid formations in *-ity* seem to always have been rare. While words in *-ableness* seem to have appeared earlier than those in *-ability*, the association of *-ity* with *-able* may have arisen already in Middle English through *-able*/*-ability* word pairs borrowed from Latin; nowadays, *-ability* is by far the more common choice, provided that the semantics of the *-able* word matches that of *-ity*. Similarly, *-ibility* has become more and more common than *-ibleness*, whereas *-iveness* seems to have been consistently somewhat favoured over *-ivity*. Other bases with which *-ity* became productive from the 15th century onwards include *-ic*, *-al* and *-ar*.

Methodologically, the historical studies overviewed above are in general less sophisticated than the PDE studies, partly because there has been too little data for proper statistical analysis. While psycholinguistic

experimentation has been impossible for obvious reasons, scholars have developed complementary methods such as counting hybrid formations or using first attestation dates from dictionaries. Here as in PDE studies, corpus linguistics is a welcome addition to the earlier dictionary-based approach, as it provides information on actual usage.

The diachronic perspective highlights some of the problems with using type counts as a measure of productivity. Firstly, the existence of a large number of types may be due to aggregation through productivity in the past rather than current productivity. Secondly, in the case of *-ity*, some words have been borrowed from French or Latin as a package including the suffix, with no productivity involved at all in English. Thirdly, an increase in the number of types between one time period and the next in a corpus does not straightforwardly mean an increase in productivity, because the types are not necessarily aggregated. Fourthly, comparing type counts obtained from (sub)corpora of varying sizes is difficult, as normalisation is not a legitimate option. Another problematic concept is that of "a point in time".

5. Research question

If we accept Romaine's (1985: 457) variationist hypothesis that competing patterns of word-formation establish themselves through social or stylistic specialisation, it becomes obvious that something crucial is missing from previous research on the productivity of *-ness* and *-ity*. While register variation has been studied to some extent by, e.g., Plag, Dalton-Puffer and Baayen (1999), Riddle (1985) and Romaine herself, the sociolinguistic point of view has so far been left unexplored. Sociolinguistic variation in word-formation in general is something which has only very recently begun to be studied, one of the few examples being Březina's (2005) diachronic study described above, and Keune, van Hout and Baayen (2006)

with their synchronic study of spoken Dutch. Nevertheless, it seems clear to me that if we wish to understand how productivity changes, we must not ignore extralinguistic factors. Therefore, this thesis attempts to answer the following research question: Does the productivity of *-ness* and *-ity* in 17th-century English letters vary between different social groups, and if so, how?

My hypothesis is that the productivity of *-ity*, a learned and etymologically foreign suffix, is both lower and more variable than the productivity of *-ness* in this material. More specifically, I expect *-ity* to be less productive with poorly educated social groups, such as women and the lower ranks, than with well-educated groups, such as men and the higher ranks. As to the productivity of *-ness*, I do not expect to find significant differences between social groups.

The choice of material, 17th-century correspondence, is motivated by several considerations. Firstly, keeping the register constant will facilitate observing the sociolinguistic variation within it (cf. Romaine 1985: 458). Secondly, while previous work has been limited to literate registers with the exception of Plag, Dalton-Puffer and Baayen's (1999) PDE study, investigating a speech-like register such as letters is of interest to historical linguistics, as speech is the primary medium of language and, according to Nevalainen and Raumolin-Brunberg (2003: 28), the most fertile ground for linguistic change. Thirdly, the time period is interesting because it is to be expected that *-ity* would by this time have spread to wider use from the more literate registers in which it entered the language (cf. Riddle 1985: 455–456). Fourthly, my pilot study (Säily 2005) using the full time span provided by the CEECS (1418–1680) showed a gender difference in the use of *-ity* in the letters of the 17th century.

A further justification for this study is offered by the fact that it proposes to improve on the methods used for measuring and comparing

productivity in previous research, especially on the historical side. As noted above, comparing productivity measures obtained from different (sub)-corpora is problematic. Previous work often relies on the corpora being approximately the same size, so that for instance type counts obtained from each corpus can be compared directly. Then, if the type counts differ by an order of magnitude, it may be possible to draw conclusions without paying attention to statistical significance (e.g., Dalton-Puffer 1996: 106). This is usually the approach taken in dictionary-based diachronic studies as well (e.g., Barber 1976: 166–195): for example, different periods in the OED are assumed to contain comparable amounts of data, even though this is untrue, as noted in the previous section. Some scholars (e.g., Riddle 1985: 452–453; Anshen and Aronoff 1989: 200–201; Dalton-Puffer 1992: 478) have compared percentages of the types of a given affix out of all words, or of similar measures, in different periods; this, too, is questionable considering that the number of types may grow at different rates for different processes, and there is still no way of establishing the statistical significance of the results.

The category-conditioned degree of productivity P, developed by Baayen and his colleagues, is essentially an approximation of the growth rate of the number of types with a given affix in a corpus (Baayen 1992: 115). Based on permutation testing, the little-known method used in this thesis makes it possible to examine the actual growth curve of the affix, providing additional information to P, the approximation of the tangent at the end of the curve. More importantly, it facilitates both comparing data obtained from corpora of varying sizes and establishing the statistical significance of the results. Similar methods have been used before, mainly in studies of lexical richness, but mostly with the purpose of verifying results from more sophisticated techniques involving inter- or extrapolation (e.g., Tweedie and Baayen 1998); previous work does not seem to have

recognised the robustness of this method on its own. Both the methodology and the corpus used in the study will be discussed further in the next section.

6. Material and methods

6.1. The corpus

The material used in this study comes from the 17th-century part of the 2.7-million-word *Corpus of Early English Correspondence* (1998 version). The CEEC is an electronic collection of 6,039 letters composed by 778 writers between the years 1410?–1681. It was compiled by Terttu Nevalainen (team leader), Jukka Keränen, Minna Nevala (née Aunio), Arja Nurmi, Minna Palander-Collin and Helena Raumolin-Brunberg. Due to lack of resources for transcribing and editing, the corpus is based on published editions of letters; however, some of the material has been checked against the originals by members of the CEEC team.

The CEEC is designed for studying the English language — more specifically, English English — in its sociohistorical context. To this end, the writers have been carefully selected to give as balanced a representation of different social categories as possible. Nevertheless, the dominance of men from the upper ranks has been unavoidable: they were the most literate group, they were considered important enough that their letters were preserved, and their letters were later considered important enough to be published. The 17th-century part of the CEEC consists of 1.4 million words covering the years 1600–1681. Unfortunately, only about a quarter of this material was written by women, as can be seen from Figure 2 below. The situation between different ranks, regions, etc. is similarly imbalanced.

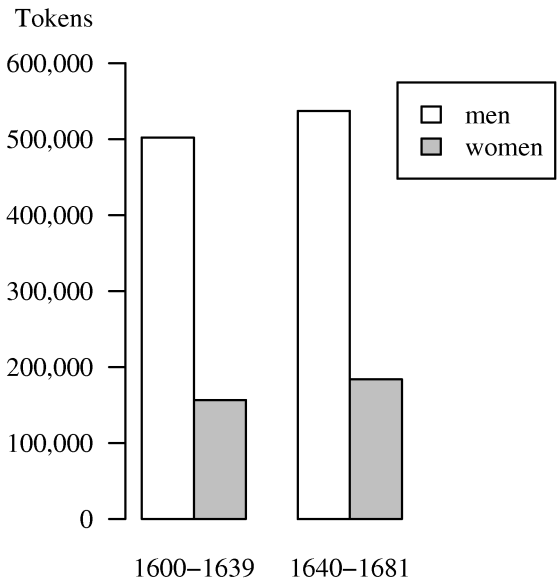

Figure 2. Running words written by men vs. women in the 17th-century part of the CEEC.

6.2. Data collection

The principles used in determining what counted as a *-ness* or *-ity* token were initially kept very accommodating; as noted by Plag (1999: 29), dropping out "non-productive formations" could mean prejudging the issue of whether the suffix is productive. Thus, prefixed forms such as *infallibility* and *undutifulness* were counted even though they could have been formed by adding the prefix to the noun rather than adding the suffix to the prefixed adjective. Furthermore, anything that etymologically contained either suffix was counted, whether or not it had an extant base at the time when the material was written.

For *-ity*, this meant taking in loanwords — not only those such as *generosity* which were originally borrowed but which had a possible base in English by the 17th century, but also words like *sagacity*, which could not have been formed in English unless one applied the truncation rule described by Aronoff (1976: 40) to the adjective (in this case *sagacious*), and even words like *quality*, for which there did not exist even a remotely

possible base in English, unless one wished to postulate a bound root (cf. Dalton-Puffer 1996: 107). Such words as *city*, which due to their shortness could not have been associated with the suffix in any way even though the ultimate Latin etymology was the same, were nevertheless left out.

Combinations of a base ending in a consonant + *-ty* such as *admiralty* were not counted as instances of *-ity*, which seems to have been the practice in most of the previous research with the exception of Romaine (1985). However, combinations of a base ending in *e* or *i*/*y* + *-ty* such as *nicety* and *jollity* were at this stage included in the data, even though they have traditionally been classified under the suffix or loan termination *-ty* (Marchand 1969: 315). This is because I considered it possible that psycho-linguistically these might have been associated with *-ity*, given that both the form and the meaning of the ending more or less matched it.

The instances of *-ness* and *-ity* were extracted from the corpus files using the WordCruncher program. Since the corpus was neither lemmatised nor grammatically tagged, this had to be done by searching for all word-forms which had a suitable ending. Different spelling variants of the suffixes were collected from the OED, the *Middle English Dictionary* (MED) and by browsing the corpus itself, after which they were used one by one in WordCruncher searches. Some of these variants, such as *-nes*, yielded a vast number of erroneous results, because many other words besides those having the suffix ended in that way, such as plurals of words ending in *-n*. These had to be weeded out by hand.

The results were printed from WordCruncher into text files containing each instance of the suffixes embedded in six lines of context, as in example (7) below (emphasis added). The last line, added automatically by WordCruncher, serves to identify the collection, year, author, page number and section of the letter in which the instance was found. As the letters in the corpus have not been divided into sections that the program could

understand, the section information is always the same, "Heading". The code for the author consists of the first initial and the last name of the writer, with an extra character added after the initial if there are several people in the corpus with the same initial and last name. In (7), the writer is Thomas Meautys I, the number indicating that there is also a later Thomas Meautys in the corpus. The collection code, COR, indicates that the letter comes from the Cornwallis collection, which was compiled from an edition of the private correspondence of Jane, Lady Cornwallis. The letter collections included in the CEEC are listed in Nevalainen and Raumolin-Brunberg (2003: 223–234).

(7) departure let them endevor to lyve soe as they may dye the servants of Almightie God. I have often called to minde a sayinge of you unto me, which for the *pyousnes* of it I must never forget, it being upon the death of your fyrst husband, when myselfe was with you and saw how exceedingly you greeved for the los of him; and I well remember that I was a lyttle (COR 1627 T1MEAUTYS 181:Heading)

I had already gone through the 17th-century letters in the *Corpus of Early English Correspondence Sampler* (CEECS) for my pilot study (Säily 2005). As the letters in the sampler are a subset of those in the full corpus, CEEC, I used my earlier results done on the CEECS files and augmented them where necessary with files from the CEEC. There was some overlap, as the files I used from the full corpus were person-specific, and there were some cases where there were more letters from a person in the full corpus than in the sampler. In cases of overlap, I used the results from the CEECS files — the texts are identical in both corpora, but the identifying line added by WordCruncher is slightly different.

Example (7) above shows the identifying line for CEECS. For CEEC, the abbreviation for the collection is replaced by a code showing the degree of authenticity of the letter: A for autograph, B for autograph by a person of whom little is known, C for copy or secretarial letter, D for unknown

autograph status. Furthermore, after the year comes an extra code for the relationship between the sender and the recipient: FN for nuclear family, FO for other family, FS for family servant, TC for close friends, T for other acquaintances. For the text in example (7), the CEEC line would be as follows.

(8) (A 1627 FN T1MEAUTYS 181:Heading)

6.3. Data processing

A combination of manual work and Perl scripts, the latter kindly provided by researcher Jukka Suomela of the Helsinki Institute for Information Technology HIIT, was used to create an Excel-readable text file containing the following information for each instance: name of source file, line number where the instance was found in the file, year of writing, author code, page number & section, the orthographic form of the token as found in the corpus, and the token in context. With the help of the OED, this was then augmented by adding after the orthographic form the lemma under which it belonged, the etymology of the base, and first attestation dates for both the base and the suffixed word, as in (9) below (shown on two lines due to lack of horizontal space).

(9)

data-ceecs/ness/nes2.txt	766	1627	T1MEAUTYS	181:Heading	pyousnes	piousness
AN/F/L	c1450	1623	departure let them endevor to lyve soe as they may dye the servants of Almightie God. I have often called to minde a sayinge of you unto me, which for the ***pyousnes*** of it I must never forget, it being upon the death of your fyrst husband, when myselfe was with you and saw how exceedingly you greeved for the los of him; and I well remember that I was a lyttle			

The lemmas were later used when counting types, which meant that the nominative and genitive singular and plural forms of a lexeme were all subsumed under one type (cf. Plag, Dalton-Puffer and Baayen 1999: 214).

The etymology in (9), AN/F/L, means Anglo-Norman/French/Latin. The first attestation date was always taken to be the earliest listed for the word in the OED, regardless of whether the sense under which it appeared matched the sense of the instance, as it would have been too time-consuming and complicated to analyse the sense of each instance. The letter c before the first attestation date of the base means "circa".

For a closer look at the types, another Excel file was prepared with the following information: number of instances, lemma, base, base related (mainly for words in *-ity* which did not have a straightforward base but there was a related adjective), base part of speech, base etymology, base first attestation, lemma first attestation. To take an example from *-ity*:

(10)

Instances	**Lemma**	**Base**	**Base Related**	**Base POS**	**Base Etymology**	**Base 1st**	**Lemma 1st**
2	incongruity	(incongrue)	incongruous	(a)	(F/L)	(1398)	1532a

The parentheses indicate that the base in question was marked in the OED as rare or obsolete by the 17th century. The part of speech, a, means adjective; the letter a after the first attestation date of the lemma means that the citation is from "ante", or before, the year 1532. This notation comes from the OED with the exception that the letter has been placed after the year rather than before it, to facilitate sorting in Excel.

For the purposes of the method for comparing type counts between subcorpora, the corpus was virtually divided into samples, each consisting of one person's letters from a 20-year period in the corpus: 1600–1619, 1620–1639, 1640–1659, and 1660–1681. The total number of samples in the corpus was 412, of which 300 consisted of letters written by men. As an example, all letters in the corpus that were written by Thomas Meautys I in 1620–1639 formed a sample called T1MEAUTYS-1620.

The division was virtual in that the actual text of the corpus was not used at all; rather, the already collected instances of *-ness* and *-ity* were grouped by person and 20-year period, and this information was recorded in a text file in the form of two incidence matrices, one for *-ness* and the other for *-ity*, along with the word count of each sample. Table 2 shows part of the incidence matrix for *-ity*. Each row corresponds to one sample and each column to one type; the numbers indicate how many tokens there are of each type in the samples. Thus, the sum of the elements on a row equals the number of *-ity* tokens in that sample, and the number of nonzero elements on the row equals the number of *-ity* types in the sample.

	ability	…	***incivility***	***incommodity***	***incongruity***	…	***sincerity***	…
…								
ASTUART-1600	0		1	0	2		2	
HOXINDEN-1640	5		1	0	0		2	
T1MEAUTYS-1620	1		0	0	0		1	
…								

Table 2. Part of the matrix representation of *-ity*.

The word counts were obtained from a database of sociolinguistic information on the writers in the CEEC, prepared by Arja Nurmi. For example, here is the information on Thomas Meautys I for the period 1620–1639, and on Arabella Stuart, the author of the two instances of *incongruity* in the corpus, for 1600–1619:

Author Code	**Gen-der**		**Domicile**	**Rank**	**Birth-date**	**Name**	**Social Mob.**	**Word Count**	**Period**
T1MEAUTYS	M	D	O	Gentry Upper/ Profes-sional		Meautys Thomas I		4,547	1620–1639
ASTUART	F	A	C	Nobility	1575	Stuart Arabella		30,473	1600–1619

Table 3. Two entries in the auxiliary database to the CEEC.

The unnamed third column in Table 3 is for the degree of authenticity of the letters, explained in 6.2 above. The domicile codes are C for the court, L for London, H for the home counties surrounding London, F for East

Anglia (both Norfolk and Suffolk), N for the north (all counties north of Lincolnshire), and O for other (cf. Nevalainen and Raumolin-Brunberg 2003: 38). The rank classifications follow the most fine-grained division presented in 3.2.2 above (Nevalainen and Raumolin-Brunberg 2003: 136, Model 1). If the person was a social aspirer, there would be a "Yes" in the column for social mobility. Out of the social parameters, only gender and birth date are objective facts, while the main domicile, status and social mobility of the writer are matters of interpretation (Nevalainen and Raumolin-Brunberg 1996: 53). Nevertheless, these were all used to form subcorpora whose type counts were compared with those from the entire 17th-century corpus.

6.4. Comparing type counts

Because the subcorpora based on social categories varied greatly in size, some sort of method was needed for comparing the productivity measures obtained from them. Normalising type counts or similar measures would not work: the number of types grows at a different rate from the number of running words in the corpus (recall Figure 1 on p. 15 above), so simply dividing the number of types by the number of running words to obtain a figure for types per, say, 100,000 running words, would not be justifiable. Taking a random sample of equal size from each subcorpus, as recommended by, e.g., Baayen and Lieber (1991: 840), would mean a loss of data from every subcorpus except the smallest one, which is especially undesirable in lexical and derivational studies, where the frequencies involved are typically small. Furthermore, there would be no way of ensuring the representativeness of the samples, and no obvious way of establishing the statistical significance of the results.

Parametric methods for estimating vocabulary growth have been used for comparisons between subcorpora (e.g., Plag, Dalton-Puffer and Baayen

1999), but they are not completely reliable; see Evert and Baroni (2005) for an evaluation. Furthermore, these methods entail simplifying assumptions about the text: for example, one common assumption is that the occurrences of the words are independent, which is strictly speaking untrue, as the choice of words may have subtle dependencies on the cotext. The method used in this thesis addresses all of the issues outlined above. It is non-parametric and assumption-free, and provides access to the actual growth curve instead of an estimation; it is based on random sampling, but there is no loss of data involved, and it provides a straightforward way of establishing statistical significance.

The basic idea behind the method is this. Instead of comparing the results from each subcorpus with each other, each result is compared with the growth curve of the suffix in the full corpus to see if it the result is significantly different from the corpus as a whole. The growth curve, also known as the type accumulation curve, is constructed as follows. Divide the corpus into small samples; as noted in the previous section, my samples consist of one person's letters from a 20-year period in the corpus. Randomly pick one sample and calculate the number of types in it, plotting the result on a graph with the number of running words in the sample on the *x* axis and the number of types on the *y* axis. Pick another sample and add it to the last one, calculate the combined length of the samples and the overall number of different types in them, again plotting the result on the graph. Pick a third sample and add it to the previous samples, again calculating the combined length of the samples and the overall number of different types in them and plotting the result on the graph. Repeat until the entire corpus has been sampled. This is how it looks:

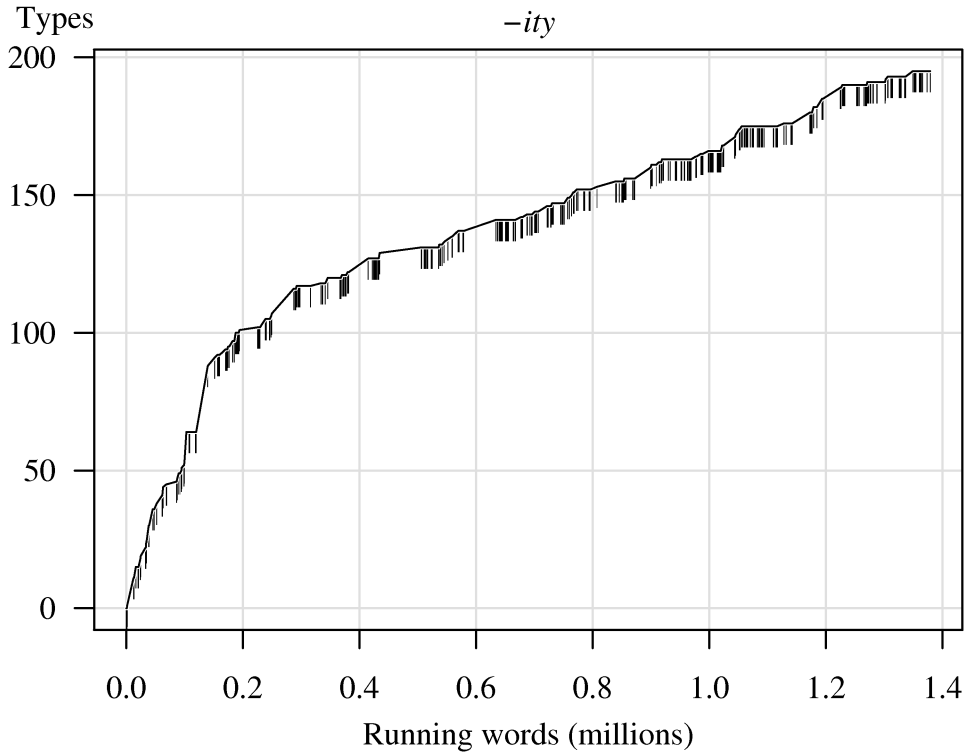

Figure 3. A type accumulation curve.

However, this is not all. To make the most of the data, this process of calculating type accumulation curves for random permutations of the corpus is repeated several times — in the present work, one million times. Of course, this is not done manually but by a computer program, for which I am indebted to researcher Jukka Suomela of the Helsinki Institute for Information Technology HIIT (see Suomela 2007). Next, all of the accumulation curves are combined into a single graph such as Figure 4 below. Lower bounds are drawn for areas containing 90%, 99%, 99.9%, 99.99%, and 99.999% of the curves; these correspond to the levels of statistical significance p = 0.1, 0.01, 0.001, and 0.0001. The same is then done for upper bounds. Now it is possible to simply plot the results from each subcorpus on the graph and see where they go. If a point falls, say, below the area containing 99.9% of the type accumulation curves for the random permutations of the corpus, this means that it is significantly different from the overall corpus at $p < 0.001$.

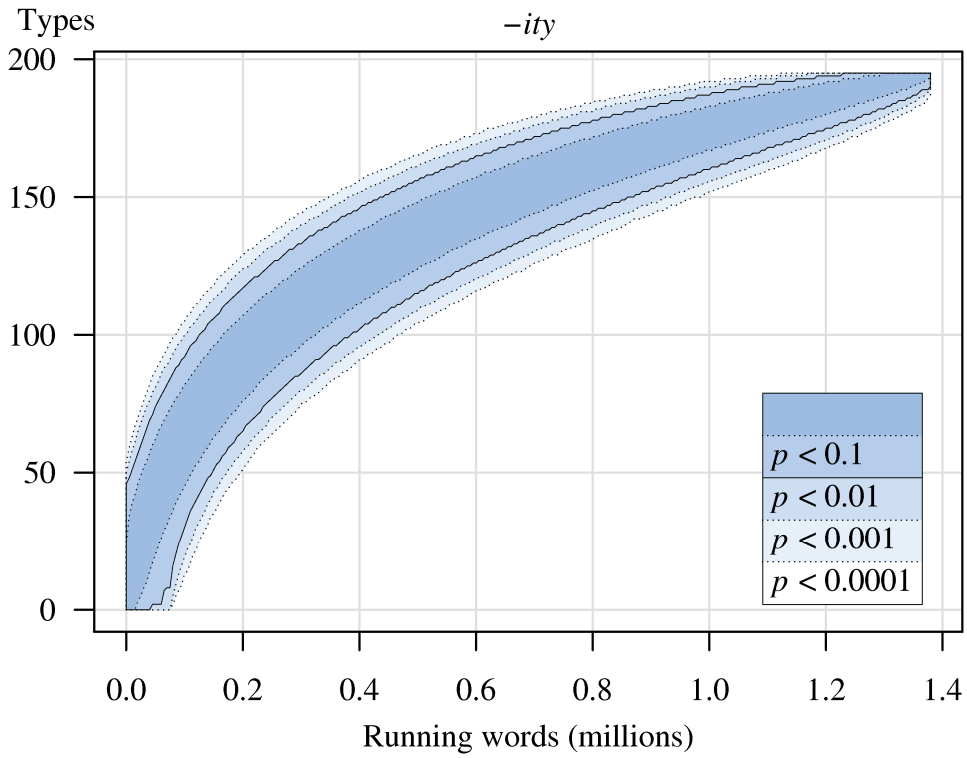

Figure 4. Bounds for 1,000,000 type accumulation curves.

It is also possible to generate accumulation curves for hapaxes instead of all types. Furthermore, it is possible to use the number of suffix tokens rather than the number of running words on the *x* axis. Figure 5 shows the bounds for 1,000,000 hapax accumulation curves as a function of the number of suffix tokens. Plotting hapax counts from subcorpora on this graph is essentially an improved version of comparing values of Baayen's (e.g., 1992: 115) category-conditioned degree of productivity P, which is the ratio between the number of hapaxes and the number of tokens with the suffix in the corpus.

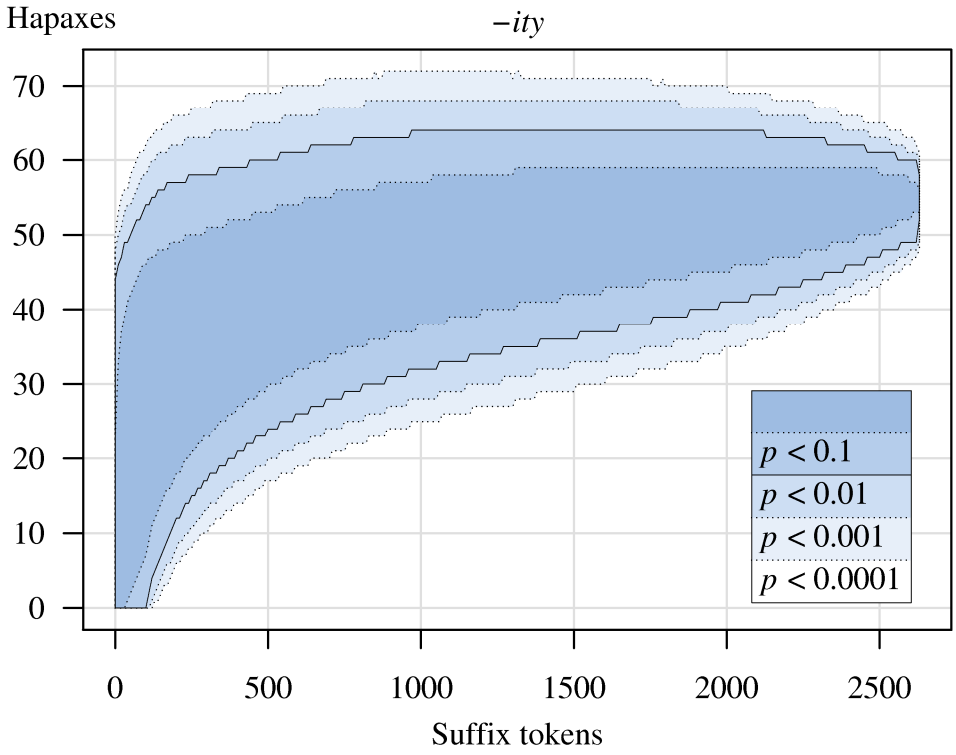

Figure 5. Upper and lower bounds for -*ity* hapaxes as a function of the number of -*ity* tokens.

For a more detailed discussion of the method and its implementation, see Säily and Suomela (forthcoming).

7. Analysis

The results of this study are presented in three sections. First, Section 7.1 introduces the full data set with some productivity statistics for the suffixes and an overview of the doublets, bases and most common types found in the corpus. Next, Section 7.2 presents the main results concerning the sociolinguistic variation in the productivity of -*ness* and -*ity*. Finally, Section 7.3 explores the effects of restricting the data set by leaving out certain groups of types which are less likely to have been formed productively.

7.1. Overview

Using the criteria described in Section 6.2 above, I found 314 -*ness* types and 195 -*ity* types in the 17th-century part of the CEEC; they are listed in

Appendix 1. Table 4 shows some statistics for this data set. As expected, all of the productivity measures (V, P, P*, A) presented in Sections 2.2.2 and 2.2.3 indicate that *-ness* is more productive than *-ity*. (I selected the threshold θ=8 for the activation level because the same threshold was used by Baayen 1993.) Furthermore, the proportion of hapaxes out of all types is greater for *-ness* than for *-ity*, while the proportion of dis legomena is very similar for both; thus, the frequency distribution of *-ness* types is more skewed towards the low-frequency end, which also points to its higher productivity (Baayen and Lieber 1991: 810).

	N	**V**	**P = n_1/N**	**N_1 ~ P***	**n_2**	**A (θ=8)**
-ness	3,941	314	0.04	157 (50% of V)	45 (14% of V)	537
-ity	2,630	195	0.02	55 (28% of V)	25 (13% of V)	346

Table 4. Productivity measures for *-ness* and *-ity* in the full data set. N = number of suffix tokens, V = extent of use (number of types), P = category-conditioned degree of productivity, P* = hapax-conditioned degree of productivity, n_1 = number of hapax legomena (types occurring once), n_2 = number of dis legomena, A = activation level.

Next, Table 5 presents a list of the *-ness* and *-ity* doublets found in the corpus. As can be seen, many of them do not share the exact same base, but they were yet close enough that I decided to include them in the list (e.g., *depravedness – depravity*). Some of the *-ity* types are clearly loanwords with no real base in English (e.g., *piety*). Nevertheless, the list does in my opinion show that the suffixes were to some degree interchangeable. Interestingly, where the number of tokens exceeds 1–2, it is usually *-ity* which has the larger number; it may be that the established form in *-ity* sometimes fails to block *-ness* due to the greater productivity of *-ness* (cf. the discussion on blocking in Section 2.2.5 above). Where the numbers are more similar to each other, the suffixes often seem to be in genuine competition with each other (e.g., *stupidness – stupidity*); this is supported by a qualitative analysis of the instances. The number of doublets, 19, is quite low, but this may be due more to the fact that lexical/derivational

studies require a large corpus to capture the low-frequency items, than to a lack of synonymy between the suffixes.

-ness		***-ity***	
barbarousness	2	*barbarity*	1
briefness	1	*brevity*	14
depravedness	1	*depravity*	2
firmness	2	*firmity*	1
generousness	1	*generosity*	4
gentleness	2	*gentility*	2
humbleness	6	*humility*	45
inactiveness	1	*inactivity*	2
infiniteness	1	*infinity*	3
insensibleness	1	*insensibility*	2
jolliness	1	*jollity*	4
niceness	1	*nicety*	1
nobleness	4	*nobility*	31
piousness	1	*piety*	25
quietness	15	*quiety*	1
rareness	2	*rarity*	13
singleness	1	*singularity*	1
soberness	1	*sobriety*	7
stupidness	1	*stupidity*	2
Total tokens	45	Total tokens	161

Table 5. -*ness* and -*ity* doublets in the 17th-century part of the CEEC.

There is, however, a difference in what kinds of bases the suffixes attach to — -*ity* only occurs with French or Latinate bases in my data, whereas -*ness* occurs with bases of all etymologies. Another difference is that only -*ness* occurs with participial adjectives such as *depraved*. For both suffixes, adjectival bases are by far the most common; -*ness* has one nominal base (*wit*) and one that could have been a verb or an adjective (*busy*), both formations (*witness*, *business*) dating from Old English, while -*ity* has no verbs and a few nouns (e.g., *author*, *infortune*, *par*), the corresponding words in -*ity* being originally loans from French or Latin.

To get an idea of how -*ness* and -*ity* were used by the most interesting social groups as predicted by my hypothesis — men, women, and the higher and lower ranks — I calculated some statistics for these as well, shown in Table 6. In this overview, statistical significance has not been

tested; this is left to the subsequent sections offering the main results of the study. For the sake of simplicity, the ranks are represented through a rough division into gentry (including royalty, nobility, gentry proper, and upper clergy) and non-gentry (the rest). There is one person (who contributed one *-ness* instance) whose rank is unknown, which is why the running words and *-ness* tokens for gentry + non-gentry do not equal the totals. The statistics given below focus on tokens because they are comparable across subcorpora of various sizes; while token counts are a poor indicator of productivity, they can still tell us something about the use of the words in *-ness* and *-ity*.

	Running words	***-ness* tokens (per 1,000 words)**	***-ness* types**	***-ity* tokens (per 1,000 words)**	***-ity* types**
Men	1,038,951	2,892 (2.78)	276	2,178 (2.10)	186
Women	340,116	1,049 (3.08)	127	452 (1.33)	82
Gentry	1,099,036	3,075 (2.80)	263	1,916 (1.74)	179
Non-gentry	279,740	865 (3.09)	142	714 (2.55)	113
C17 total	1,379,067	3,941 (2.86)	314	2,630 (1.91)	195

Table 6. Tokens and types of *-ness* and *-ity* among different social groups in the 17th-century part of the CEEC.

For instance, the relative frequency of *-ness* tokens in the corpus is somewhat higher than that of *-ity*. Women and non-gentry seem to use *-ness* tokens more frequently than men and gentry. On the other hand, women have relatively few *-ity* tokens, while non-gentry have more than the other groups.

For a closer look, I prepared a list of the most common types in each group. Consider Table 7 and Table 8, listing the favourite *-ness* types of women and men, respectively.

Women's top ten types	Women's tokens (per 1,000 words)	Men's tokens (per 1,000 words)
business	237 (0.70)	947 (0.91)
kindness	198 (0.58)	182 (0.18)
happiness	143 (0.42)	257 (0.25)
highness	53 (0.16)	93 (0.09)
sickness	45 (0.13)	150 (0.14)
goodness	33 (0.10)	107 (0.10)
witness	26 (0.08)	111 (0.11)
thankfulness	24 (0.07)	60 (0.06)
weakness	22 (0.06)	41 (0.04)
unkindness	19 (0.06)	16 (0.02)

Table 7. Women's most used -*ness* types.

Men's top ten types	Men's tokens (per 1,000 words)	Women's tokens (per 1,000 words)
business	947 (0.91)	237 (0.70)
happiness	257 (0.25)	143 (0.42)
kindness	182 (0.18)	198 (0.58)
sickness	150 (0.14)	45 (0.13)
witness	111 (0.11)	26 (0.08)
goodness	107 (0.10)	33 (0.10)
highness	93 (0.09)	53 (0.16)
holiness	63 (0.06)	1 (0.00)
thankfulness	60 (0.06)	24 (0.07)
readiness	54 (0.05)	4 (0.01)

Table 8. Men's most used -*ness* types.

The most frequent type for both men and women is the highly lexicalised *business*, as in (11) below:

(11) All your delit, is wall heare, and I shall pray and long to heare of your prospring in your *besnes* and good settelment agine; […]
(BAS 1651 FBASIRE 109:Heading, emphasis added to this and all of the following examples)

Most of the other types in the top ten are shared by both genders as well. In terms of relative frequency, however, men only have one really dominant type, the above-mentioned *business*, while women display high frequencies of *kindness* and *happiness* as well. It seems that women's letters often thank people for their kindness (12) and wish them happiness (13):

(12) Itt is now high tyme that I acknowledge the reseit of both your letters by this bearer, and withall how I aprehend your *kindnes* so many ways expressed to me in them; […]
(COR 1619 LRUSSEL 61:Heading)

(13) I thinke you weare never trobled with so tedious a letter, and therfore now I will conclude, with a hartie wish of all *hapines* to you and yours; and soe, sweet Madam, ons more farewell.
(COR 1632? DRANDOLPH 250:Heading)

The most dramatic difference between men and women seems to be in the use of *holiness*, which is frequent with men but only used once by women. The word is most commonly used in combination with the pronoun *his* to mean the pope:

(14) And yet nevertheles I have not understood that his *hollynes* hath rejected, or that he dislyketh that mocion, but that he only differreth it for some tyme, […]
(A 1608 T TFITZHERBE 15:Heading)

The one use by a woman, however, is abstract:

(15) My good Ned — The Lord in mercy blles you, and giue you interest in his sonne Christ, and such a measure of *holyness*, that you may liue heare like his child.
(HAR 1639 BHARLEY 30:Heading)

This could mean that women did not have much occasion to talk about the pope — while noblewomen could have access to royalty and thus refer to *his* or *her highness*, they could not have any dealings with the pope, unlike male diplomats or clergy such as the Catholic Thomas Fitzherbert of example (14) above. On closer inspection, it turns out that most of the men's *holiness* types were in fact produced by Thomas Fitzherbert alone, but the explanation may still be valid; even Catholic women could probably not be in close contact with the pope.

As for the two types that make it into the women's top ten but not the men's, *weakness* and *unkindness*, both describe personal qualities and could convey a negative attitude, while the men's *readiness* is usually a positive state. It is tempting, if hopelessly speculative, to ascribe this usage to the fact that culturally women were supposed to be the weaker and more passive gender, whereas men were construed as strong and active. This is exemplified in (16), written by the nun Winefrid Thimelby (see also (20)

below for an example from nobility), and in (17) by the dramatist Ben Jonson.

(16) I must proclame my ioys, though 'twill discover much of my *weaknis* to be so esily transported from won passion to a nother […]
(TIX 1670S? WTHIMELBY 42:Heading)

(17) I do not only w^{t}h all *readynesse* offer my seruice, but will p~forme it w^{t}h as much integrity, as yo^{r} particular Fauor, or his Maiesties Right in any Subiect he hath, can exact.
(A 1605 T BJONSON 202:Heading)

This pattern seems to be discernible on the *-ity* side as well (see Table 9 and Table 10): women have types such as *vanity* and *importunity*, which are negative personal qualities that they often apply to themselves, whereas the men's *ability* and *authority* are assertive ones related to power, whether their own or that of others. Example (18) is from a gentlewoman, Maria Thynne, while (19) was written by Arthur Capel, the earl of Essex.

(18) My good mother, I assure you it is not any desire I have to offend you with my *importunities*, which maketh me so often trouble you with the testaments of my grieved mind […]
(E 1602 FO MTHYNNE 29:Heading)

(19) His Mat^{y} by ye employment hee has given mee hath putt mee above ye taking that satisfaction as gentlemen in like cases use to doe, and therefore I conceive it fitt my Resentment should bee exprest in such a way as ye *Authority* I have in my hands will permitt me to take, for which reason I have given away his troop, and forbid him to appear in my presence or to come within ye park of wch hee is Ranger, wherein I usually take ye aire […]
(C 1677 FN ACAPEL 132:Heading)

Women's top ten types	**Women's tokens (per 1,000 words)**	**Men's tokens (per 1,000 words)**
opportunity	50 (0.15)	231 (0.22)
necessity	38 (0.11)	154 (0.15)
quality	32 (0.09)	61 (0.06)
civility	20 (0.06)	36 (0.03)
charity	17 (0.05)	54 (0.05)
vanity	17 (0.05)	29 (0.03)
extremity	15 (0.04)	38 (0.04)
felicity	14 (0.04)	23 (0.02)
importunity	14 (0.04)	17 (0.02)
prosperity	13 (0.04)	34 (0.03)

Table 9. Women's most used *-ity* types.

Men's top ten types	Men's tokens (per 1,000 words)	Women's tokens (per 1,000 words)
opportunity	231 (0.22)	50 (0.15)
necessity	154 (0.15)	38 (0.11)
security	95 (0.09)	4 (0.01)
commodity	93 (0.09)	8 (0.02)
authority	80 (0.08)	11 (0.03)
university	65 (0.06)	3 (0.01)
quantity	64 (0.06)	10 (0.03)
quality	61 (0.06)	32 (0.09)
ability	56 (0.05)	6 (0.02)
charity	54 (0.05)	17 (0.05)

Table 10. Men's most used -*ity* types.

Like *kindness* and *happiness*, the qualities or states of *civility*, *felicity* and *prosperity* seem to usually be attributed by women to their correspondents or a third party and not to themselves. Example (20), penned by the noblewoman Anne Conway, illustrates further the women's tendency to commend others and to deprecate themselves.

(20) […] I pray S^{r} What you terme courtship in my former Letter, Lett me intreat you to account as a reall truth: for such in earnest are all the expressions that I can make of my esteeme of yo^{r} favoures; the commendations you give the enclosed copy I sent you I must attribute to yo^{r} great *civility*, and yet I shall rest confident that you will both pardon the *weaknesses* of y^{t} paper and alsoe beare wth any other I shall send you of the Like kind […]
(A 1651 TC ACONWAY 493:Heading)

Finally, men talk more about business because they were normally the ones to conduct it — hence their higher use of *security*, *commodity* and *quantity* (see example (21) by the merchant George Richards). Men also lead in the use of *university*, because only they could attend such an institution of higher education.

(21) Here are also good *quantetyes* of Russia hides sold and its a staple *commodity*, but the benefitt might be by haveing them cheape bought in.
(D 1679 T GRICHARDS 546:Heading)

The second potentially significant social category is that of rank. A rough division of the data into gentry and non-gentry reveals an interesting difference: whereas the gentry likes to talk about *vanity*, the non-gentry is

more concerned with *humility*, as was deemed proper for the lower ranks at the time. Example (22) was written by a schoolmaster to a patron of the school:

(22) Obliged to your worship in all the offices of *humilitie* and thankfullnes,
Adrian Carew.
(HUT 1608 ACAREW 210:Heading)

Some of the other differences may be due to women influencing gentry more than non-gentry, as most literate women came from the gentry.

Top ten types for non-gentry	**Non-gentry tokens (per 1,000 words)**	**Gentry tokens (per 1,000 words)**
opportunity	85 (0.30)	196 (0.18)
necessity	49 (0.18)	143 (0.13)
university	46 (0.16)	22 (0.02)
commodity	40 (0.14)	61 (0.06)
authority	36 (0.13)	55 (0.05)
quantity	35 (0.13)	39 (0.04)
security	29 (0.10)	70 (0.06)
quality	20 (0.07)	73 (0.07)
humility	16 (0.06)	29 (0.03)
infirmity	15 (0.05)	34 (0.03)

Table 11. Top ten *-ity* types for non-gentry.

Top ten types for gentry	**Gentry tokens (per 1,000 words)**	**Non-gentry tokens (per 1,000 words)**
opportunity	196 (0.18)	85 (0.30)
necessity	143 (0.13)	49 (0.18)
quality	73 (0.07)	20 (0.07)
security	70 (0.06)	29 (0.10)
commodity	61 (0.06)	40 (0.14)
charity	60 (0.05)	11 (0.04)
authority	55 (0.05)	36 (0.13)
ability	49 (0.04)	13 (0.05)
vanity	46 (0.04)	0 (0.00)
civility	45 (0.04)	11 (0.04)

Table 12. Top ten *-ity* types for gentry.

For *-ness*, there do not seem to be any real differences. Only the gentry have *readiness* in the top ten, but it is #11 for the non-gentry; only the non-gentry talk much about *holiness*, but these instances are mostly due to a single clergyman, Thomas Fitzherbert.

Top ten types for non-gentry	Non-gentry tokens (per 1,000 words)	Gentry tokens (per 1,000 words)
business	186 (0.66)	998 (0.91)
kindness	93 (0.33)	287 (0.26)
holiness	61 (0.22)	3 (0.00)
happiness	48 (0.17)	352 (0.32)
witness	39 (0.14)	98 (0.09)
highness	38 (0.14)	108 (0.10)
goodness	38 (0.14)	102 (0.09)
sickness	32 (0.11)	163 (0.15)
thankfulness	25 (0.09)	59 (0.05)
weakness	23 (0.08)	40 (0.04)

Table 13. Top ten -*ness* types for non-gentry.

Top ten types for gentry	Gentry tokens (per 1,000 words)	Non-gentry tokens (per 1,000 words)
business	998 (0.91)	186 (0.66)
happiness	352 (0.32)	48 (0.17)
kindness	287 (0.26)	93 (0.33)
sickness	163 (0.15)	32 (0.11)
highness	108 (0.10)	38 (0.14)
goodness	102 (0.09)	38 (0.14)
witness	98 (0.09)	39 (0.14)
thankfulness	59 (0.05)	25 (0.09)
weakness	40 (0.04)	23 (0.08)
readiness	36 (0.03)	22 (0.08)

Table 14. Top ten -*ness* types for gentry.

7.2. Sociolinguistic analysis using type accumulation curves

My hypothesis was that -*ity*, as a learned and etymologically foreign suffix, would be less productive with poorly educated social groups, such as women and the lower ranks, than with well-educated groups, such as men and the higher ranks. Let us first examine the category of gender. Figure 6 and Figure 7 show the type counts from the subcorpora of men and women plotted on the bounds for -*ity* and -*ness* types in the full corpus. Figure 6 confirms the hypothesis: women use a significantly low number of -*ity* types ($p < 0.001$). With -*ness* types, neither group is significantly different from the corpus as a whole.

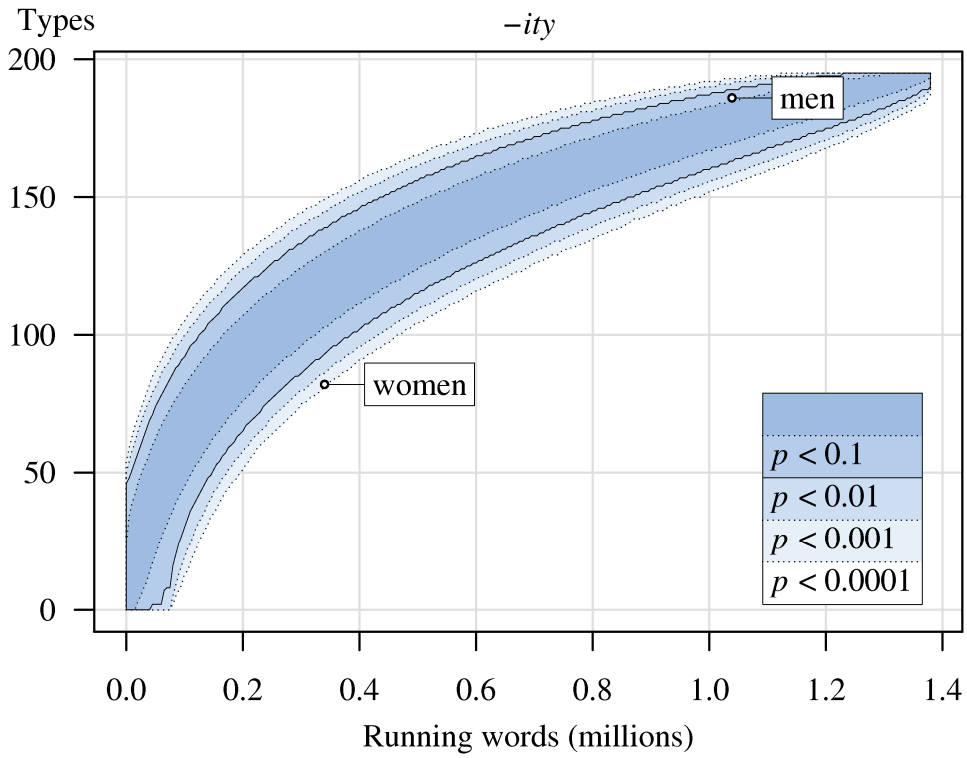

Figure 6. Gender and -*ity* types in the 17th-century part of the CEEC.

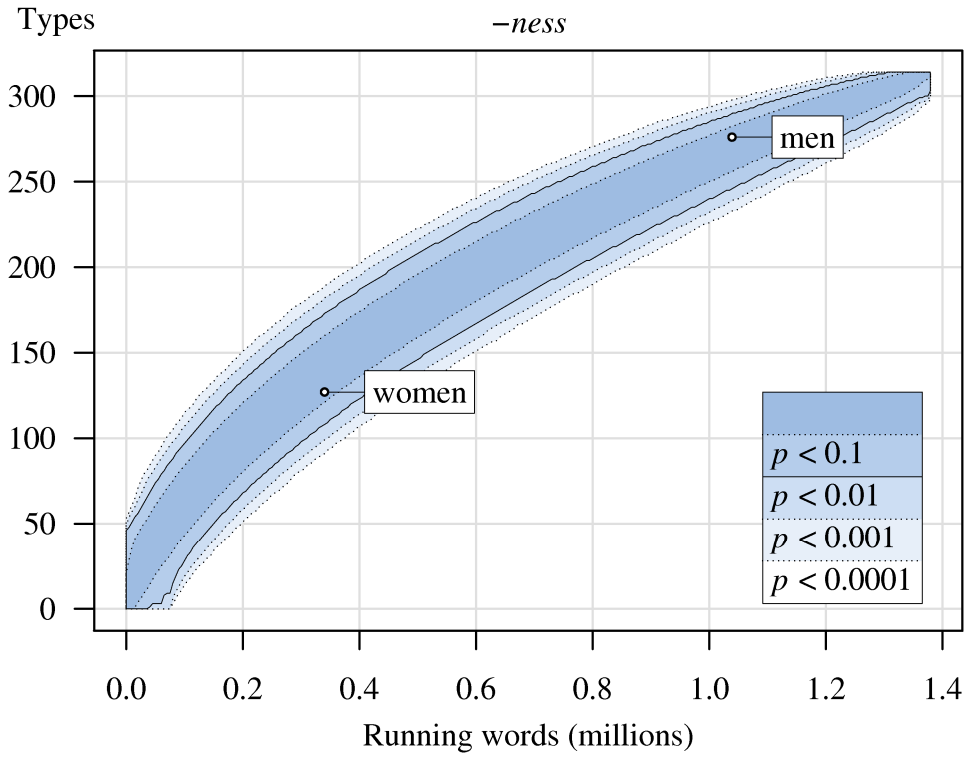

Figure 7. Gender and -*ness* types in the 17th-century part of the CEEC.

The bounds for hapax counts as a function of the number of *-ity* tokens turned out to be too wide for a significant difference to emerge (see Figure 8). It may be that this measure requires more data to become usable. As can be seen from Figure 9, the bounds for *-ness* hapaxes are somewhat narrower; nevertheless, the measure will not be used in the remainder of this section due to the problem with *-ity*.

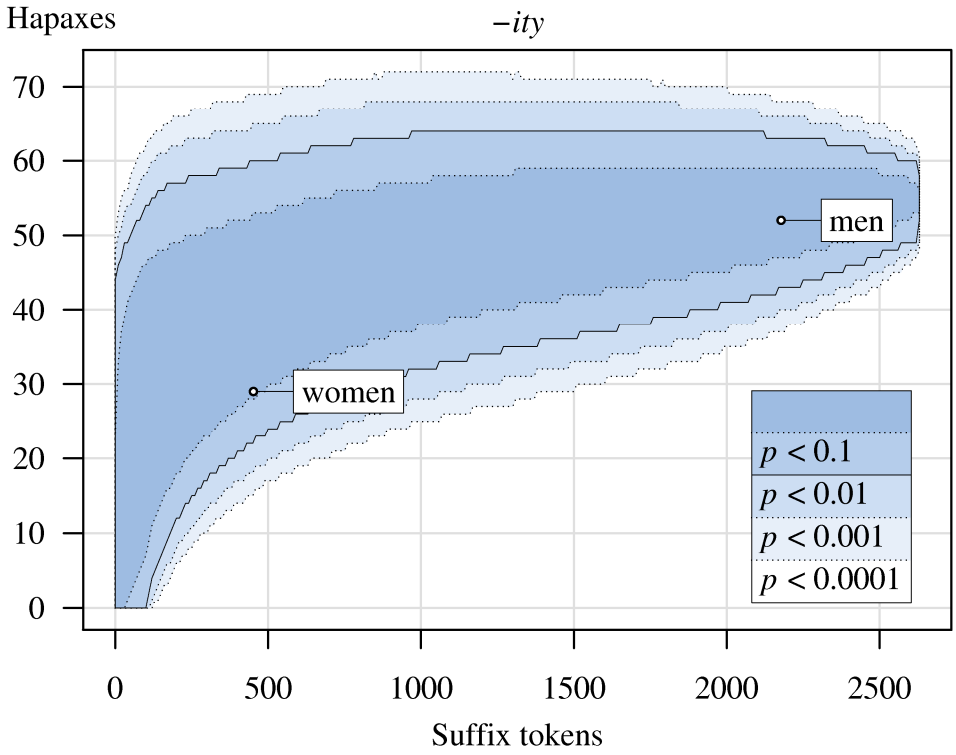

Figure 8. Gender and *-ity* hapaxes in the 17th-century part of the CEEC.

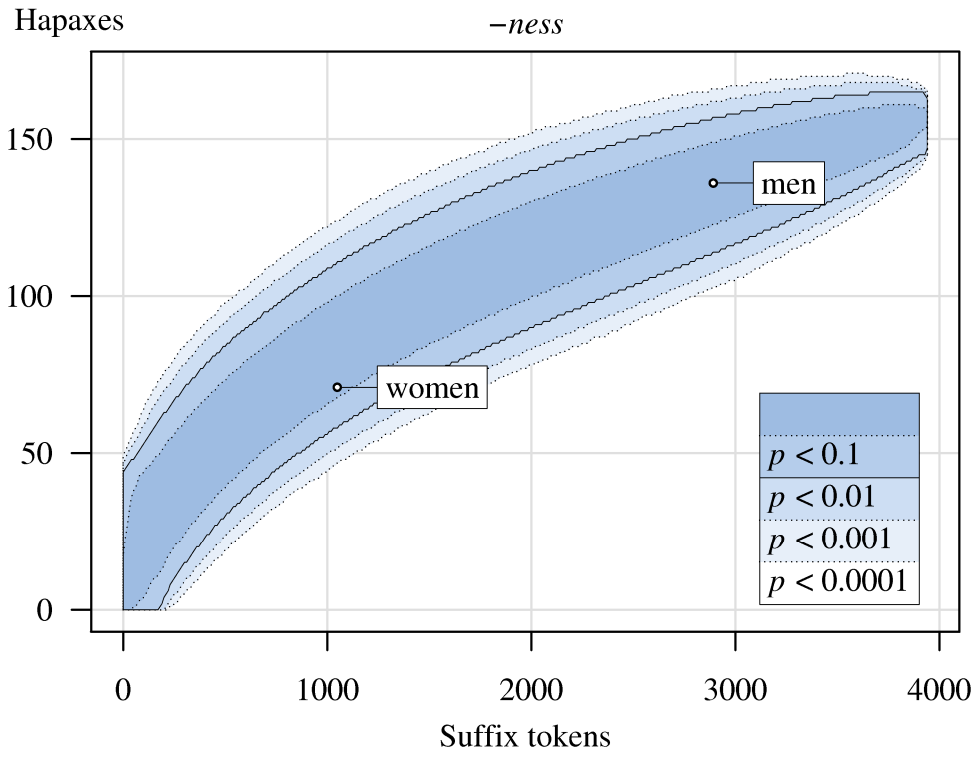

Figure 9. Gender and -*ness* hapaxes in the 17th-century part of the CEEC.

The next category of interest is that of social rank. Here some exploratory data analysis was needed in order to determine the most suitable level of granularity for the rank divisions (recall Section 3.2.2 above). Figure 10 shows the results for the most fine-grained hierarchy suggested by Nevalainen and Raumolin-Brunberg (2003: 136, Model 1). This is clearly too fine-grained in that the amount of data per subcorpus becomes too small for any differences to show. The figure for -*ness* is similar and thus not shown here; unless otherwise noted, this holds for the following figures as well. Nevalainen and Raumolin-Brunberg's (2003: 136) Model 2, in which the different kinds of gentry proper are merged, as are the clergy, also yields no results with these data.

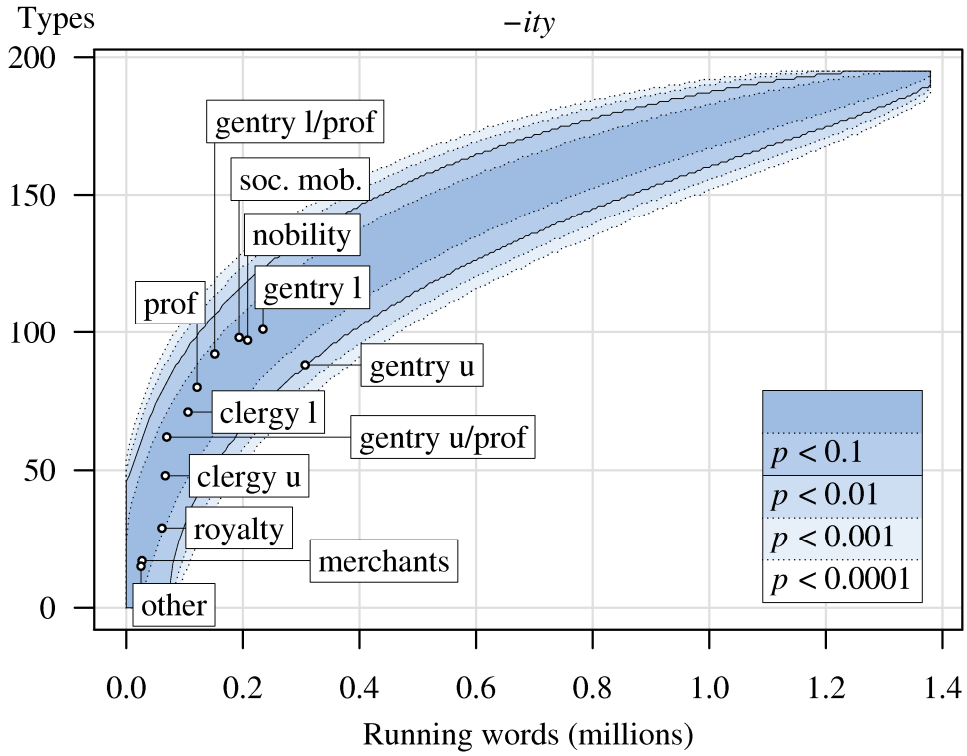

Figure 10. Rank (fine-grained hierarchy) and *-ity* types in the 17th-century part of the CEEC.

Figure 11 shows the results for Nevalainen and Raumolin-Brunberg's (2003: 136) Model 4, with only a three-way division; their Model 3 is the same with the addition of the socially mobile group, which can be seen in Figure 10. Unfortunately, the most interesting group is that of the poorest sort, who are also poorly educated, and there is still too little data from them. Finally, Figure 12 presents a rough two-way division into gentry — consisting of royalty, nobility, gentry proper, and upper clergy — and non-gentry, including professionals, merchants, lower clergy, and others, such as yeomen. It seems that there would be enough data for a difference to show up if there were one. I also tested the effect of leaving women out of the calculations, but still no difference emerged. Therefore, I conclude that this division is not relevant in terms of the use of *-ity*; professionals could be as well-educated as the gentry, and it was only the poorest (who did constitute the vast majority of the population!) whose education was the most lacking.

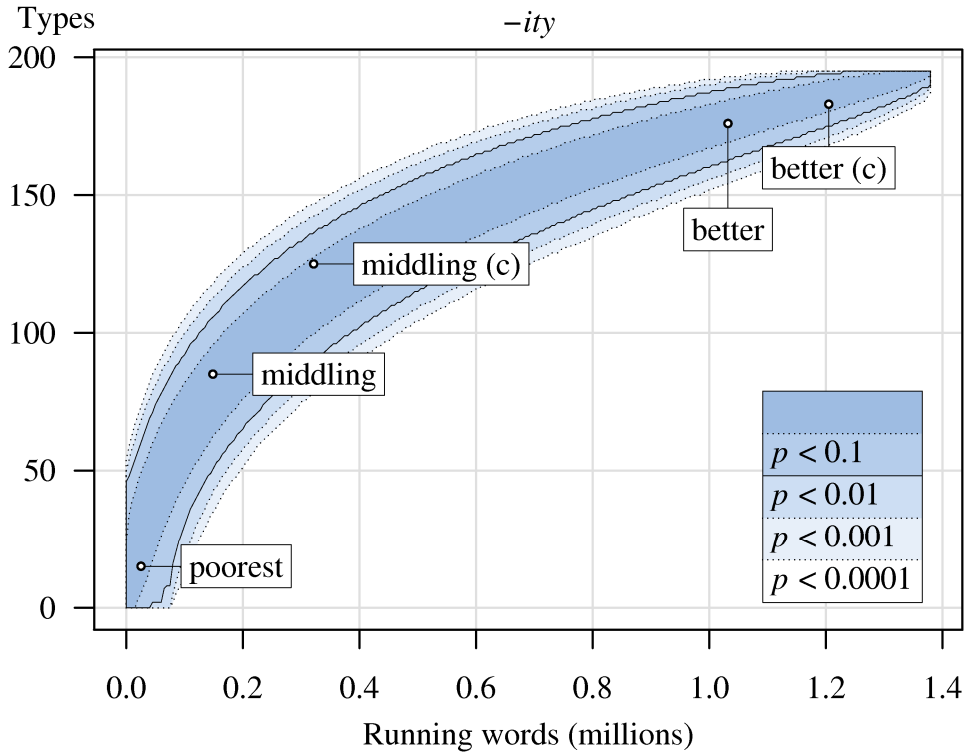

Figure 11. Rank (tripartite division) and *-ity* types in the 17th-century part of the CEEC. Alternative results are shown for clergy (c) included in the middling sort and in the better sort, respectively.

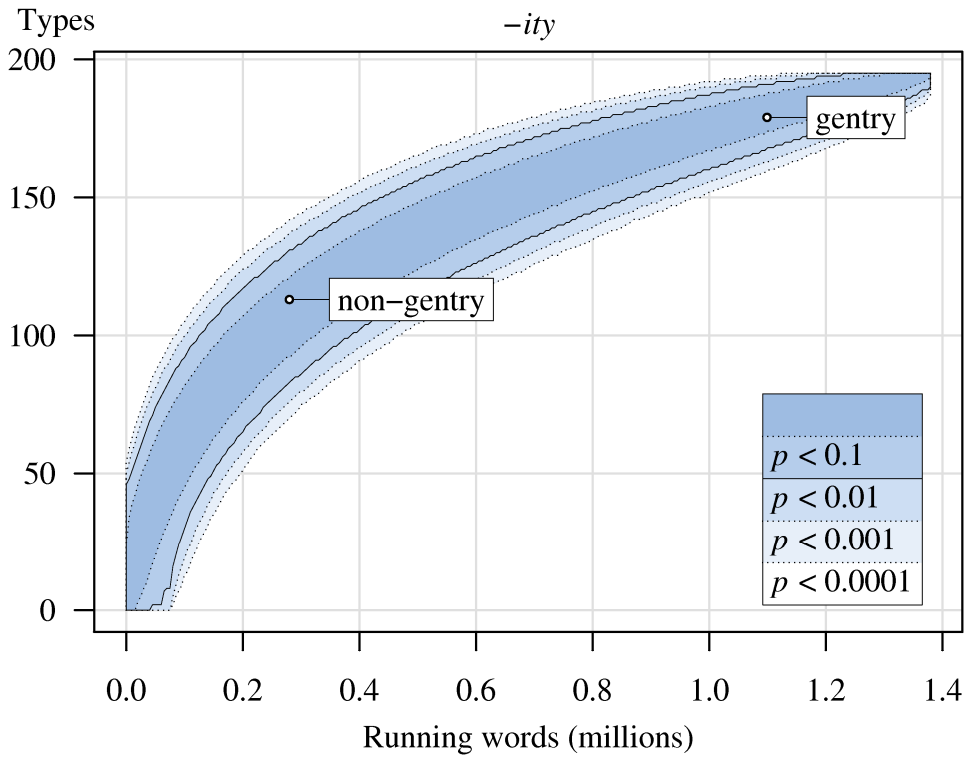

Figure 12. Rank (dichotomous model) and *-ity* types in the 17th-century part of the CEEC.

Besides the categories predicted to be significant by my hypothesis, I analysed the effects of region and time period. Unsurprisingly, type counts from subcorpora based on the main domicile of the writer revealed no significant results. This could, however, be due to the fact that the amount of data was very small for all areas except for 'other', as can be seen from Figure 13.

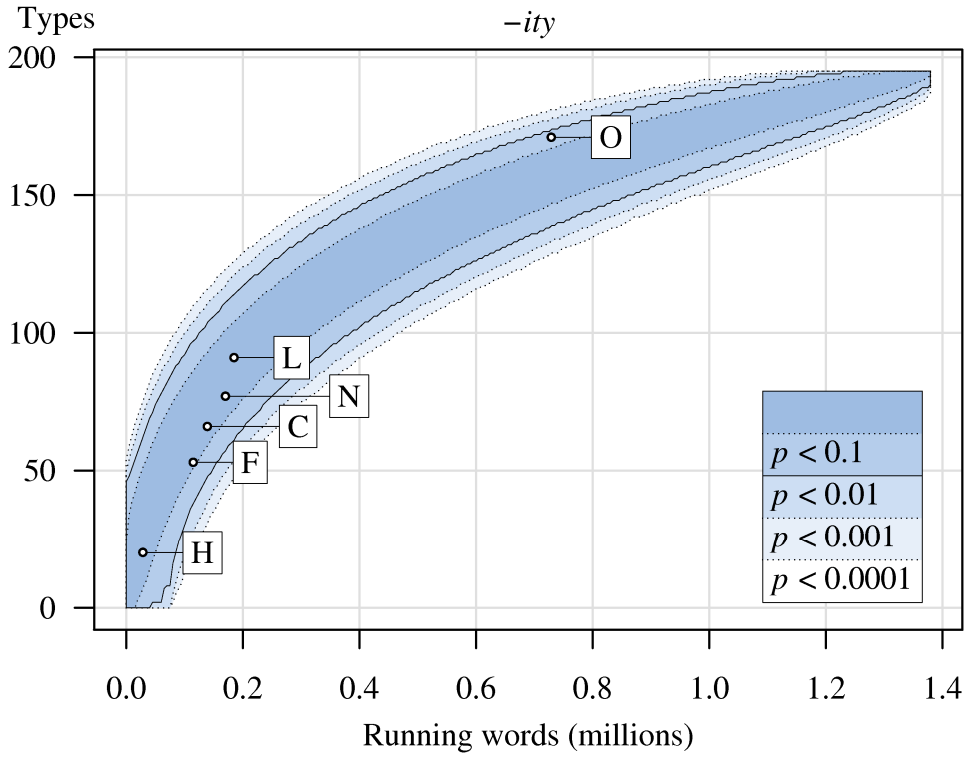

Figure 13. Domicile and *-ity* types in the 17th-century part of the CEEC.

More surprising results were yielded by a chronological comparison. With a division into two c. 40-year periods (created to ensure sufficient data by combining the 20-year periods given in the auxiliary database), it became apparent that the earlier period, 1600–1639, had a significantly low number of *-ity* types ($p < 0.001$) — see Figure 14. A tentative explanation for this could be that the use of *-ity*, as that of many Latinate features, was register-specific, and spread from more formal contexts to less formal ones during the 17th century (cf. Nevalainen and Tieken-Boon van Ostade 2006: 281–

282; Riddle 1985: 455–456). With *-ness* types, there was no significant difference, as evident from Figure 15.

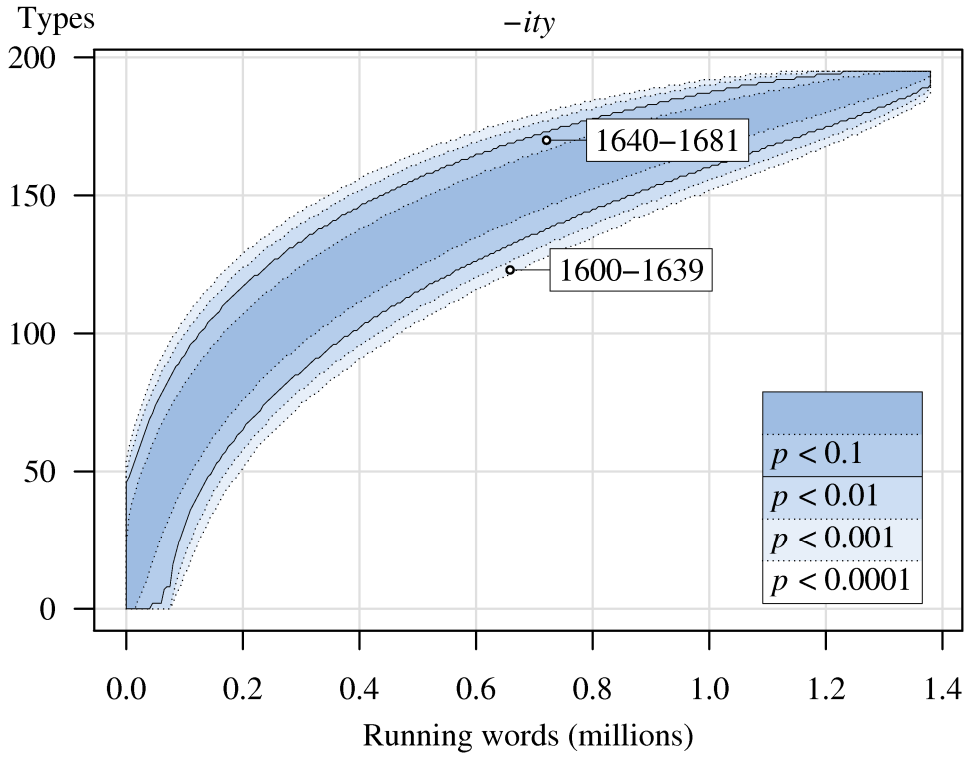

Figure 14. Time period and *-ity* types in the 17th-century part of the CEEC.

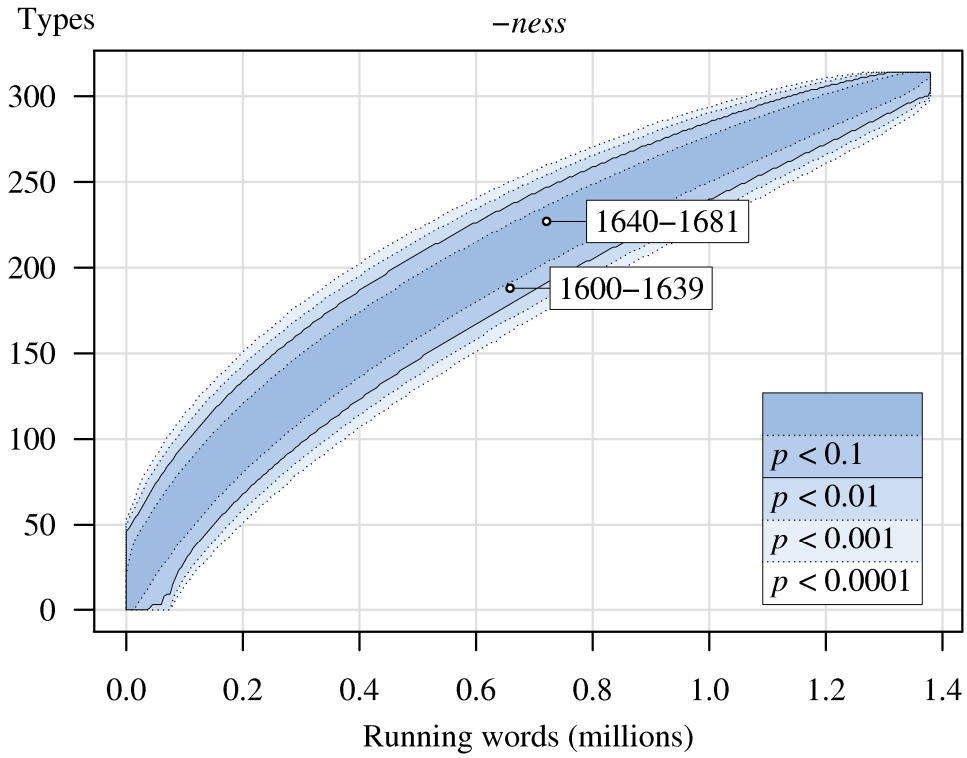

Figure 15. Time period and *-ness* types in the 17th-century part of the CEEC.

7.3. Restrictions on type counts

The results presented above were obtained with the full data set; this included types with no extant base (usually loanwords), prefixed types that could have been formed through prefixation on a suffixed base rather than suffixation on a prefixed base, as well as types that had been established in the language for centuries. This section explores the effect on the results of restricting the kinds of types that are counted. The goal is to ensure that the results pertain to morphological productivity rather than just the vocabulary of the social groups under study. The effect of each restriction will first be examined separately, and finally all restrictions will be applied simultaneously.

The first obvious restriction was to leave out the types with no extant base in the 17th century (or a base that could only have been used by applying a truncation rule, such as *sagacious* minus *-ous* for *-ity*), as determined with the help of the OED and its attestation dates. As OED entries are easily antedated, however, one 18th-century base was included, namely *inactive*.[3] These types are marked with a B in Appendix 1. Also left out were types with bases that were rare or nonce-formations or ones that required a special kind of manipulation for use with *-ity* (e.g., *quiet* + *-ity* → *quiety* rather than *quietity*); these are marked with B* in Appendix 1. Thus, the overall number of baseless *-ity* types (B + B*) amounted to 54,

[3] To illustrate the ease of antedating OED entries, here are some *-ness* and *-ity* words which occur in my data at an earlier date than that provided by the OED as the first attestation date: *depravity* (1635 in my data / 1641 in the OED), *impenetrability* (1651/1665), *solicitousness* (1632/1636), *helpfulness* (1642/1643), *unsatisfiedness* (1631/1646), *healthiness* (1665/1670), *uselessness* (1641/1690), *imperviousness* (1665?/1727), *oversweetness* (1653/1759), *unfittingness* (1647/1861), *obstructedness* (1662/unlisted).

while there were only three such *-ness* types: *ear-witness*, *eye-witness* and *wilderness*.

Table 15 shows some statistics for the remaining data set. In comparison with Table 4 above, the numbers for *-ity* have decreased drastically. However, there is a slight increase in its P value and in the percentage of hapax legomena out of all types. Thus, these productivity measures seem to show quite rightly that *-ity* with extant bases is more likely to be productive than *-ity* with both extant and non-extant bases; but taking into account the problems with comparing these figures, this may be a mere coincidence.

	N	**V**	**P = n_1/N**	**n_1 ~ P***	**n_2**	**A (θ=8)**
-ness	3,928	311	0.04	157 (50% of V)	44 (14% of V)	524
-ity	1,730	141	0.03	44 (31% of V)	17 (12% of V)	238

Table 15. Productivity measures for *-ness* and *-ity* instances with an extant base.

Despite the restriction on what was counted, the results from the full data set were replicated with this one, as can be seen from Figure 16 and Figure 17 below. As with the full data set, no other statistically significant results emerged. This was also the case when only those types marked with a B were removed from the data set, instead of both B and B*.

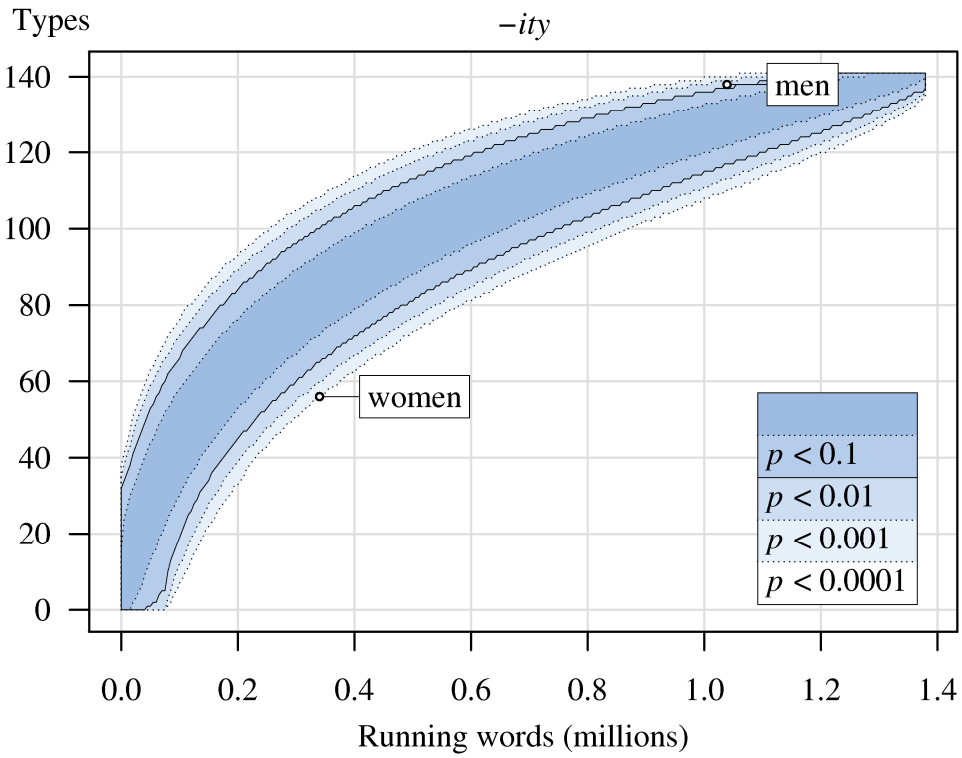

Figure 16. Gender and *-ity* types with an extant base in the 17th-century part of the CEEC.

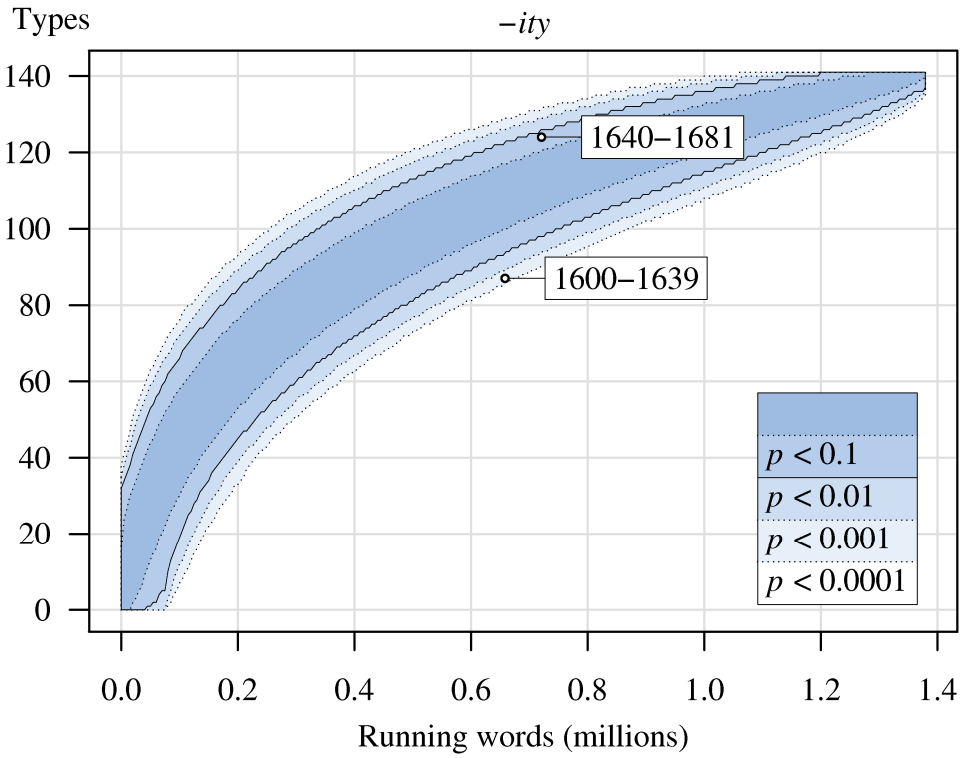

Figure 17. Time period and *-ity* types with an extant base in the 17th-century part of the CEEC.

The next restriction was to remove prefixed types from the full data set. This included any types that could conceivably have been formed by adding a prefix to a complex base in *-ness* or *-ity*. In general, this meant that the base needed to be extant in the 17th century according to the OED; but as the OED does not provide perfect coverage, some later bases were also accepted, if the adjective from which the base was formed did exist in the 17th century. The types categorised as prefixed can be seen in the list of *-ness* and *-ity* types in Appendix 1; the statistics for the remaining data set are shown in Table 16.

	N	**V**	**P = n_1/N**	**N_1 ~ P***	**n_2**	**A (θ=8)**
-ness	3,838	274	0.04	127 (46% of V)	41 (15% of V)	483
-ity	2,459	156	0.03	41 (26% of V)	19 (12% of V)	258

Table 16. Productivity measures for non-prefixed *-ness* and *-ity* instances.

Here, too, the P value for *-ity* increases as compared with Table 4. The number of hapaxes, on the other hand, decreases for both *-ness* and *-ity*, which shows that the prefixed types had the effect of exaggerating the hapax-conditioned productivity of both processes. Of course, some of the types categorised as prefixed could have been formed by adding the suffix to the prefixed adjective, so the numbers could now be somewhat too low. In any case, the results from the full data set were again replicated; the statistically significant differences are shown in Figure 18 and Figure 19 below.

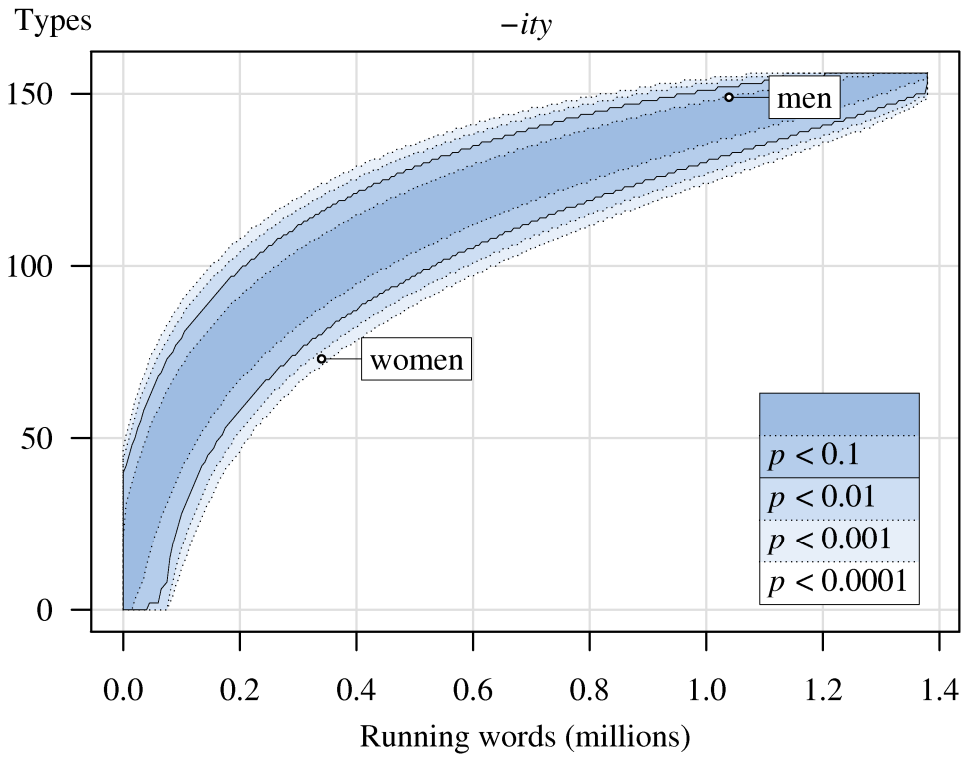

Figure 18. Gender and non-prefixed -*ity* types in the 17th-century part of the CEEC.

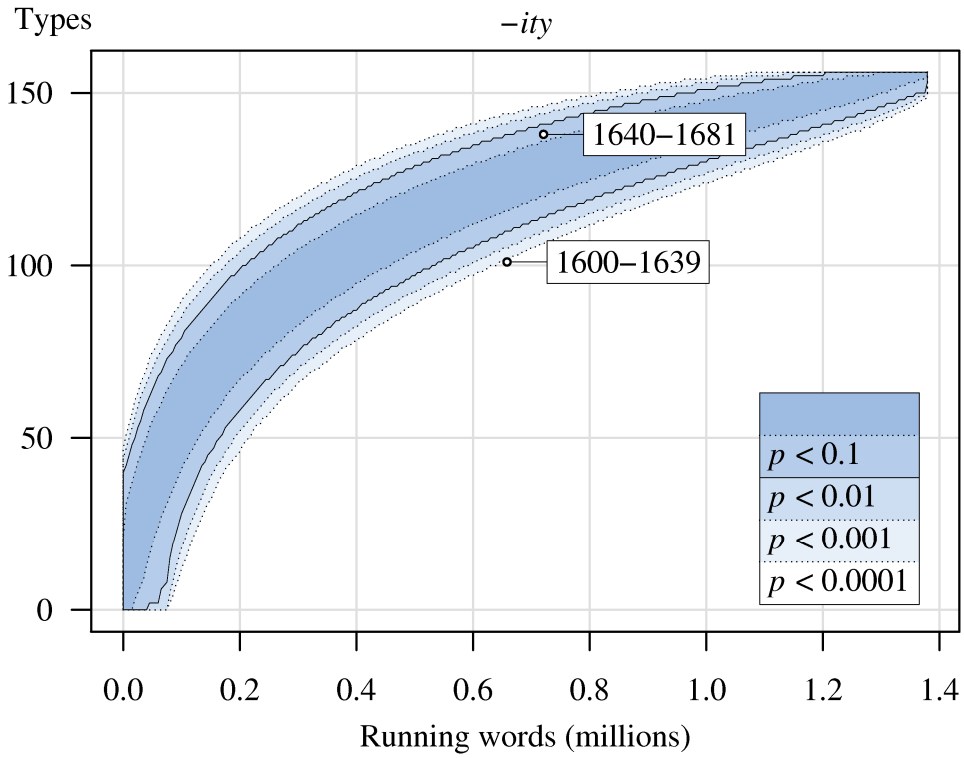

Figure 19. Time period and non-prefixed -*ity* types in the 17th-century part of the CEEC.

The last restriction was to leave out types, whether originally loanwords or productively formed, which had been established in the language for centuries. This was accomplished by dividing the full data set into types whose first attestation date in the OED was before 1500, and those with a first attestation date of 1500 or later. The year 1500 was chosen as the cut-off point because words do not diffuse instantly into the entire population: it would be unfair to exclude, say, a word first used in a medical treatise in 1550, as it may still be new to the greater part of the language community in 1650.

	N	**V**	**P = n_1/N**	**n_1 ~ P***	**n_2**	**A (θ=8)**
-ness	869	159	0.12	101 (64% of V)	28 (18% of V)	242
-ity	350	64	0.08	27 (42% of V)	10 (16% of V)	119

Table 17. Productivity measures for *-ness* and *-ity* instances with a first attestation date of 1500 or later in the OED.

As can be seen from Table 17, the numbers of types and tokens of both *-ness* and *-ity* have decreased quite drastically. The proportion of hapax and dis legomena out of all types, however, has increased considerably for both, as has the category-conditioned degree of productivity P, indicating that the probability of encountering a new type is greater in this data set than in the original. The results from calculating upper and lower bounds for the number of types remain the same; the significant deviations can be seen in Figure 20 and Figure 21.

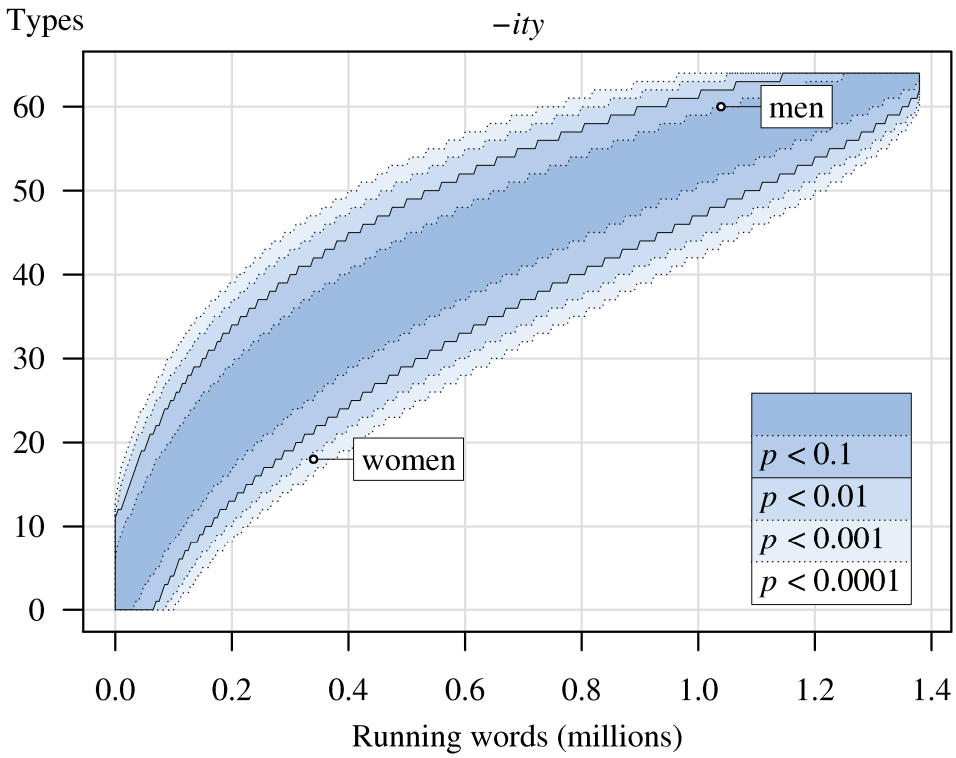

Figure 20. Gender and *-ity* types first attested in or after 1500, in the 17th-century part of the CEEC.

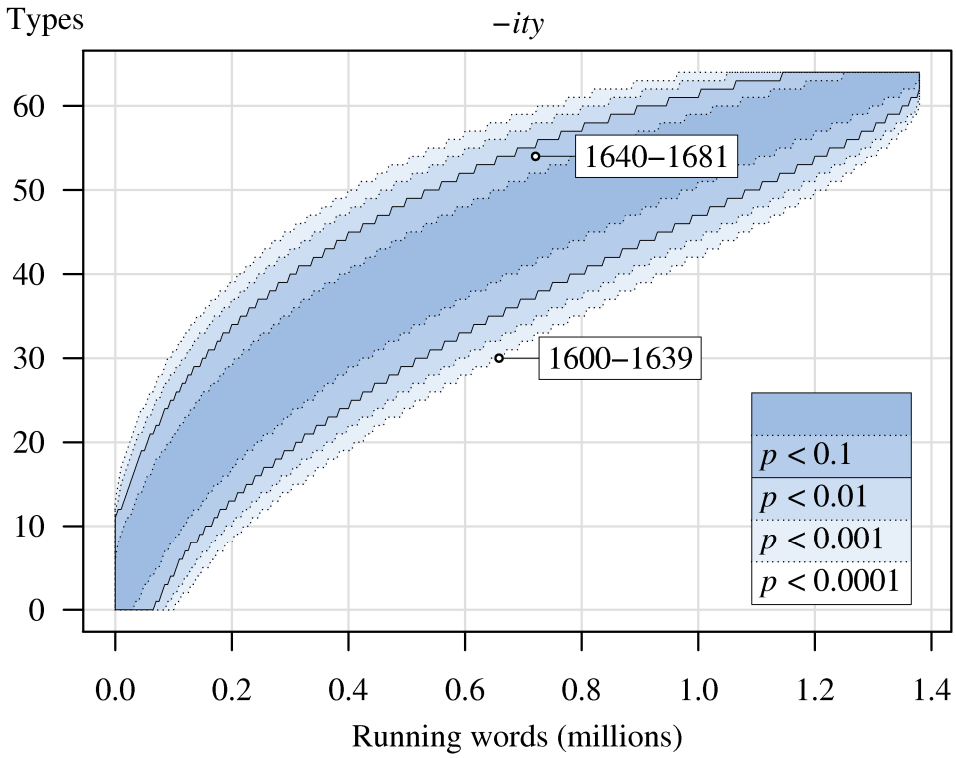

Figure 21. Time period and *-ity* types first attested in or after 1500, in the 17th-century part of the CEEC.

Finally, the results were put to the acid test: what would happen if all three restrictions were applied at once? The statistics for this data set are shown in Table 18. The numbers of tokens and types for *-ity* are very low, especially when compared with *-ness*. Nevertheless, the significant results remain the same, although the difference in time period receives a lower level of statistical significance; see Figure 22 and Figure 23 below. Thus, it is now safer to say that the results actually pertain to morphological productivity rather than just the lexicon.

	N	**V**	**P = n_1/N**	**n_1 ~ P***	**n_2**	**A (θ=8)**
-ness	814	130	0.10	80 (62% of V)	23 (18% of V)	201
-ity	162	29	0.07	12 (41% of V)	3 (10% of V)	46

Table 18. Productivity measures for non-prefixed *-ness* and *-ity* instances with an extant base and a first attestation date of 1500 or later.

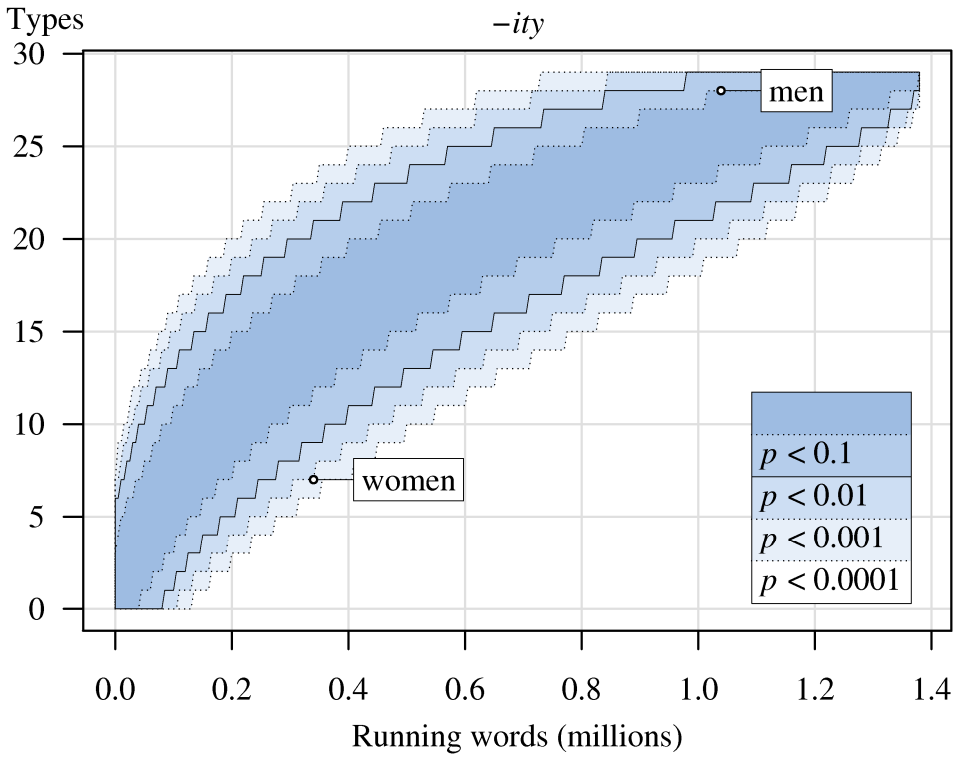

Figure 22. Gender and *-ity* types which are non-prefixed, first attested in or after 1500, and have an extant base, in the 17th-century part of the CEEC.

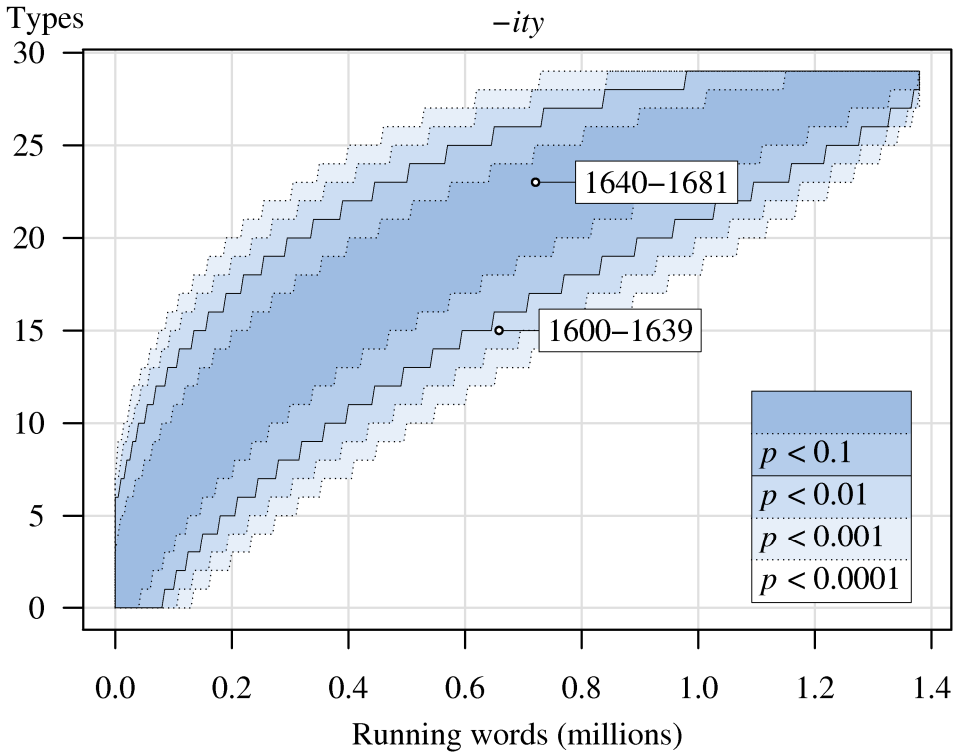

Figure 23. Time period and *-ity* types which are non-prefixed, first attested in or after 1500, and have an extant base, in the 17th-century part of the CEEC.

8. Discussion

It is now time for a critical examination of the study. Section 8.1 evaluates the methods used in the study, while Section 8.2 discusses the results and offers some explanations for them.

8.1. Evaluation of methods

This study has tested the method of calculating upper and lower bounds for type accumulation curves of random permutations of the corpus. The theoretical advantages of the method are enumerated in Section 6.4 above; Sections 7.2 and 7.3 show that it is possible to gain linguistically interesting and statistically significant results using this method even with a relatively small corpus and small numbers of types. Clearly, the way in which this method makes the most of the data through permutation testing is useful and much needed in, e.g., historical studies, in which the amount

of data is very limited to begin with, as well as studies in which a great many divisions into subcorpora are necessary to capture the variation across registers, different social groups, and so on.

Nevertheless, even this method does require a reasonable amount of data in order to yield statistically significant results; the lack of a significant difference for the lowest ranks in this study seems to be due to an insufficient amount of data from them (only 25,338 running words). Furthermore, even if there were enough data, it would be somewhat cumbersome to use this method to test all possible combinations of categories that could be significant together (e.g., professional men in 1640–1659). For this, a method based on factor analysis (e.g., Biber 1995) might be better, provided that it solved the problem of comparing type counts in a satisfactory way.

The bounds for hapax accumulation curves turned out too wide to be usable in this study (see 7.2 above). It would be interesting to test this measure with a larger corpus to see if an increase in the amount of data would render the measure more practical. It is also interesting that the shape of the bounds for -*ity* hapaxes is much wider than that of -*ness*, while the shapes of the bounds for types are quite similar for both -*ness* and -*ity*. According to Jukka Suomela (personal communication), this could mean that -*ity* is dominated by more 'temporary' hapaxes than -*ness*: when the size of the corpus is increased, further instances of the same type are found, and the hapax thus ceases to exist. Linguistically, this could be interpreted as a sign that -*ity* is less productive than -*ness*.

If the problem of wide bounds for hapax accumulation curves persists in larger corpora, this could have implications for the use of hapax-based productivity measures such as Baayen's (e.g., 1993) P and P*. Is the measure in any way reliable if it is so much a matter of chance which number of hapaxes occurs in a corpus of a given size, as implied by the

wideness of the bounds? This shows the danger of compressing productivity information into a single figure such as P = 0.04 for *-ness* and P = 0.02 for *-ity* in my data; it would be better to also examine the shape of the frequency distribution of the types, as well as the bounds for the accumulation curves. As a further drawback, the comparison of P values, as well as those of P* and A, is problematic, as they are dependent on the number of tokens of a particular morphological category; in addition, there is no obvious way of estimating the statistical significance of the results.

Granted, the type-based approach to productivity adopted in this study, even if statistically sound, is not ideal: as noted at the end of Section 4.2, the number of types is not the best possible measure of productivity, either. However, the restrictions applied in Section 7.3 should improve the reliability of the results, even if the group of remaining types was still not necessarily formed productively. The question arises how close to observing the productivity of *-ness* and *-ity* in action it is possible to get in a relatively small corpus like this.

Could some of the antedatings to the OED mentioned on page 4 (the non-prefixed ones) be counted as productively formed? This would be akin to Baayen and Renouf's (1996: 75) restriction that a new word must not be listed in a major dictionary. Taking the viewpoint of the individual user of the language, on the other hand, any sufficiently rare complex word could be formed productively, which again leads us to the hapaxes. But what is rare in this corpus, in addition to being largely a matter of chance, might not be so rare in a corpus representing, say, another register, even if the writers were the same people. This calls for further research with more data and different registers.

Finally, I would like to address two questions that could be raised about the design of the study. Firstly, if the hypothesis was that it was people's education which was the deciding factor in the productivity of *-ity*,

why not analyse the difference between people with and without a classical education rather than men vs. women or higher vs. lower ranks? The answer to this question is that a 17th-century person's education is known far less often than his or her rank or, indeed, gender. In addition, the information on education had not been coded into the auxiliary database at my disposal. The second question is why the study did not use the information given by the OED about how the words had been formed (e.g., through prefixation or borrowing). In short, the answer is that the classifications of the OED are not perfect and that I wanted to keep an open mind: for example, even a word which was originally borrowed could have been formed productively by the individual user.

8.2. Evaluation and explanation of results

Let us begin with an overview of the productivity of *-ness* and *-ity*. As expected, the productivity measures P, P* and A consistently showed *-ness* to be more productive than *-ity* in all of the data sets examined in Sections 7.1 and 7.3. Despite the problems with comparing these figures, this result seems clear enough. If *-ity* was simply less common than *-ness* rather than less productive, the proportion of *-ity* hapaxes out of all types would be expected to be larger, as would the P value. The above discussion on hapaxes, however, adds some uncertainty to the result.

Another source of uncertainty comes from the functions of word-formation discussed in Section 4: if *-ity* was already at this stage used more for labelling than syntactic recategorisation, it would be natural for it to have fewer hapaxes, as concepts tend to be mentioned more than once in a discussion. On the other hand, as remarked in 4.1, morphological processes used mostly for labelling are usually less productive than others. Nevertheless, it may be that hapax-based productivity measures such as P and P* are not the best way of assessing the productivity of *-ity*. Taking

into account the phenomenon of persistence studied by, e.g., Szmrecsanyi (2005), the repetition of the same type in a text could be partially due to language users being creatures of habit who "re-use recently used or heard linguistic options whenever they can" (Szmrecsanyi 2005: 113). Based on this, it could be argued that only counting hapaxes is insufficient for any morphological process, not just *-ity*.

In any case, a qualitative examination of the *-ity* types gives a definite impression of low productivity. Even with all three restrictions in force, there seems to be something wrong with many of the remaining 29 *-ity* types: either the base is somehow odd (e.g., *barbar* when there would be a more common word, *barbarous*) or some of the instances are plural (indicating that the meaning may have drifted from abstract to concrete), etc. Most of the words seem to be regarded by the OED as adaptations from Latin or loans from French rather than independent formations. The 130 *-ness* types, on the other hand, seem all quite normal.

The lower productivity of *-ity* is at variance with Nevalainen's (1999: 398) description of both *-ity* and *-ness* as "very productive" in Early Modern English. This may be a question of register: letters were an informal register close to spoken language, and *-ity* was probably more common in the more literate registers (cf. Plag, Dalton-Puffer and Baayen 1999 for the situation in present-day English). Incidentally, the most common *-ness* and *-ity* types may also reflect the register in which they were written: for instance, the most common *-ity* type in my material, *opportunity*, seems to be often used specifically about the favourable set of circumstances which enables the writing or conveying of the present letter.

Moving on to the main research question of this study, the results show that the productivity of *-ity* in my material varies across certain sociolinguistic categories, while the productivity of *-ness* does not exhibit statistically significant sociolinguistic variation. In accordance with my

hypothesis, women use *-ity* less productively than men; the case of rank, however, remains inconclusive due to lack of data from the lowest ranks.

The difference between men and women can be explained by the fact that *-ity* is a 'learned' and etymologically foreign (French/Latinate) suffix that is both semantically and phonologically more opaque than *-ness*; as women in this period had much less access to education than men (especially the sort of classical education given in universities, see Section 3.3 above), they were also less competent in the use of *-ity*. In other words, men were better equipped to do the strange things commented on above: they could adapt words from Latin into English, take short Latinate bases and add *-ity* to the end, etc. All this could have been done almost unconsciously by such bi- or trilingual individuals as the highly educated men of the period often were.

Unpredicted by my hypothesis, the productivity of *-ity* also seems to vary diachronically: the productivity of *-ity* in the earlier period, 1600–1639, is significantly low in comparison with the corpus as a whole. This could be interpreted as a linguistic change in progress: in the course of the 17th century, the use of *-ity* becomes more common in personal letters. Perhaps the increased use of non-native words and affixes, which seems to have peaked around the year 1600 in learned texts (see 4.2 above), shows up with a delay of several decades in personal letters, a more informal register. Another possible factor is the so-called Civil-War effect (e.g., Raumolin-Brunberg 1998): the increase in loose-knit social networks during the war of 1642–1649 may have accelerated the diffusion of many linguistic changes, including this one. On a pragmatic level, the use of *-ity* could have become more fashionable.

In my opinion, the main results of this study can be considered fairly reliable: the corpus is as good and unbiased as possible, the method is statistically sound, and the measure of type counts, while not perfect, is

commonly used in research on productivity. I will now consider some of the directions which the study could have taken, or which it invites further research to take.

In comparing the productivity of *-ness* and *-ity*, the study merely touched upon structural constraints. More could have been done in, e.g., examining the effect of different base types such as *-able* and *-ive*, as this has been shown to be an important factor in previous research (see Section 4 above). Furthermore, the phonological transparency of each formation could have been analysed and taken into account. As for pragmatic constraints, the function of each instance could have been analysed to determine, e.g., whether *-ity* was indeed used more for labelling than syntactic recategorisation, and what (if any) was the role of the attitudinal function mentioned in 2.2.4 above.

As for the sociolinguistic side, the category of age would be an interesting addition to the study. How would the diachronic change show up in different age groups? Another possible avenue of study would be the influence of the relationship between the sender and the recipient of the letter, which has been coded into the CEEC (see Section 6.2). Furthermore, all sorts of micro-level studies of selected individuals and their social networks might be of interest, if there were enough data. The discussion of possible future research will be continued in the next section.

9. Conclusion

First, Section 9.1 presents a summary of the main points of the thesis. In 9.2, the focus is widened for a consideration of the implications of the study for future research.

9.1. Summary

The main purpose of this study has been to find out whether there is sociolinguistic variation in the morphological productivity of the roughly synonymous nominal suffixes *-ness* and *-ity* in personal letters of the 17th century. As hypothesised, such variation is indeed observable in the productivity of the 'learned' suffix *-ity*, while the default suffix for forming abstract nouns from adjectives, *-ness*, shows no significant variation. The productivity of *-ity* is found to be significantly low in letters written by women, as well as in letters written during the period 1600–1639.

Women's lower productivity is explained by their restricted access to education, which was then necessary for a full command of the intricacies of *-ity*; it is probable that the lowest ranks would also exhibit a lower productivity for the same reason if there were enough data from them. The variation over time can be interpreted as linguistic change in progress: perhaps *-ity* spreads from the literate registers in which it first appeared to the more speech-like letters during the 17th century, or an increase in its use in literate registers shows up with a delay in more speech-like registers. The change may have been accelerated in the 1640s by the Civil-War effect, as there was much more contact between different kinds of people and an increase in weak social ties during the war.

The second focus of the study has been on methodology. In collaboration with researcher Jukka Suomela of the Helsinki Institute for Information Technology HIIT, a little-known solution has been presented to the problem of comparing type counts obtained from (sub)corpora of varying sizes (see Säily and Suomela forthcoming, Suomela 2007). Based on type accumulation curves and the statistical technique of permutation testing, the method is an assumption-free, highly visual way of determining whether a subcorpus is significantly different from the corpus as a whole, in terms of either the number of types or hapax legomena. The latter

measure, however, was shown to be unpractical at least in the corpus used in this study, as the upper and lower bounds for hapaxes turned out to be too wide for any significant differences to emerge. Therefore, the results mentioned above were obtained with the measure of type counts. With the help of this method, it has been possible to gain linguistically interesting and statistically significant results even though the amount of data has been relatively small, c. 1.4 million words divided into various subcorpora.

9.2. Implications for future research

To begin with methodology, the method presented in this study would certainly merit wider use, especially in historical linguistics, where un-lemmatised corpora and the scarcity of data have hitherto often prevented a proper statistical treatment of the results. Furthermore, as noted in Section 8.1, the problem observed with hapax legomena should be tested in larger corpora as well as corpora representing different registers. If persistent, it could call into question the use of productivity measures such as Baayen's (e.g., 1993) P and P*, which are based on hapaxes. On an interdisciplinary note, the same method could be applied to ecological studies of species richness (cf. Gotelli and Colwell 2001).

As mentioned in Section 8.2, the sociolinguistic portion of the study could be enlarged by taking into account the category of age, as well as some of the more micro-level aspects such as sender–recipient relationships and social networks. The morphological portion could be improved through a more systematic inspection of the structural and pragmatic constraints affecting the productivity of the suffixes, in particular the functions of labelling and syntactic recategorisation. These functions seem to illustrate the fuzzy boundaries between morphology and lexis on the one hand, and morphology and syntax on the other.

To widen the scope somewhat, the study could be extended into the rest of the CEEC corpora, which together cover four centuries, or the period 1402–1800.[4] The 18th-century part contains much more data from the lowest ranks than the 17th century, so the question of rank variation could perhaps be answered conclusively for the 18th century. Moreover, the corpora would offer better possibilities for observing historical change or variation in the productivity of the suffixes, and thus an opportunity to check whether the explanation given in this thesis for the low productivity of *-ity* in 1600–1639 seems feasible.

For a more comprehensive picture of the variation and change in the productivity of *-ness* and *-ity*, similar studies could be done in corpora focussing on other registers. Considering the tentative explanation for the change in the productivity of *-ity*, material such as scholarly texts from the 16th and 17th centuries would be of especial interest. Furthermore, the range of suffixes could be extended to cover, e.g., all suffixes forming abstract nouns. Besides productivity, their lexical aspects could also be studied, e.g., by analysing the words most frequently used by different groups, as was briefly done in Section 7.1 above.

My final remark pertains once more to the issue of productivity. While comparing the productivity of a single affix across different (sub)corpora is significantly facilitated by the method presented in this thesis, comparing the productivity of different affixes remains problematic. In addition to developing better methods for comparing affixes, more studies are needed of both sociolinguistic and register variation in productivity; among other things, they are sure to enhance our understanding of the mechanisms of change.

---

[4] See <http://www.helsinki.fi/varieng/domains/CEEC.html>.

Acknowledgements

My biggest thanks are due to Jukka Suomela and Terttu Nevalainen for their tireless support, advice and collaboration. The CEEC team have my gratitude for allowing me to use their unpublished corpus and auxiliary materials. I would also like to thank Harald Baayen, the members of VARIENG and the audience at ICAME 28 for comments and discussions.

Bibliography

ADAMSON, S., 1989. With Double Tongue: Diglossia, Stylistics and the Teaching of English. In: M. SHORT, ed, *Reading, Analysing and Teaching Literature*. London: Longman, pp. 204–240.

ANSHEN, F. and ARONOFF, M., 1989. Morphological Productivity, Word Frequency and the Oxford English Dictionary. In: R. FASOLD and D. SCHIFFRIN, eds, *Language Change and Variation*. Amsterdam: John Benjamins, pp. 197–202.

ARONOFF, M., 1976. *Word Formation in Generative Grammar*. Linguistic Inquiry Monograph One. Cambridge, Massachusetts: The MIT Press.

ARONOFF, M. and ANSHEN, F., 1998. Morphology and the Lexicon: Lexicalization and Productivity. In: A. SPENCER and A.M. ZWICKY, eds, *The Handbook of Morphology*. Cambridge, MA: Blackwell Publishers, pp. 237–247.

BAAYEN, R.H., 1992. Quantitative Aspects of Morphological Productivity. In: G. BOOIJ and J. VAN MARLE, eds, *Yearbook of Morphology 1991*. Dordrecht: Kluwer Academic Publishers, pp. 109–149.

BAAYEN, R.H., 1993. On Frequency, Transparency and Productivity. In: G. BOOIJ and J. VAN MARLE, eds, *Yearbook of Morphology 1992*. Dordrecht: Kluwer Academic Publishers, pp. 181–208.

BAAYEN, R.H. and LIEBER, R., 1991. Productivity and English Derivation: A Corpus-Based Study. *Linguistics*, 29, pp. 801–843.

BAAYEN, R.H. and NEIJT, A., 1997. Productivity in Context: A Case Study of a Dutch Suffix. *Linguistics*, 35 (3), pp. 565–587.

BAAYEN, R.H. and RENOUF, A., 1996. Chronicling the Times: Productive Lexical Innovations in an English Newspaper. *Language*, 72 (1), pp. 69–96.

BARBER, C., 1976. *Early Modern English*. London: André Deutsch.

BAUER, L., 1983. *English Word-Formation*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

BAUER, L., 2001. *Morphological Productivity*. Cambridge Studies in Linguistics 95. Cambridge: Cambridge University Press.

BIBER, D., 1995. *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge: Cambridge University Press.

BIBER, D., CONRAD, S. and REPPEN, R., 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

BOLINGER, D.L., 1948. On Defining the Morpheme. *Word*, 4, pp. 18–23.

BŘEZINA, V., 2005. *The Development of the Prefixes* un- *and* in- *in Early Modern English with Special Regard to the Sociolinguistic Background*. Unpublished MA thesis, Faculty of Arts, Charles University in Prague.

CAWDRY, R., 1604. *A table alphabeticall conteyning and teaching the true writing, and vnderstanding of hard vsuall English wordes, borrowed from the Hebrew, Greeke, Latine, or French, &c. With the interpretation thereof by plaine English words, gathered for the benefit & helpe of ladies, gentlewomen, or any other vnskilfull persons. Whereby they may the more easilie and better vnderstand many hard English wordes, vvhich they shall heare or read in scriptures, sermons, or elswhere, and also be made able to vse the same aptly themselues.* London: Printed by I. R[oberts] for Edmund Weauer.

CEEC = *Corpus of Early English Correspondence*, 1998. Compiled by the Sociolinguistics and Language History project team (T. Nevalainen, J. Keränen, M. Nevala, A. Nurmi, M. Palander-Collin, H. Raumolin-Brunberg) at the Department of English, University of Helsinki.

<http://www.helsinki.fi/varieng/domains/CEEC.html>.

CEECS = *Corpus of Early English Correspondence Sampler*, 1998. Compiled by the Sociolinguistics and Language History project team (T. Nevalainen, J. Keränen, M. Nevala, A. Nurmi, M. Palander-Collin, H. Raumolin-Brunberg) at the Department of English, University of Helsinki. <http://www.helsinki.fi/varieng/domains/CEEC.html>.

CHAMBERS, J.K., 1995. *Sociolinguistic Theory: Linguistic Variation and Its Social Significance*. Language in Society 22. Oxford: Blackwell.

COWIE, C. and DALTON-PUFFER, C., 2002. Diachronic Word-Formation and Studying Changes in Productivity over Time: Theoretical and Methodological Considerations. In: J.E. DÍAZ VERA, ed, *A Changing World of Words: Studies in English Historical Lexicography, Lexicology and Semantics*. Amsterdam: Rodopi, pp. 410–437.

CULPEPER, J. and CLAPHAM, P., 1996. The Borrowing of Classical and Romance Words into English: A Study Based on the Electronic Oxford English Dictionary. *International Journal of Corpus Linguistics*, 1 (2), pp. 199–218.

DALTON-PUFFER, C., 1992. The Status of Word Formation in Middle English: Approaching the Question. In: M. RISSANEN, O. IHALAINEN, T. NEVALAINEN and I. TAAVITSAINEN, eds, *History of Englishes: New Methods and Interpretations in Historical Linguistics*. Berlin: Mouton de Gruyter, pp. 465–482.

DALTON-PUFFER, C., 1996. *The French Influence on Middle English Morphology: A Corpus-Based Study of Derivation*. Topics in English Linguistics 20. Berlin: Mouton de Gruyter.

EVERT, S. and BARONI, M., 2005. Testing the Extrapolation Quality of Word Frequency Models. In: P. DANIELSSON and M. WAGEN-MAKERS, eds, *Proceedings of Corpus Linguistics 2005*. The Corpus

Linguistics Conference Series 1. URL: <http://www.corpus.bham.ac.uk/PCLC/>.

GOTELLI, J. and COLWELL, R., 2001. Quantifying Biodiversity: Procedures and Pitfalls in the Measurement and Comparison of Species Richness. *Ecology Letters*, 4, pp. 379–391.

HAY, J., 2001. Lexical Frequency in Morphology: Is Everything Relative? *Linguistics*, 39 (6), pp. 1041–1070.

HAY, J. and BAAYEN, R.H., 2003. Phonotactics, Parsing and Productivity. *Italian Journal of Linguistics*, 15 (1), pp. 99–130.

HAY, J. and PLAG, I., 2004. What Constrains Possible Suffix Combinations? On the Interaction of Grammatical and Processing Restrictions in Derivational Morphology. *Natural Language & Linguistic Theory*, 22 (3), pp. 565–596.

HLAVA, M., 2005. *Aspects of Morphological Productivity: A Corpus-Based Study of Foreign and Native Prefixes in Early Modern English*. Unpublished MA thesis, Department of English, University of Vienna.

HOFFMANN, S., 2004. Using the *OED* Quotations Database as a Corpus — a Linguistic Appraisal. *ICAME Journal*, 28, pp. 17–30.

HUDSON, R.A., 1996 [1980]. *Sociolinguistics*. 2nd edition. Cambridge: Cambridge University Press.

KASTOVSKY, D., 1986. The Problem of Productivity in Word Formation. *Linguistics: An Interdisciplinary Journal of the Language Sciences,* 24 (3 [283]), pp. 585–600.

KEUNE, K., VAN HOUT, R. and BAAYEN, R.H., 2006. Socio-Geographic Variation in Morphological Productivity in Spoken Dutch: A Comparison of Statistical Techniques. In: J.-M. VIPREY, ed, *Proceedings of the 8th International Conference on the Statistical Analysis of Textual Data, vol. 2*. Besançon: Presses Universitaires de Franche-Comte, pp. 571–581.

LABOV, W., 1978 [1972]. *Sociolinguistic Patterns*. Oxford: Basil Blackwell.

LABOV, W., 1994. *Principles of Linguistic Change, Volume 1: Internal Factors*. Language in Society 20. Oxford: Blackwell.

LASLETT, P., 1965. *The World We Have Lost*. New York: Charles Scribner's Sons.

MARCHAND, H., 1969. *The Categories and Types of Present-Day English Word-Formation: A Synchronic-Diachronic Approach*. 2nd edition. Munich: C.H. Beck'sche Verlagsbuchhandlung.

MARLE, J. VAN, 1986. The Domain Hypothesis: The Study of Rival Morphological Processes. *Linguistics*, 24 (3), pp. 601–627.

MED = *Middle English Dictionary*. 2001 edition. Electronic version. URL: <http://ets.umdl.umich.edu/m/med/>.

MENDELSON, S. and CRAWFORD, P., 1998. *Women in Early Modern England, 1550–1720*. Oxford: Clarendon Press.

MILROY, L. and GORDON, M., 2003. *Sociolinguistics: Method and Interpretation*. Language in Society 34. Malden, MA: Blackwell.

NEVALAINEN, T., 1999. Early Modern English Lexis and Semantics. In: R. LASS, ed, *The Cambridge History of the English Language, III: 1476–1776*. Cambridge: Cambridge University Press, pp. 332–458.

NEVALAINEN, T. and RAUMOLIN-BRUNBERG, H., 1996. The Corpus of Early English Correspondence. In: T. NEVALAINEN and H. RAUMOLIN-BRUNBERG, eds, pp. 39–54.

NEVALAINEN, T. and RAUMOLIN-BRUNBERG, H., 2003. *Historical Sociolinguistics: Language Change in Tudor and Stuart England*. London: Pearson Education.

NEVALAINEN, T. and RAUMOLIN-BRUNBERG, H., eds, 1996. *Sociolinguistics and Language History: Studies Based on the Corpus of Early English Correspondence.* Amsterdam: Rodopi.

NEVALAINEN, T. and TIEKEN-BOON VAN OSTADE, I., 2006. Standardisation. In: R. HOGG and D. DENISON, eds, *A History of the English Language*. Cambridge: Cambridge University Press, pp. 271–311.

OED = *Oxford English Dictionary*. 2nd edition, 1989. OED Online. URL: <http://dictionary.oed.com>.

PLAG, I., 1999. *Morphological Productivity: Structural Constraints in English Derivation*. Topics in English Linguistics 28. Berlin: Mouton de Gruyter.

PLAG, I., 2003. *Word-Formation in English*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

PLAG, I., 2006. Productivity. In: B. AARTS and A. MCMAHON, eds, *Handbook of English Linguistics.* Oxford: Blackwell, pp. 537–556.

PLAG, I., DALTON-PUFFER, C. and BAAYEN, R.H., 1999. Morphological Productivity across Speech and Writing. *English Language and Linguistics*, 3 (2), pp. 209–228.

RAINER, F., 1988. Towards a Theory of Blocking: The Case of Italian and German Quality Nouns. In: G. BOOIJ and J. VAN MARLE, eds, *Yearbook of Morphology 1988*. Dordrecht: Foris Publications, pp. 155–185.

RAUMOLIN-BRUNBERG, H., 1998. Social Factors and Pronominal Change in the Seventeenth Century: The Civil-War Effect? In: J. FISIAK and M. KRYGIER, eds, *Advances in English Historical Linguistics (1996)*. Berlin: Mouton de Gruyter.

RENOUF, A., 2006. Tracing Lexical Productivity and Creativity in the British Media: The *Chavs* and the *Chav-Nots*. Lecture presented at the Department of English, University of Helsinki, 26 October.

RIDDLE, E.M., 1985. A Historical Perspective on the Productivity of the Suffixes *-ness* and *-ity*. In: J. FISIAK, ed, *Historical Semantics; Historical Word-Formation*. Berlin: Mouton de Gruyter, pp. 435–461.

ROBINS, R.H., 1990 [1967]. *A Short History of Linguistics*. 3rd edition. New York: Longman.

ROMAINE, S., 1982. *Socio-Historical Linguistics: Its Status and Methodology*. Cambridge: Cambridge University Press.

ROMAINE, S., 1985. Variability in Word Formation Patterns and Productivity in the History of English. *Papers from the 6th International Conference on Historical Linguistics*. Amsterdam: John Benjamins, pp. 451–465.

SÄILY, T., 2005. Use of the Suffixes *-ity* and *-ness* in Early English Letters: Was Gender a Factor? Unpublished seminar paper, Department of English, University of Helsinki.

SÄILY, T. and SUOMELA, J., forthcoming. Comparing Type Counts: The Case of Women, Men and *-ity* in Early English Letters. Submitted to *Proceedings of ICAME 28*.

SUOMELA, J., 2007. *Type and Hapax Accumulation Curves*. Computer program. URL: <http://www.cs.helsinki.fi/jukka.suomela/types/>.

SZMRECSANYI, B., 2005. Language Users as Creatures of Habit: A Corpus-Based Analysis of Persistence in Spoken English. *Corpus Linguistics and Linguistic Theory*, 1 (1), pp. 113–150.

TWEEDIE, F.J. and BAAYEN, R.H., 1998. How Variable May a Constant be? Measures of Lexical Richness in Perspective. *Computers and the Humanities*, 32, pp. 323–352.

WERMSER, R., 1976. *Statistische Studien zur Entwicklung des englischen Wortschatzes*. Schweizer Anglistische Arbeiten 91. Bern: Francke.

WRIGHTSON, K., 1993. *English Society, 1580–1680*. London: Routledge.

Appendix 1. List of the *-ness* and *-ity* types found in the corpus

H=hapax, B=baseless (*=strict), P=prefixed, A=1st attestation ante 1500

-ness

abruptness				*craziness*			
accurateness	H			*crossness*			
advisedness	H			*daintiness*	H		
aptness				*dangerousness*	H		
assuredness	H			*darkness*			A
averseness				*dearness*			A
backwardness	H			*deceitfulness*	H		
badness				*dejectedness*			
barbarousness				*depravedness*	H		
barrenness	H		A	*devoutness*	H		A
baseness				*dirtiness*	H		
beholdingness	H			*disorderliness*	H	P	
bigness			A	*distrustfulness*	H	P	
bitterness			A	*dizziness*	H		A
blindness			A	*doubtfulness*			
boldness			A	*drowsiness*	H		
briefness	H		A	*drunkenness*			A
brightness			A	*dryness*	H		A
briskness	H			*dullness*			A
business			A	*duskishness*	H		
calmness				*dutifulness*	H		
captiousness				*eagerness*			A
carefulness			A	*earnestness*			
carelessness			A	*ear-witness*		B	
chargeableness	H			*easiness*			A
cheapness	H			*emptiness*			
cheerfulness				*evenness*			A
churlishness	H			*exactness*			
clearness			A	*eye-witness*		B	
clownishness	H			*fairness*			A
coldness			A	*faithfulness*			A
comeliness			A	*faultiness*			
communicativeness	H			*feverishness*	H		
completeness				*fineness*	H		A
composedness	H			*firmness*			
consciousness				*fitness*			
contentedness	H			*fixedness*	H		
coolness			A	*fogginess*	H		
cordialness	H			*fondness*			A
costiveness	H		A	*foolishness*			A
courageousness	H		A	*forgetfulness*			A
covetousness			A	*forgiveness*			A
coyness	H			*forwardness*			

foulness			A
frankness	H		
freeness	H		A
freshness	H		A
frowardness	H		A
fruitfulness	H		A
fullness			A
generousness	H		
gentleness			A
gladness			A
godliness			
goodness			A
graciousness	H		A
gratefulness			
greatness			A
greediness	H		A
grievousness	H		A
grossness	H		A
happiness			
hardiness	H		A
hardness			A
harshness			A
hastiness			A
healthiness	H		
heaviness			A
heinousness			
helpfulness	H		
highness			A
hoarseness			A
holiness			A
hopefulness	H		
humbleness			A
idleness			A
illness			
imperviousness	H	P	
inactiveness	H	P	
incorrigibleness	H	P	
infiniteness	H	P	
inquisitiveness			
insensibleness	H	P	
inwardness	H		A
irksomeness	H		A
jolliness	H		A
justness			A
kindness			A
lameness			
largeness			A
lateness	H		A
lawfulness			A
laziness			
leanness			A
lewdness	H		A
lightness			A
likeness			A
listlessness	H		
litherness			A
loathness	H		A
loneliness	H		
looseness			A
loveliness	H		A
lowness			A
madness			A
maliciousness	H		A
meanness			
meekness			A
mildness			A
mindfulness			
monstrousness	H		
muchness			A
muddiness	H		
nakedness			A
nearness			A
niceness	H		
niggardliness	H		
nobleness			A
notoriousness	H		
numbness	H		
obstinateness	H		
obstructedness	H		
oddness			A
officiousness	H		
over-boldness	H	P	A
oversweetness	H	P	
painfulness	H		A
peevishness			A
peremptoriness	H		
perfidiousness	H		
perverseness	H		A
piousness	H		
plain-heartedness	H		

plainness		A		*silliness*	H		
pleasantness	H	A		*singleness*	H		
poroseness	H			*slackness*			A
preciousness	H	A		*slipperiness*	H		
preciseness	H			*slovenliness*	H		
princeliness	H			*slowness*			A
procliveness	H			*smallness*			A
profaneness				*smartness*	H		A
profuseness	H			*smoothness*			A
quickness		A		*soberness*	H		A
quietness		A		*softness*	H		A
rapidness	H			*solicitousness*	H		
rareness				*solitariness*			
rashness				*soreness*			A
readiness		A		*soundness*			A
reasonableness	H			*sourness*			A
redness	H	A		*sprightfulness*	H		
refractoriness				*stateliness*			
religiousness	H			*steadfastness*	H		A
remissness	H			*stomachfulness*	H		
remoteness				*straitness*			A
reservedness	H			*strangeness*	H		A
restiness	H			*strictness*			
retiredness				*stubbornness*			A
richness		A		*stupidness*	H		
ridiculousness	H			*subtleness*	H		A
righteousness		A		*suddenness*			A
rigidness	H			*sullenness*			
ripeness	H	A		*sumptuousness*	H		
rottenness	H	A		*sureness*			A
roughness	H	A		*sweetness*			A
rudeness		A		*talkativeness*	H		
ruggedness	H			*tediousness*			A
sadness		A		*tenderness*			A
saltness		A		*thankfulness*			
sauciness	H			*thickness*			A
schismaticalness	H			*thinness*	H		A
sensibleness				*towardliness*	H		
serviceableness	H			*unableness*	H	P	A
shallowness				*unactiveness*	H	P	
shamefastness	H	A		*unadvisedness*	H	P	A
sharpness		A		*uncouthness*	H	P	A
shortness		A		*undutifulness*	H	P	
sickliness				*uneasiness*	H	P	A
sickness		A		*unevenness*	H	P	A

unfaithfulness	H		P	A
unfeignedness			P	
unfitness	H		P	
unfittingness	H		P	
unfortunateness			P	
unhappiness			P	A
unkindness			P	A
unlawfulness	H		P	
unmercifulness	H		P	
unpassableness	H		P	
unpractisedness	H		P	
unpreparedness	H		P	
unquietness			P	
unreasonableness	H		P	
unruliness	H		P	
unsatisfiedness	H		P	
unseasonableness			P	
unskilfulness	H		P	A
unthankfulness			P	
unthriftiness	H		P	A
untowardness			P	A
unwellness	H		P	
unwillingness			P	
unworthiness			P	A
uprightness				
usefulness	H			
uselessness	H			
vileness	H			A
voluntariness	H			
voluptuousness	H			
wantonness	H			A
wariness	H			
watchfulness				
weakness				A
weariness				A
wearisomeness				
wellness	H			
wickedness				A
wideness	H			A
wilderness		B		A
wildness	H			A
wilfulness				A
willingness				
witness				A
worldliness	H			A
worthiness				A
wretchedness				A
yellowness	H			A

-ity

ability				A
absurdity	H			
activity				
adversity				A
affinity			P	A
alacrity		B		
ambiguity		B		A
amity		B		A
animosity				A
annuity		B		A
antiquity				A
anxiety		B		
austerity	H			A
authority				A
barbarity	H			
benignity				A
brevity		B		
calamity		B		A
capability	H			
capacity				A
captivity				A
carnality	H			A
charity		B		A
christianity				A
civility				A
commodity				A
conformity			P	A
consanguinity	H		P	A
contrariety				A
credulity		B		A
curiosity				A
debility	H			A
deformity	H		P	A
deity		B		A
depravity			P	
dexterity				
dignity				A
disability			P	
disparity	H	B*	P	
diuturnity	H			A
diversity				A
divinity				A
enormity				A
entity		B*		

equality				A
equity		B		A
eternity				A
extremity				A
facility				
fallibility	H			
falsity				A
familiarity				A
felicity		B		A
ferocity	H	B*		
fertility	H			A
festivity	H			A
fidelity				A
firmity	H			A
formality				
fraternity	H	B*		A
gaiety				
generality				A
generosity				A
gentility				A
gratuity		B*		
gravity				
hospitality				A
hostility				
humanity				A
humility		B*		A
imbecility	H			
immortality			P	A
immunity				A
impartiality	H		P	
impenetrability	H		P	
impiety		B	P	A
importunity				A
impossibility			P	A
improbability	H		P	
inability		B	P	A
inactivity			P	
incapacity		B	P	
incivility			P	
incommodity			P	A
incongruity		B*	P	
incredulity		B*	P	A
indemnity		B*		A
indignity			P	

inequality			P	A
infallibility	H		P	
infidelity			P	
infinity		B*	P	A
infirmity			P	A
infortunity	H		P	A
ingenuity		B		
inhumanity			P	A
iniquity			P	A
insecurity			P	
insensibility			P	
instability	H		P	A
integrity				A
invalidity	H		P	
irregularity			P	A
jollity				A
laity				
Latinity	H			
legality				A
lenity		B		
liberality				A
magnanimity				A
malignity				A
maturity				A
mediocrity	H			A
minority				A
morality				A
mortality				A
multiplicity		B		A
mutability				A
mutilletie	H	B		
mutuality	H			
nativity				A
necessity		B		A
neutrality				A
nicety	H			A
nobility				A
nonconformity	H		P	
nonentity		B	P	
nullity				
obdurity	H		P	
obliquity	H			A
obscurity				A
opportunity				A
parity	H			
partiality				A
particularity				
perplexity				A
perspicacity	H	B		
piety		B		A
plurality	H			A
popularity	H			
possibility				A
posterity		B		A
priority		B		A
privity		B*		A
probability				
prodigality				A
profundity	H			A
prolixity	H			A
propriety		B		A
prosperity		B*		A
proximity		B*		A
punctuality				
purity				A
quality		B		A
quantity		B		A
quiety	H	B*		A
rarity				
reality				
rusticity	H			
sagacity	H	B		
sanctity	H	B		A
satiety	H	B		
scarcity				A
security				A
seniority				A
sensuality				A
serenity				A
severity				A
simplicity		B*		A
sincerity				
singularity	H			A
sobriety		B		A
society		B		
solemnity				A
solidity	H			
stability	H			A
stupidity				
superfluity		B*		A

superiority	H			
temerity		B		A
temporality	H			A
tranquillity				A
trinity				A
ubiquity	H	B		
unanimity	H			A
uncharity	H	B	P	
uniformity	H		P	A
unity		B		A
universality	H			A
university				A
validity				
vanity				A
variety		B*		
verity	H			A
virginity	H			A