Variation in the productivity of adjective comparison in Present-day English

Tanja Säily, Victorina González-Díaz, Jukka Suomela

This is an Accepted Manuscript of a book chapter published by Routledge in *Corpus Approaches to Contemporary British Speech: Sociolinguistic Studies of the Spoken BNC2014* on 31 May 2018, available online: http://www.routledge.com/9781138287273

1. Introduction

English adjective comparison has received a great deal of attention in corpus-based research, particularly in the functional competition between inflectional (*-er*) and periphrastic (*more*) strategies (e.g. Mondorf 2003, González-Díaz 2008, Matsui 2010). There is, however, a key area of competition that remains relatively unexplored, namely, the productivity of either comparative strategy. The received wisdom is that inflectional affixes are fully productive, which would suggest lack of variation within the productivity of *-er*. However, recent research using novel methodologies (Säily 2014) crucially shows *sociolinguistic* variation in the productivity of extremely productive derivational suffixes. Whether the same variation applies to the productivity of inflectional processes remains therefore an open question.

Our study explores intra- and extra-linguistic variation in the productivity of comparative strategies. Intra-linguistic factors include syntactic position, the presence of premodifiers, complements and a second term of comparison, and the length of the adjective. The extra-linguistic determinants focus on gender, age, socio-economic status, conversational setting and roles of the interlocutors. Rather than limiting ourselves to the relatively small number of adjective types in which both inflectional and periphrastic comparison can occur, we take a holistic approach and consider the entire range of types within each strategy using the methodology and software recently developed by Säily and Suomela (2009) and Suomela (2016) for the study of derivational productivity (*types*2). The Spoken BNC2014 (Love et al. 2017) is instrumental to the project, as it is the only up-to-date corpus of Present-day English (PDE) providing access to both intra- and extra-linguistic information across a representative sample of British society. To provide our study with a diachronic dimension, we will compare the corpus with the original *British National Corpus* (BNC1994).

Our research constitutes a timely contribution to current knowledge of adjective comparison and morphological theory-building. It not only provides greater descriptive adequacy as regards the factors shaping the growth and development of the English comparative system, but also deals with much-debated issues concerning analytic vs. synthetic trends in the history of English. Past empirical work on analyticity and syntheticity has often excluded the study of derivational morphology (e.g. Szmrecsanyi 2012), partly because of the strict compartmentalisation of

inflectional and derivational morphology. The idea that there is a derivation-to-inflection cline rather than a sharp divide between the two has of course been expressed in previous literature (e.g. Brinton and Traugott 2005; Bauer 2004). However, more empirical evidence is still needed to support this hypothesis. If our results were to show variation and change in the productivity of inflectional comparison (as an example of inflectional morphology) similar to that previously observed in derivational morphology (e.g. Säily 2014), then this would provide further support for the 'cline' view. Furthermore, the existence of a cline would mean that *both* derivation and inflection contribute to syntheticity, which is also the view expressed by Danchev (1992).

The chapter is organised as follows. Section 2 summarises previous research on adjective comparison and morphological productivity. Section 3 describes the material and methods used. Sections 4 and 5 focus on the data analysis and discussion of the findings, whereas Section 6 considers the theoretical implications of the study.

2. Background

2.1. English adjective comparison

Adjective comparison has been a long-standing topic of interest in English linguistics from the early 20th century onwards. Lexicographical works of the early 20th century (e.g. Poutsma 1914, Curme 1931, Jespersen 1949) provide a description of the two structures available in PDE (inflectional comparison – e.g. *friendlier* – and periphrastic comparison, e.g. *more friendly*) and of the main factors governing the choice of one (inflectional) or the other (periphrastic) strategy (see also Quirk et al. 1985 or Huddleston and Pullum 2002). These factors are normally of a phonological and morphosyntactic nature, i.e. the number of syllables of the adjective (adjectives of more than three syllables normally take periphrastic comparison and monosyllabic ones prefer the inflectional form) and/or its ending and stress pattern (e.g. *-ive, -ous-* or *-ful* adjectives normally take periphrastic comparison). Another frequently-discussed issue in early lexicographical works is the origin of the periphrastic construction (e.g. Mossé 1952, Strang 1970; but see also Knüpfer 1921).

Scholarly interest in adjective comparison grew in the second half of the twentieth century, especially in works couched within the generative tradition. These studies often focused on the development of theoretical models that could account for the semantics and syntax of comparatives in an efficient manner (see Cresswell 1976, Rusiecki 1985 or, more recently van Rooij 2010, 2011), as well as on the constraints leading to the derivation of comparatives (e.g. Huddleston 1967, Bresnan 1977). Controversies about the nature of over-generalisations also sparked an interest in the acquisition and use of comparatives (see Gathercole 1979, 1985; Gitterman and Johnston 1983 or, more recently, Graziano-King 1999; Graziano-King and Cairns 2005; Hohaus et al. 2014).

The growth of computer-based English linguistics in the late 1980s elicited new synchronic and diachronic interest in the comparative system. On the synchronic front, recent corpus-based scholarship has primarily focused on the factors governing the functional distribution of inflectional and periphrastic strategies in PDE. Thus, alongside length and ending, a number of other prosodic, syntactic, semantic and cognitive-pragmatic determinants of variation have been put forward, e.g. syntactic position and presence of premodifiers, complements and a second term of comparison (syntax); concrete vs. abstract meanings (semantics); frequency of adjectival use, complexity of the context and previous mention in discourse (cognitive-pragmatic) (see, among others, Leech and Culpeper 1997; Lindquist 2000; Mondorf 2003, 2007, 2009; Szmrecsanyi 2005; González-Díaz 2008; Hilpert 2008, 2010). Although some discrepancies can be found as regards how influential each of these factors is, these works present a unified picture as far as they show that "the true extent of variability in this area appears to have been underestimated in the past" (Mondorf 2009, xiii).

Diachronic studies of adjective comparison are relatively less numerous. Some concomitances can nevertheless be found across works: namely, an interest in ascertaining possible usage trends in comparative strategies across time (Kytö 1996 and Kytö & Romaine 1997, 2006 observe a tendency for the inflectional strategy to prevail over the periphrastic one over time), and a willingness to further explore previous claims on the origin of the periphrastic construction. In addition, accounts of genre-based distribution of comparatives as well as socio-stylistic analyses of non-standard comparative strategies (i.e. double comparison e.g. *more friendlier*) are also found (see Kytö 1996; González-Díaz 2004, 2006, 2008; Mondorf 2009).

The brief outline above attests to the wealth of research already produced on the English comparative system. There are, however, particular aspects where the application of recent developments, be they corpora or methodologies, may lead to new insights about comparison. Corpus-wise, BNC2014 constitutes a valuable resource to confirm and/or reject previously observed shifts and trends in the recent history of the British English comparative. On the methodological front, previous corpus-based research on the competition between inflectional and periphrastic forms has traditionally focused on a small number of adjective types (normally disyllabic adjectives) which can take both comparative strategies. New software such as the *types*2 program used in this paper allows for reliable comparisons of adjective types regardless of their default comparative preferences. This in turn opens up new avenues of investigation – in our case, it allows us to study variation within the productivity of either comparative strategy across internal and external factors of change. Note, in this connection, that although the received wisdom is that inflection is fully productive – and consequently we would expect to find no variation within the productivity of synthetic *-er* comparatives – the existence of sociolinguistic variation in the productivity of extremely productive derivational suffixes has been observed (see Säily 2014). Whether the same variation applies to the productivity of inflectional processes remains an open question, which we will pursue in the following sections.

2.2. Morphological productivity

The concept of morphological productivity is difficult to define. Bolinger (1948, 18) refers to a "degree of animation" as an essential property of morphemes that amounts to a "statistically determinable readiness with which an element enters into new combinations". Baayen (e.g. 2009, 1992, 1993) specifies three aspects of productivity: realised, expanding and potential productivity, measured in different ways. The measures are based on the frequencies of tokens (all words containing the affix or morphological category in question), types (different words containing the affix) and hapax legomena (words occurring only once) in a corpus. **Realised productivity** is measured in type frequency: the number of different words formed using the affix estimates how the productivity has been realised up to the point or period in time represented by the corpus. **Potential productivity** is measured as the proportion of hapax legomena containing the affix out of all tokens containing the affix: this assesses the growth rate of the category. Put simply, the reasoning behind this is that hapax frequency approximates the number of new types, as it is among hapax legomena that most new types are found (Baayen 1993, 189). Finally, **expanding productivity** is measured as the proportion of hapax legomena containing the affix out of all hapax legomena in the corpus: this estimates the rate at which the category is expanding relative to the overall lexicon. When comparing the productivity of affixes, they may be ranked in a different order depending on the measure used, as the measures represent different facets of productivity. Where possible, therefore, all three measures should be taken into account when estimating productivity; however, it has been shown that measures based on hapax legomena yield unreliable results in smaller corpora, including the demographically sampled spoken subcorpus of BNC1994 (Baayen 1993; Säily and Suomela 2009; Säily 2011; see further Section 3.2 below).

Studies of morphological productivity have tended to focus on derivation (e.g. Plag 1999). Inflection is traditionally regarded as more productive than derivation (Stump 1998, 16) and has even been claimed to be fully productive (Haspelmath 2002, 75; Plag 2003, 16). Previous sources, however, make further distinctions between contextual (syntax-dependent) and inherent inflection (which is not required by syntax; cf. e.g. Booij 1996, 2–3). Examples of inherent inflection include number-marking in nouns, or comparative and superlative degree for adjectives. Booij (1996) argues that inherent inflection is more similar to derivation and may be subject to similar lexicalisation tendencies and constraints on productivity. To give an example, some nouns in English cannot be pluralised because their semantics does not allow for the addition of an affix meaning 'more than one instance of'. Along the same lines, Gaeta (2007) studies a range of derivational and inherently inflectional affixes in Italian and shows not only that the potential productivity of the affixes forms a cline from derivation to inflection but also that inflectional productivity varies across both affixes and entire inflectional categories. The study also suggests that some of this variation may be due to competing periphrastic forms.

Productivity, then, is not just a property of morphology. As recent work within Construction Grammar shows, productivity can be applied to different types of form–content pairings ('constructions') at different levels of granularity (see Perek 2016; Zeldes 2009, 2013; Hilpert 2013). Furthermore, as noted above, productivity may be either constrained or influenced by various factors. Plag (2006, 550–51) lists a number of structural factors, which are mostly process-specific and include phonological, morphological, syntactic and semantic constraints. External factors have chiefly been studied with respect to derivational productivity and focus on pragmatic, stylistic and sociolinguistic considerations (the latter including gender, age, region, education, socio-economic status, and register in terms of participant relations; see Štekauer et al. 2005; Keune et al. 2006; Schröder 2008; Palmer 2009; Gardner 2013; Säily 2014). Even though social factors were already regarded as important by Romaine (1983), they have only recently started to be studied in more detail; moreover, the extent to which these factors are applicable to non-derivational kinds of productivity is yet to be determined.

Our goal is therefore to extend previous work that has challenged the strict separation of derivation and inflection. While Gaeta (2007) found a cline between the productivity of inflectional and derivational affixes, we wish to see if there is variation and change within the productivity of individual inflectional affixes (in our case, in the use of the *-er* comparative strategy) and if so, whether it is similar to that discovered in derivational affixes (e.g. Säily 2014). In this respect, we are particularly interested in whether the productivity of inflectional comparison may be constrained by external factors of change. We will also study productivity beyond morphology by including in our analysis the periphrastic comparative construction [*more* +ADJ]. Although our work cannot be couched within any specific theoretical framework, it will pay attention to both internal and external factors as both are considered crucial to understanding language variation and change (cf. the embedding problem presented by Weinreich et al. 1968).

3. Materials and methods

3.1. Materials

Our data comes from two main sources, i.e. the spoken subcorpus of BNC1994 and the Early Access Subset of the newly compiled Spoken BNC2014 (2012–2015). BNC1994 is a c. 100 million word compilation of (mainly) British English of the second half of the 20th century (1960s–1990s). Its spoken subcorpus amounts to c. 10 million words and consists of "orthographic transcriptions of unscripted informal conversations … and spoken language collected in different contexts, ranging from formal business or government meetings to radio shows and phone-ins" (http://www.natcorp.ox.ac.uk/corpus/; last accessed 30/01/2017). The Spoken BNC2014 corpus will, when completed, also be a 10 million word compilation of "informal, spoken interactions between speakers of British English from across the United Kingdom" (http://cass.lancs.ac.uk/; last accessed 30/01/2017). At present, it contains c. 5 million

words. The reason behind this dual-corpus choice was not only to increase the robustness of our claims (as noted in the previous section, we adopt a holistic approach to type analysis) but also to allow for a short-term diachronic dimension in our study.

We considered all of the BNC2014 data made available to us (see further below). From BNC1994, however, we took a random sample of 500 speakers from the 'demographically sampled' part of the collection (recorded c. 1991–1993), which is the part of the BNC1994 subcorpus that matches best the new BNC2014 (meta)data and therefore allows for greatest reliability of short-term diachronic comparisons. Overall, our dataset comprises c. 6 million words: 1.33 million words from BNC1994, and 4.76 million words from BNC2014. In the rest of this paper, we will refer to these subsets as BNC1994-s and BNC2014-s, respectively.

3.2. Methods

The data was retrieved via CQPweb (Hardie 2012) using a combination of lemma and POS tag searches and then post-processed with Python scripts; we used the search term `*er_AJC` or `*er_JJR` for inflectional comparatives (e.g. *friendlier*, *colder*) and `more *_AJ0` or `more *_JJ` for the periphrastic ones (e.g. *more friendly*, *more interesting*).[1] The initial dataset was later manually checked and pruned down to relevant types and tokens. Discarded examples comprised:

- Incorrectly tagged lexemes. For instance, the nouns *lighter* or *cooler* (e.g. *do you have a **lighter**?*) or verbs such as *lower* (e.g. ***lower** the tax rate*).
- Instances where the relevant comparative forms are part of a set phrase or expression (e.g. *the **bigger** the **better**; it makes it all the **more serious***). In these cases, the productivity of the individual lexical element may be compromised by its function within a bigger unit.
- Cases where the comparative is part of a nominal, adjectival or verbal lexeme (e.g. ***higher** education*, *the **greater** good*).
- Examples featuring comparative forms in adverbial functions (e.g. *I cycle so much **faster** / you could have delivered **funnier** that kind of thing*), as well as cases where *more* performs determinative functions (e.g. *we need **more** modern literature in the department / I have **more** grey hair than she does*).
- Double comparative tokens (e.g. *He is **more poorer** than X*). The frequency of these structures was, across BNC compilations, too low to draw any meaningful results from the analyses.

[1] Note that the POS tagset varies slightly across compilations: for example, `AJ0` is the POS tag for a general adjective not inflected for degree in BNC1994 (see the CLAWS 5 POS tagset); the corresponding tag in BNC2014 is `JJ` (see the CLAWS 6 POS tagset) For more information on the different CLAWS tagsets, please see http://ucrel.lancs.ac.uk/claws/.

- Cases where the speaker hesitates and repeats the comparative (e.g. *it is **easier** it is **easier** to see*). The first comparative token of the repetition was systematically discarded from the tally (i.e. see underlined *easier* above), the second (repeated) form was included.

In line with previous literature, the relevant tokens were classified according to a set of intra- and extra-linguistic factors. Intra-linguistically, the data was coded for the following well-established morpho-syntactic considerations:

- Number of syllables of the positive adjective
- Syntactic position (attributive/predicative/postpositive comparatives)
- Presence/absence of a second term of comparison (i.e. a *than* phrase)
- Presence of adverbial premodifiers (e.g. ***much/a bit/a lot*** more friendly)
- Presence of complements (e.g. *he was more aware **[of X]*** / *I am a lot more careful **[with X]***)

The external variables considered included the **gender**, **age** and **social class** of the speaker. Speakers whose gender and/or age group was unknown were discarded from the tally. Note also that some slight mismatches were observed in the age and social class categories used across BNC compilations. To elaborate, BNC1994 works with the age groups [0–14], [15–24], [25–34], [35–44], [45–59], [60+], whereas BNC2014 prefers the categories [0–10], [11–18], [19–29], [30–39], [40–49], …, [90–99]. To make the categories more comparable, we collapsed them into three groups for both corpora: [0–24], [25–44], [45–99] for BNC1994, and [0–29], [30–49], [50–99] for BNC2014. Similarly, the social class distinctions were somewhat different in the two corpora, but we were provided with an automatic mapping from the BNC2014 categories to the BNC1994 categories, which are based on Social Grade, determined by the occupation of the head of the household (National Readership Survey 2015). For ease of analysis, we collapsed the categories into two groups: middle class (A+B+C1) and working class (C2+D+E). In BNC2014, we also had access to the external factor of **education**: secondary school, college/sixth form and university. Regional variation was not examined as our initial explorations did not yield (socio)linguistically interesting results.

We also considered the influence of **register**, in order to check whether our sociolinguistic results remained the same when focusing on speech in a similar setting. In particular, we were interested in everyday discussions among family and close friends at home, as this is the setting that represents the widest range of social groups (as opposed to e.g. 'work', which only represents those who work and is male-skewed at least in BNC1994). Again, the differences in the corpus metadata made such comparisons somewhat challenging: in BNC1994 we had access to the locale of the conversation, while in BNC2014 locale was not directly available but there was information on the relationship between the speakers. Hence for register-specific comparisons, in BNC1994 we focused on conversations that took place at home (e.g. 'home', 'kitchen', 'bedroom') and in BNC2014 we focused on conversations among close family, partners, and very close friends.

In our quantitative analysis of productivity, we focused on realised productivity, as measured by **type frequency** (for a justification of this, see Säily 2014, 238–239).[2] For example, to assess whether there was change in the productivity of inflectional adjective comparison in time, we initially counted the number of different adjectives used in such comparisons in both BNC1994-s and BNC2014-s and compared these numbers with each other. Similarly, to study sociolinguistic variation in productivity, we first considered similar comparisons across subcorpora that represented different social categories. Here we faced three methodological challenges:

- How to compare type frequencies among **corpora of different sizes**? If we observe 100 types in a corpus with 1M words, how many types would we expect to see in a corpus with 2M words?
- How large a difference in type frequencies is **statistically significant**? Could the findings be explained by mere random chance?
- How to deal with **outliers**? How do we know if our findings are representative of the entire social category, and not e.g. the influence of a speaker's idiolectal preferences.

To address these challenges, we first divided the corpus into relatively large **samples** so that e.g. individual idiosyncrasies or one-off events only influenced a small number of samples. In our case, we made the choice that one sample corresponded to all utterances of **one speaker in one setting** (in total 2,800 samples in BNC1994-s and 1,493 samples in BNC2014-s). This allowed us to unambiguously associate both speaker-specific metadata (e.g. age and social class) and register-specific metadata (e.g. locale) with each sample. Then we constructed a large number of **random re-orderings** (permutations) of the samples, in order to learn the typical relationship between type frequency and token frequency in the corpus under study. This information can be represented as a **type accumulation curves** (see Figure 1).

[2] In brief, it could be argued that by only using type frequency and leaving out hapax-based measures, we end up measuring lexical diversity rather than productivity in the sense of 'readiness to enter into new combinations'. However, measuring change in type frequency between two time periods does provide us with access to formations that are 'new' in the more recent time period in the corpus. While many of the 'corpus-new' formations may not be new to the language of the community as a whole, they may be productive uses in that they are not stored – or are only weakly activated, cf. Baayen (1993) – in the speaker's mental lexicon, so that the speaker produces the formation from its more strongly activated components in the speech situation (e.g. *posh* + *-er* → *posher*). This is more likely to happen if the formation is much less frequent than its base (e.g. *posh*: 396 instances in the entire BNC1994, *posher*: 17 instances; cf. Hay and Baayen 2002). Hence, many types that are not 'corpus-new' can also be formed productively. We would therefore argue that even though it is clearly imperfect, type frequency can still be a useful measure of productivity.

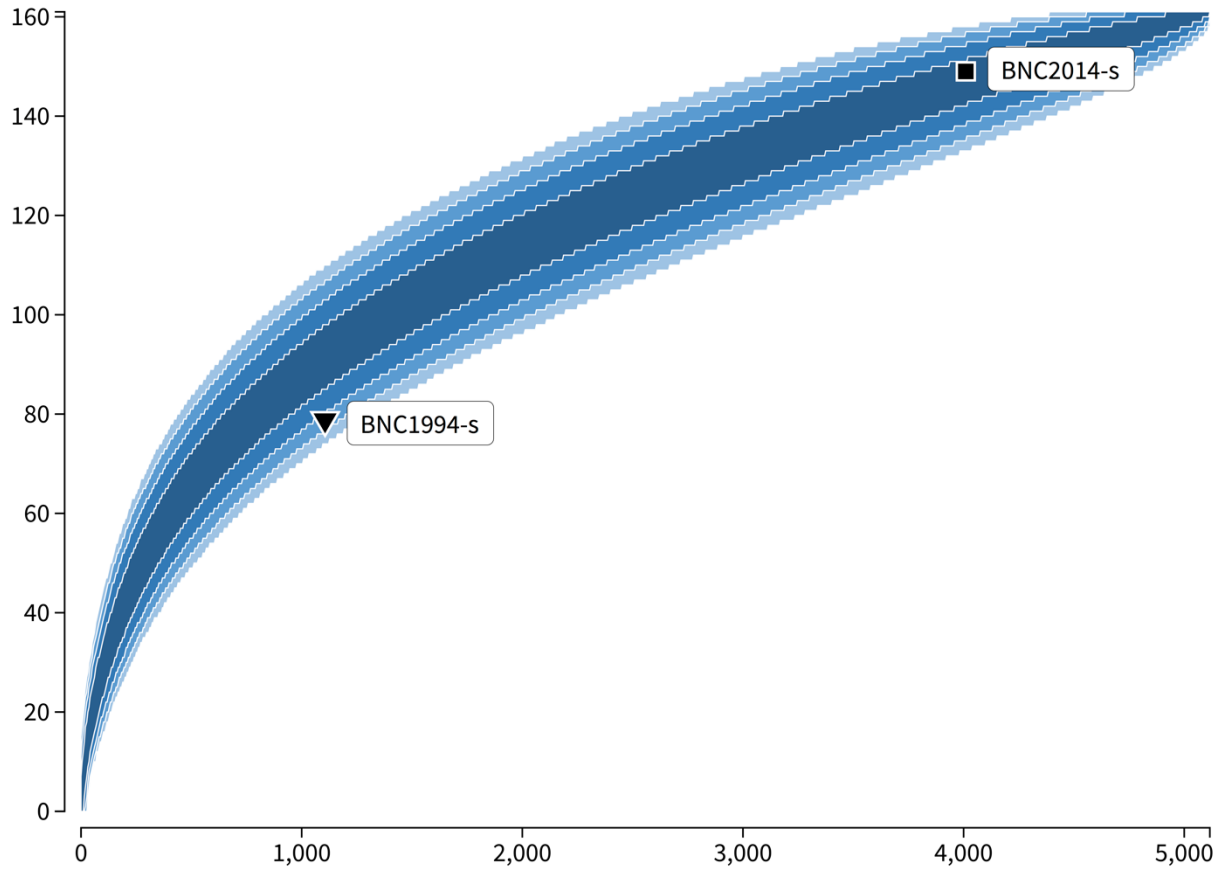


Figure 1. Type accumulation curve: type frequency (*y* axis) vs. token frequency (*x* axis). Inflectional adjective comparison in the entire corpus (BNC1994-s + BNC2014-s). Dark shading indicates the range of the most common values.

Figure 1 presents the type accumulation curve for inflectional adjective comparison in the entire dataset that consists of both BNC1994-s and BNC2014-s. The *x* axis shows the token frequency, while the *y* axis records the type frequency. The darker shading indicates more typical values; for example, a random collection of samples with 1,000 tokens is expected to contain c. 80–95 types, whereas a collection with 2,000 tokens is expected to contain c. 105–120 types.

The method allows us to pick a subcorpus of interest – e.g. the subcorpus that consists of BNC1994-s only, or the subcorpus of female speakers – and compare the number of types and tokens in it with the overall type accumulation curve. We can also zoom in on internal factors: to give an example, we can calculate type and token frequencies for only those adjective comparisons with adverbial premodifiers and see if similar variation between subcorpora is present in this dataset (see Section 4 below). Note, however, that in order to compare internal factors with each other, a slightly different (and cruder) approach was taken: we split the corpus into samples that consisted of just one occurrence of adjective comparison. Here, more care was

needed in the interpretation of the results, as a single speaker was potentially represented by a large number of samples.[3]

Overall, the three above-mentioned challenges were met in that:

- We do not need to compare e.g. BNC1994-s and BNC2014-s with each other; we can compare e.g. BNC1994-s with random collections of samples from BNC1994-s and BNC2014-s that happen to have the same token frequency as BNC1994-s.
- We can directly assess the statistical significance of the findings (see Section 4.1 for a concrete example). In essence, we are testing hypotheses using the statistical technique of permutation testing.
- The findings are robust to outliers: a small number of highly atypical samples cannot have much influence on the findings.

Some computer programs were naturally needed to conduct this kind of study. We imported our data to the ***types*2 software** (Suomela 2016), which takes care of the numerical calculations related to the permutation testing, and produces interactive visualisations that can be used to explore type and token frequencies in different subcorpora. The tool also takes care of the issue of **false discovery rate** (FDR) control in studies in which we test a large number of hypotheses. For more details on the methodological background and on the software that we use, see Säily and Suomela (2009, forthcoming).

4. Analysis

4.1. Overview

Let us first examine the overall type and token frequencies of inflectional and periphrastic comparison in our two corpora. In BNC1994-s, we find 78 types and 1,106 tokens for inflectional comparison, and 79 types and 114 tokens for periphrastic comparison.[4] As periphrastic comparison achieves a similar number of types to inflectional comparison in a much lower number of tokens, it is clearly the more productive of the two strategies. However, the difference between the strategies is much less obvious when we focus on disyllabic adjective types (where the two strategies alternate), and we are unable to determine which of the two strategies is more productive in this case.[5] In BNC2014-s, we find 149 types and 4,010 tokens

[3] In particular, the *p*-values reported in Tables 1 and 2 were derived with token-level samples.

[4] We have included the highly frequent inflectional type *better*, which has no base, in our analyses. If *better* is removed, the results become similar but weaker.

[5] Not all disyllabic bases can take both kinds of comparison (e.g. disyllabic adjectives ending in *-ful* or *-ous* do not normally take inflectional comparison; e.g. *?cautiouser, ?carefuller*). In order to make our results fully comparable, we could have further restricted the bases under analysis. In this paper, however, we are more interested in variation and change *within* the productivity of either strategy as a whole and, therefore, such considerations were discarded.

for inflectional comparison, and 525 types and 1,249 tokens for periphrastic comparison, making the latter again clearly more productive. This time similar results are also obtained for disyllabic types.

Figure 1 shows the type accumulation curve for **inflectional comparison** in BNC1994-s and BNC2014-s combined, plotting each corpus on the curve. BNC1994-s has few types compared to randomly sampled subcorpora of the same size: it contains 1,106 tokens and 78 types, and only c. 0.32% random collections of samples of this size have such a low number of types. The significantly low productivity of inflectional adjective comparison in BNC1994-s ($p < 0.0032$) implies an **increase in the productivity** of inflectional comparison over time, which will be studied further in Section 4.2 below.

As regards **periphrastic comparison**, a noticeable **increase in token frequency** is observed, i.e. from 85.5 instances per million words in BNC1994-s to 262.6 in BNC2014-s. Moreover, this change does not seem to be due to a small number of outliers, but it is representative of the whole corpus (see Figure 2). However, the results do not record any significant change in the productivity, or type frequency, of periphrastic comparison.

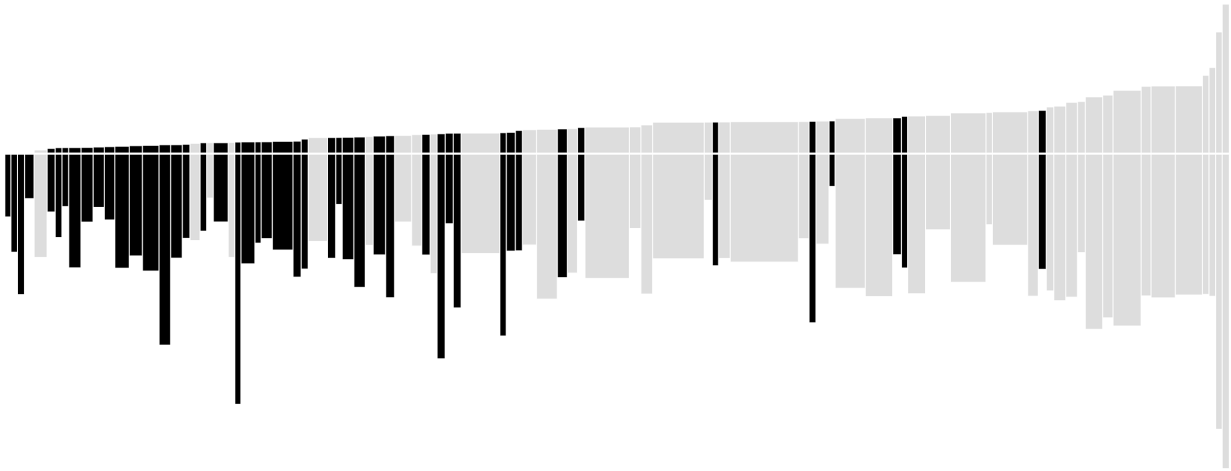


Figure 2. Speakers with at least 20,000 words in the corpus; black bars represent speakers in BNC1994-s (this also includes speakers outside our random sample of 500) and grey bars represent speakers in BNC2014-s. The top part indicates the relative frequency of *more* followed by a general adjective not inflected for degree (typically periphrastic adjective comparison), while the bottom part indicates the relative frequency of other occurrences of *more*. The speakers are ordered by the frequency of *more* + adjective; clearly most speakers from BNC1994-s have a low frequency, while most speakers from BNC2014-s have a high frequency.

Another question of interest concerns the influence of intra- and extra-linguistic factors on the productivity of either comparative strategy. The data analysis reveals **no sociolinguistic variation in the productivity of periphrastic comparison** in either corpus, even when the corpora are considered jointly as in Figure 1. Although BNC2014-s shows a non-significant tendency for older speakers and those with a college education to use periphrastic comparison less productively, the trend disappears when we restrict the dataset by register (see Section 3.2 above). Inflectional comparatives feature a somewhat different distribution, as social class and gender considerations do appear to have an impact on their productivity across corpora (see Section 4.2.2 below).

Moving on to a consideration of intra-linguistic factors, **periphrastic comparison** in BNC2014-s appears to be significantly *un*productive when accompanied by an **infinitival or prepositional complement** (see Figure 3), and highly productive when no complement/modifier is present. Similar tendencies in terms of infinitival and no complementation are found in BNC1994-s; while there is no evidence for the influence of prepositional complementation/postmodification, this may be due to data restrictions, as we took a sample of 500 speakers from the corpus rather than the whole corpus dataset. In other words, the internal factors influencing the productivity of periphrastic comparison seem to have remained qualitatively similar over time.

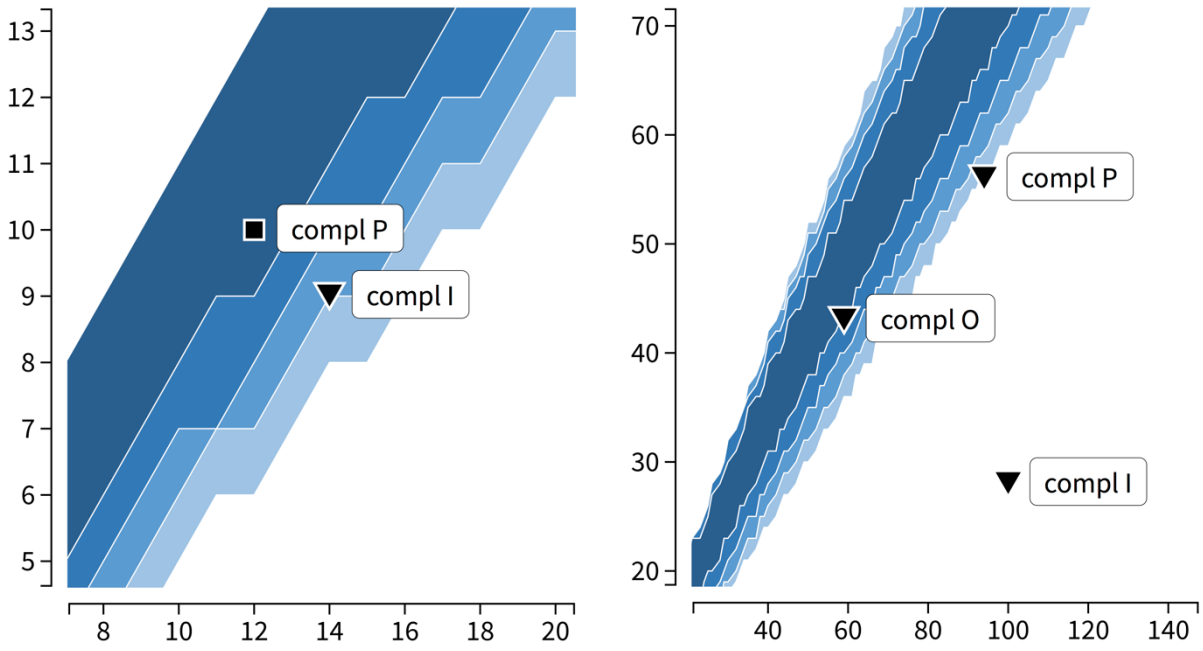


Figure 3. Productivity of periphrastic comparison across different complementation patterns in BNC1994-s (left) and BNC2014-s (right), types (*y* axis) vs. tokens (*x* axis). I = infinitival, P = prepositional, O = other.

Along the same lines, **inflectional comparison** is used unproductively with **infinitival complements/modifiers** in both corpora, as can be seen in Table 1 (the other complementation patterns do not reach significance with the inflectional strategy). In other words, the number of different adjectives that are used with an infinitival complement/postmodifier is very low across the board. In BNC2014-s, for example, only 14 different adjectives are used in inflectional comparison with infinitival complementation: *easier* (the most common type, e.g. *easier to see*), *better*, *cheaper*, *harder*, *quicker*, *nicer*, *longer*, *clearer*, *safer*, *faster*, *higher*, *simpler*, *slower*, *warmer*. The list is similar in BNC1994-s, to the extent that the five most common types are the same, although in a slightly different order. The default option is clearly to have no complement/postmodifier at all, and all other complementation options are quite infrequent.

Internal factor	Strategy	Corpus	Types	Tokens	*p*-value
infinitival complement/modifier	inflectional	BNC1994-s	10	65	< 0.00001
		BNC2014-s	14	259	< 0.00001
	periphrastic	BNC1994-s	9	14	0.008
		BNC2014-s	28	100	< 0.00001

Table 1. Adjective comparison with infinitival complements/modifiers. In each case the number of types is low compared with the expected number of types in a random collection of tokens taken from the same corpus (significance indicated by *p*-values).

In addition, there are two internal factors that seem to influence inflectional but not periphrastic comparison. Inflectional comparison is used highly productively (1) in a **predicative position** and (2) with a **premodifying adverb**. Both seem to be new developments in BNC2014-s (see Table 2); however, there is also some indication in BNC1994-s of the connection between the predicative position and productivity, as the productivity of inflectional comparison is low in the other syntactic positions. When we consider the corpora jointly, we find further factors according to which the productivity of inflectional comparison has changed, such as the lack of a term of comparison; these will be discussed in the next section.

Internal factor	Strategy	Corpus	Types	Tokens	*p*-value
predicative position	inflectional	BNC1994-s	77	898	(0.4)
		BNC2014-s	137	2,685	0.02
premodifying adverb	infectional	BNC1994-s	43	221	(0.6)
		BNC2014-s	96	725	0.0005

Table 2. Internal factors that are associated with high productivity of inflectional comparison in BNC2014-s.

Overall, the data analysis indicates that, while shifts in usage are attested in periphrastic comparatives, it is the inflectional strategy that has undergone significant changes in productivity, both intra- and extra-linguistically, from the 1990s onwards. The changes and the motivations behind them will be further discussed in Section 4.2 below.

4.2. Change in the productivity of inflectional comparison: close-up analysis

4.2.1. Internal factors

Exploratory analysis using *types*2 shows that inflectional comparison is used less productively in BNC1994-s than in BNC2014-s with respect to several structural factors: with a premodifying adverb, in the postpositive and predicative positions, with disyllabic adjectives, and when no term of comparison is present (see Table 3). This implies that the productivity of inflectional comparison has increased over time especially within these categories.

Strategy	Corpus	Internal factor	Type richness	*p*-value
inflectional	BNC1994-s+ 2014-s	(overall)	low in BNC1994-s	0.003
		premodifying adverb	low in BNC1994-s	0.004
		postpositive position	low in BNC1994-s	0.004
		disyllabic adjective base	low in BNC1994-s	0.005
		predicative position	low in BNC1994-s	0.005
		without term of comparison	low in BNC1994-s	0.008

Table 3. Exploring changes in productivity between BNC1994-s and BNC2014-s: top results.

How can we interpret these results? As the vast majority of all instances of inflectional comparison are used in the predicative position and/or without a term of comparison, these two categories naturally behave like the corpus as a whole. Nevertheless, the increase in the productivity of inflectional comparison in **predicative positions** is potentially interesting for, historically, the preference for inflectional forms is less marked in these predicative (as well as postpositive) environments, particularly from the 18th century onwards (see González-Díaz 2008, 82).

More generally, the results seem to suggest a recent ongoing expansion of the functional realm of inflectional forms in Present-day (British) English. This functional expansion is not syntax-specific only: one of the traditional semantic-pragmatic differences between inflectional and periphrastic comparison is that periphrastic comparison allows for a greater emphasis on the actual comparative/degree meaning by having a separate lexical element (*more*) (Curme 1931, 504; Mondorf 2003; Mondorf 2009, 90ff), whereas the inflectional strategy not only lacks this possibility but also places the comparative element at the end of the word, which in a stress-timed language such as English is less than ideal. It is therefore possible that the **co-occurrence of inflectional comparatives with a degree adverb** (see Table 3 above, e.g. *a bit happier, much colder*) may have become a functional measure to compensate for the semantic difference between the two strategies. It should also be noted here that degree adverbs not only convey emphasis but are often used as indicators of social meanings, particularly as in-group markers (see Macaulay 2002 among others). Social factors relevant to this change as well as to the overall increase in the productivity of inflectional comparison will be investigated in the next section.

4.2.2. External factors

After computing the type accumulation curve for BNC1994-s and BNC2014-s combined (Figure 1), we can also plot subcorpora based on social groups onto the curve (Figure 4). Gender-based subcorpora show no significant differences, but when looking at social class, we find that working-class speakers ("C2+DE") use inflectional comparison significantly unproductively in BNC1994-s, whereas in BNC2014-s neither working-class nor middle-class speakers differ significantly from the corpus as a whole. Thus, working-class speakers seem to have caught up with middle-class speakers within the twenty-odd years that have elapsed between the two corpora (provided that we trust the categorisation of social classes in both corpora; cf. Section 3.2 above, Section 5 below).

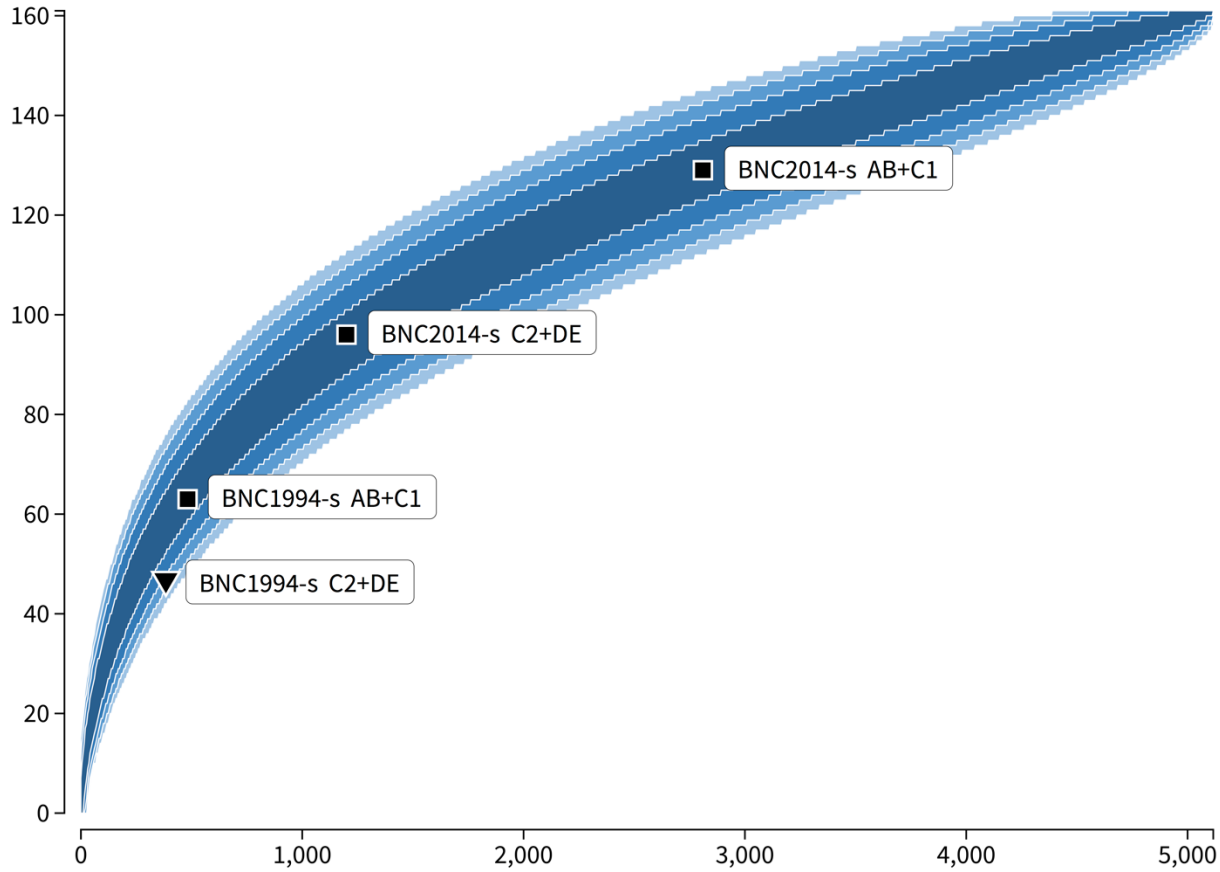
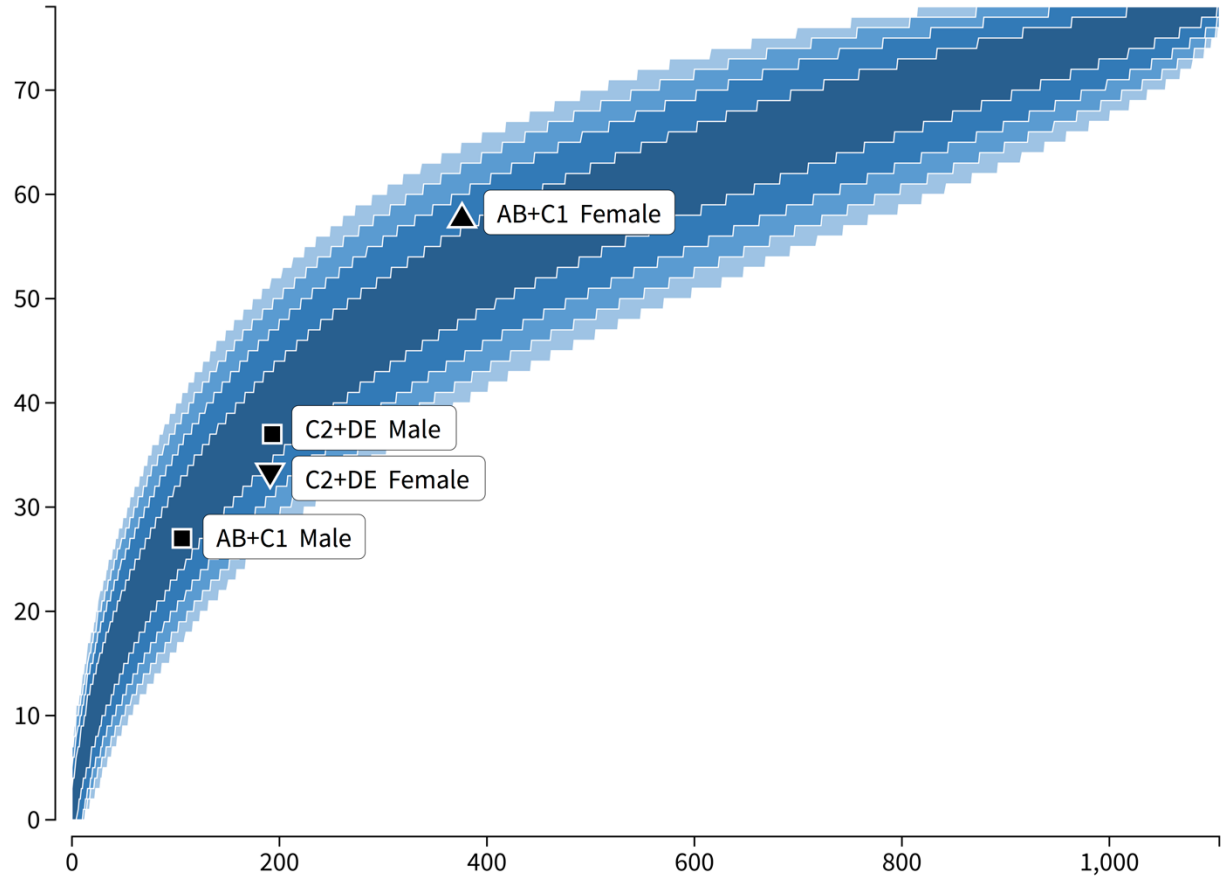


Figure 4. Working-class speakers ("C2+DE") use inflectional comparison unproductively in BNC1994-s (cf. Figure 1). Type frequency (*y* axis) vs. token frequency (*x* axis).

A closer examination of BNC1994-s reveals that it is particularly working-class *women* who use inflectional comparison unproductively (similar results are obtained when we restrict the dataset by register to home settings only). Although without reaching statistical significance, Figure 5 shows that there is approximately the same number of inflectional comparative tokens from both working-class men ("C2+DE Male") and working-class women ("C2+DE Female"), but the women use fewer adjective types, which implies lower productivity and a higher token frequency per type. Consider, in this connection, example (1) below. As previously noted (see fn. 4), *better* is, overall, one of the most frequently used inflectional forms in our data, and it is high-frequency types like this that are repeated more often by the women.

(1) Don't need that on. That's **better** innit?
(BNC1994, KCU 6538–6539, PS0GF: female, working class, age 24)[6]



[6] References to the corpora are given in the following format: corpus, text, speaker: speaker attributes.

Figure 5. In BNC1994-s, especially working-class women ("C2+DE Female") use inflectional comparison unproductively. Type frequency (*y* axis) vs. token frequency (*x* axis).

Säily (2011, 130) finds a similar class-based gender difference in the productivity of the nominal suffix *-ness*. It is as if lexical diversity is more of a concern for working-class men in BNC1994-s, whereas women of the same socio-economic status tend to be more interested in keeping the conversation going by repeating the same few types. Of course, the difference in this case is small; an analysis of the entire demographically sampled spoken section of BNC1994 would be needed to be able to gain more reliable results.

As noted in the previous section, the increased productivity of inflectional comparatives when premodified by a degree adverb could also be a question of style. Here we find a different pattern: in both BNC1994-s and BNC2014-s, male speakers have a (non-significant) tendency to use inflectional comparison productively with a premodifying adverb (see Figure 6), while social class has no effect. The gender difference seems to have grown more pronounced in BNC2014-s, where male speakers use roughly the same number of types as female speakers despite the fact that there is much more data from the latter. While these results become weaker when the dataset is restricted by register, the tendency remains clear, especially in BNC2014-s. Furthermore, there seems to be a combined effect of gender and age in BNC2014-s such that the most unproductive speakers are women representing the youngest age group, 0–29. Thus, in this case the change is not about a certain social group catching up with the rest but rather about an increasing difference between two social groups that also becomes more fine-grained over time. In (2), for instance, a male speaker uses *posher*, a hapax legomenon in the corpus. This is representative of the above-mentioned trend, where males consistently combine the *-er* inflectional strategy with a wider range of adjectival bases than their female counterparts.

(2) and older scouser and he moved away to the Wirral so he's probably a bit a bit **posher**
(BNC2014, BNCJPM003, 0250: male, middle class, age 26)

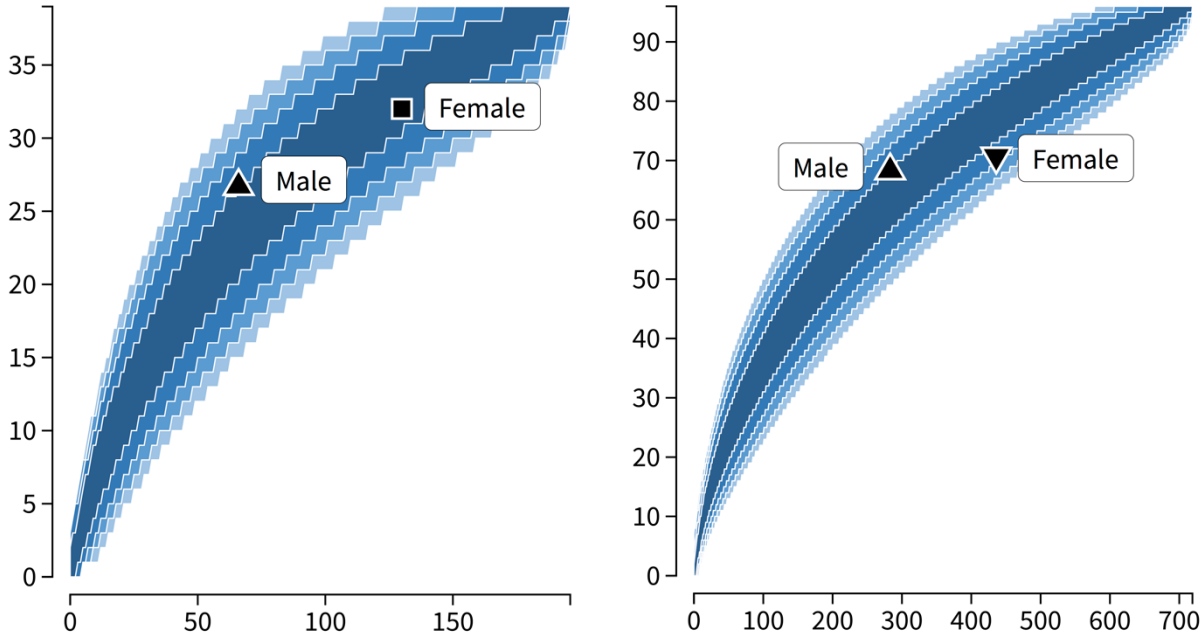


Figure 6. Gender variation in inflectional comparison with a premodifying adverb (BNC1994-s left, BNC2014-s right). Type frequency (*y* axis) vs. token frequency (*x* axis).

5. Discussion

The sections above record a noticeable difference in the behaviour of inflectional and periphrastic adjectival comparatives. Some internal factors (presence/absence of complements) seem to have some impact on the productivity of the periphrastic strategy in PDE; nevertheless, the overall picture is that of stability in its productivity over time and across social categories. By contrast, the productivity of inflectional comparison appears to have significantly increased in the recent history of British English and – more importantly for the overarching theme of the present volume – some of this change is clearly due to social factors.

Our results suggest that the change in the productivity of inflectional comparison seems to have been partly the result of working-class women increasing their usage of the strategy, which now closely matches that of the other social groups. This could perhaps be explained by recent changes in the social position of UK working-class women. In our corpora, the proportion of housewives/women not in the workforce has dropped from 13% to 0.5% between the 1990s and 2010s. This drop in the corpus figures partially matches recent statistics on female employment in Britain, which shows a consistent increase in the figures of working women (from 61.8% to

68.8%) in the period 1993–2015.[7] It would therefore not be unreasonable to assume that their incorporation into the UK workforce has had an indirect impact on working women's linguistic identity, causing a levelling of some aspects of their speech style with respect to other social groups. This interpretation, which focuses on both macro-level social categories and style as a means of identity construction, thus surfs on both the first and third waves of sociolinguistics (Eckert 2012).

It is possible that some of the sociolinguistic variation and change we observed, particularly when it did not reach statistical significance, could be due to chance (cf. Rissanen's (1989, 18) "mystery of vanishing reliability", which refers to dividing the data into ever more detailed categories that end up being too small and thus unrepresentative). However, even when a change is statistically significant, the question remains whether it is really social, or an artefact of the corpus. While BNC1994-s is based on the demographically sampled section of the *British National Corpus* (Burnard 2007, 1.5.1), the new Spoken BNC2014 has not been sampled demographically. Thus, although specific dialectal areas have been targeted, the sampling procedure has been somewhat "opportunistic" (McEnery and Hardie 2012, 11) in that students at the participating universities were asked to record their conversations, and contributions by the general public were invited through an open call via traditional and social media rather than systematically contacting representatives of specific social groups (Love et al. 2017). Hence, more research is needed to verify our results and to assess the representativeness of the entire Spoken BNC2014, of which we had at our disposal the Early Access Subset (BNC2014-s).

Note, also, that gender variation seems to remain alive and well in other environments, as evidenced by the result that male speakers tend to use inflectional comparison with premodifying adverbs more productively than female speakers in both BNC1994-s and BNC2014-s. Previous research has consistently shown that, in the case of other constructions where the conveyance of degree is at stake (e.g. adverbial intensifiers in adjective modifying functions, as in ***very*** *good,* ***really*** *important,* ***pure*** *fast*), it has been women and young people who have tended to lead the usage and renewal of lexical elements (e.g. Macaulay 2006, Tagliamonte 2008, Tagliamonte and Denis 2014). However, our point of view is different in that we are not focusing on lexical change in individual adverbs, but on variation and change in a morpho-syntactic system and, more specifically, in the variety of comparative adjective types used with any premodifying adverb (most often *much*, *a bit*, *a lot*). It may be worth noting here that some studies of derivational productivity (e.g. Keune et al. 2006) have found the most productive speakers to be highly educated older men, so our results align well with theirs.

[7] See https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/timeseries/lf25/lms (last accessed 10/02/2017). The word *partially* needs indeed to be stressed here, as the UK *Office for National Statistics* does not, to the best of our knowledge, provide specific indication of the socio-economic status (working- or middle-class) of the female workforce recorded in their statistics.

Finally, the internal factors we have analysed as potentially influencing productivity have previously been used to study the choice between the inflectional and periphrastic strategies. It is noteworthy that the factors seem to behave differently in the holistic study of productivity on the one hand and in the variationist study of strategy choice on the other. While we have found that the productivity of inflectional comparison is high in the predicative position and in the presence of a premodifying adverb, previous research has found that where both inflectional and periphrastic strategies are possible, it is the periphrastic form that is chosen more frequently (in terms of token frequency) in these contexts (González-Díaz 2008, 110ff). However, strategy choice has not yet been studied in the new BNC2014, so it is possible that preferences may have shifted towards the inflectional form after the early 1990s (cf. the discussion of functional expansion in 4.2.1 above). We therefore leave further investigation of this topic to future research.

6. Conclusion

The previous section has focused on the system-specific findings of the study. In this section, we briefly point to some of its wider implications.

Our findings have crucially shown that social factors – especially gender and social class – play a role in the productivity of inflectional morphology, opening new areas of inquiry to sociolinguistic research. Put differently, processes of variation and change seem to operate in similar ways across inflectional and derivational productivity. These findings contribute to morphological theory-building by providing concrete evidence of the existence of a cline between derivation and inflection, complementing previous research by e.g. Gaeta (2007) and Bauer (2004). How that cline is operationalised, i.e. whether any trends could emerge in terms of e.g. the incidence of factors within and across inflectional / derivational categories, needs to be ascertained by further research.

Furthermore, the study also engages with long-standing diachronic debates about synthetic vs. analytic shifts in English (see Kytö and Romaine 1997; Szmrecsanyi 2012, among others). In line with Schwegler (1994), we subscribe to the idea that analyticity and syntheticity are relative terms that can only be applied to particular systems and/or constructions within languages (as opposed to languages per se). Our research further warns about the relativity of such notions, as it records variable results depending on whether one (token) or another (type) principle is taken as starting point: the increase found in the (type) productivity of inflectional comparison would seem to point towards the increasing syntheticity of the comparative system in PDE. By contrast, the rise in token frequency of periphrastic comparison appears to contribute to increasing analyticity. Most importantly to our project, our findings support the hypothesis of a cline between derivation and inflection, which implies that derivation should be viewed as contributing to syntheticity alongside inflection. The rich array of productive derivational affixes in English could then further challenge the idea of increasing analyticity.

Acknowledgements

The authors wish to thank the anonymous reviewers for their helpful feedback. Thanks are also due to CSC – IT Center for Science, Finland, for computational resources. This work was supported in part by the Academy of Finland grant 276349 to the project 'Reassessing language change: the challenge of real time'.

References

Baayen, R. H. 1992. "Quantitative Aspects of Morphological Productivity." In *Yearbook of Morphology 1991*, edited by Geert Booij and Jaap van Marle, 109–49. Dordrecht: Kluwer.

Baayen, R. H. 1993. "On Frequency, Transparency and Productivity." In *Yearbook of Morphology 1992*, edited by Geert Booij and Jaap van Marle, 181–208. Dordrecht: Kluwer.

Baayen, R. H. 2009. "Corpus Linguistics in Morphology: Morphological Productivity." In *Corpus Linguistics: An International Handbook*, edited by Anke Lüdeling and Merja Kytö, 899–919. Berlin: Mouton de Gruyter.

Bauer, Laurie. 2004. "The Function of Word-Formation and the Inflection-Derivation Distinction." In *Words in Their Places: A Festschrift for J. Lachlan Mackenzie*, edited by Henk Aertsen, Mike Hannay, and Rod Lyall, 283–92. Amsterdam: Vrije Universiteit.

BNC1994 = The *British National Corpus*, version 3 (BNC XML edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. http://www.natcorp.ox.ac.uk/.

Bolinger, Dwight L. 1948. "On Defining the Morpheme." *Word* 4:18–23.

Booij, Geert. 1996. "Inherent versus Contextual Inflection and the Split Morphology Hypothesis." In *Yearbook of Morphology 1995*, edited by Geert Booij and Jaap van Marle, 1–16. Dordrecht: Kluwer.

Bresnan, Joan. 1977. "Variables in the Theory of Transformations." In *Formal Syntax*, edited by Peter W. Culicover, Thomas Wasow, and Adrian Akmajian, 115–22. New York: Academic Press.

Brinton, Laurel J., and Elizabeth Closs Traugott. 2005. *Lexicalization and Language Change*. Cambridge: Cambridge University Press.

Burnard, Lou, ed. 2007. "Reference Guide for the British National Corpus (XML Edition)." Published for the British National Corpus Consortium by the Research Technologies Service at Oxford University Computing Services. http://www.natcorp.ox.ac.uk/docs/URG/.

Cresswell, Max J. 1976. "The Semantics of Degree." In *Montague Grammar*, edited by Barbara H. Partee, 261–92. London & New York: Academic Press.

Curme, George O. 1931. *A Grammar of the English Language. Volume II: Syntax*. D. C. Heath & Company.

Danchev, Andrei. 1992. “The Evidence for Analytic and Synthetic Developments in English.” In *History of Englishes: New Methods and Interpretations in Historical Linguistics*, edited by Matti Rissanen, Ossi Ihalainen, Terttu Nevalainen, and Irma Taavitsainen, 25–41. Berlin: Mouton de Gruyter.

Eckert, Penelope. 2012. “Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation.” *Annual Review of Anthropology* 41:87–100.

Gaeta, Livio. 2007. “On the Double Nature of Productivity in Inflectional Morphology.” *Morphology* 17(2):181–205.

Gardner, Anne. 2013. “Derivation in Middle English: Regional and Text Type Variation.” PhD dissertation, University of Zurich.

Gathercole, Virginia. 1979. *Birdies Like Birdseed the Better than Buns: A Study of Relational Comparatives and their Acquisition*. PhD dissertation, University of Kansas.

Gathercole, Virginia. 1985. “More and More about More.” *Journal of Experimental Child Psychology* 40(1):73–104.

Gitterman, D., and Judith R. Johnston. 1983. “Talking about Comparisons: A Study of Young Children’s Comparative Adjective Usage.” *Journal of Child Language* 10:605–21.

González-Díaz, Victorina. 2004. “Adjectival Double Periphrastic Comparison in EModE: A Socio-stylistic Analysis.” *Folia Linguistica Historica* 25:177–210.

González-Díaz, Victorina. 2006. “The Origin of English Periphrastic Comparison.” *English Studies* 84:707–34.

González-Díaz, Victorina. 2008. *English Adjective Comparison: A Historical Perspective*. Amsterdam: John Benjamins.

Graziano-King, Janine. 1999. *Acquisition of Comparative Forms in English*. PhD dissertation, The Graduate School of the City University of New York.

Graziano-King, Janine, and Helen Smith Cairns. 2005. “Acquisition of English Comparative Adjectives.” *Journal of Child Language* 32:345–73.

Hardie, Andrew. 2012. “CQPweb – Combining Power, Flexibility and Usability in a Corpus Analysis Tool.” *International Journal of Corpus Linguistics* 17(3):380–409.

Haspelmath, Martin. 2002. *Understanding Morphology*. London: Arnold.

Hay, Jennifer, and R. H. Baayen. 2002. “Parsing and Productivity.” In *Yearbook of Morphology 2001*, edited by Geert Booij and Jaap van Marle, 203–235. Dordrecht: Kluwer.

Hilpert, Martin. 2008. “The English Comparative – Language Structure and Language Use.” *English Language and Linguistics* 12(3):395–417.

Hilpert, Martin. 2010. “Comparing the Comparative: A Corpus-based Study of Comparative Constructions in English and Swedish.” In *Contrastive Studies in Construction Grammar*, edited by Hans Boas, 22–41. Amsterdam: John Benjamins.

Hilpert, Martin. 2013. *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge University Press.

Hohaus, Vera, Sonja Tiemann, and Sigrid Beck. 2014. “Acquisition of Comparative Constructions.” *Language Acquisition* 21(3):215–49.

Huddleston, Rodney. 1967. “More on the English Comparative.” *Journal of Linguistics* 3:91–102.

Huddleston, Rodney, and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Jespersen, Otto. 1949. *A Modern English Grammar on Historical Principles. Part VII: Syntax*. Copenhagen: Ejnar Munksgaard.

Keune, Karen, Roeland van Hout, and R. H. Baayen. 2006. “Socio-Geographic Variation in Morphological Productivity in Spoken Dutch: A Comparison of Statistical Techniques.” In *Actes de JADT 2006 : 8es Journées Internationales D’analyse Statistique Des Données Textuelles*, edited by Jean-Marie Viprey and Lexicometrica, 571–81. Besançon: Université de Franche-Comté.

Knüpfer, Hans. 1921. “Die Anfange der periphrastischen Komparation im Englischen.” *Englische Studien* 55:321–89.

Kytö, Merja. 1996. “’The Best and Most Excellentest Way’: The Rivalling Forms of Adjective Comparison in Late Middle and Early Modern English.” In *Words. Proceedings of an International Symposium, Lund, 25-26 August 1995*, edited by Jan Svartvik, 124–44. Stockholm: Kungl. Vitterhets Historie och Antikvitets Akademien.

Kytö, Merja, and Suzanne Romaine. 1997. “Competing Forms of Adjective Comparison in Modern English: What Could be More Quicker and Easier and More Effective?” In *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*, edited by Terttu Nevalainen and Leena Kahlas-Tarkka, 329–52. Helsinki: Société Néophilologique.

Kytö, Merja, and Suzanne Romaine. 2006. “Adjective Comparison in Nineteenth-century English.” In *Nineteenth-century English: Stability and Change*, edited by Merja Kytö, Mats Rydén and Erik Smitterberg, 194–214. Cambridge: Cambridge University Press.

Leech, Geoffrey, and Jonathan Culpeper. 1997. “The Comparison of Adjectives in Recent British English.” In *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*, edited by Terttu Nevalainen and Leena Kahlas-Tarkka, 353–73. Helsinki: Société Néophilologique.

Lindquist, Hans. 2000. “Livelier or More Lively? Syntactic and Contextual Factors Influencing the Comparison of Disyllabic Adjectives.” In *Corpora Galore: Analyses and Techniques in describing English*, edited by John M. Kirk, 125–32. Amsterdam: Rodopi.

Love, Robbie, Claire Dembry, Andrew Hardie, Vaclav Brezina, and Tony McEnery. 2017. “The Spoken BNC2014: Designing and Building a Spoken Corpus of Everyday Conversations.” *International Journal of Corpus Linguistics* 22(3).

Macaulay, Ronald. 2002. “Extremely Interesting, Very Interesting, or Only Quite Interesting? Adverbs and Social Class.” *Journal of Sociolinguistics* 6(3):398–417.

Macaulay, Ronald. 2006. “Pure Grammaticalization: The Development of a Teenage Intensifier.” *Language Variation and Change* 18(3):267–83.

Matsui, Chie. 2010. "*-er* type or *more* type adjectives of comparison in English." *Studia Neophilologica* 82(2):188–202.
McEnery, Tony, and Andrew Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
Mondorf, Britta. 2003. "Support for *more*-support." In *Determinants of Grammatical Variation in English*, edited by Günter Rohdenburg and Britta Mondorf, 251–304. Berlin & New York: Mouton de Gruyter.
Mondorf, Britta. 2007. "Recalcitrant Problems of Comparative Alternation and New Insights Emerging from Internet Data." In *Corpus Linguistics and the Web*, edited by Marianne Hundt, Nadja Nesselhauf, and Carolin Biewer, 211–32. Amsterdam: Rodopi.
Mondorf, Britta. 2009. *More Support for* more-*support. The Role of Processing Constraints on the Choice between Synthetic and Analytic Comparative Forms*. Amsterdam: John Benjamins.
Mossé, Fernand. 1952. *A Handbook of Middle English*. Baltimore & London: The John Hopkins University Press.
National Readership Survey. 2015. "Social Grade." http://www.nrs.co.uk/nrs-print/lifestyle-and-classification-data/social-grade/.
Palmer, Chris C. 2009. "Borrowings, Derivational Morphology, and Perceived Productivity in English, 1300–1600." PhD dissertation, University of Michigan.
Perek, Florent. 2016. "Using Distributional Semantics to Study Syntactic Productivity in Diachrony: A Case Study." *Linguistics* 54(1):149–88.
Plag, Ingo. 1999. *Morphological Productivity: Structural Constraints in English Derivation*. Berlin: Mouton de Gruyter.
Plag, Ingo. 2003. *Word-Formation in English*. Cambridge: Cambridge University Press.
Plag, Ingo. 2006. "Productivity." In *The Handbook of English Linguistics*, edited by Bas Aarts and April McMahon, 537–56. Oxford: Blackwell.
Poutsma, Hendrik. 1914. *A Grammar of Late Modern English*. Groningen: Noordhoff.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
Rissanen, Matti. 1989. "Three Problems Connected with the Use of Diachronic Corpora." *ICAME Journal* 13:16–9.
Romaine, Suzanne. 1983. "On the Productivity of Word Formation Rules and Limits of Variability in the Lexicon." *Australian Journal of Linguistics* 3(2):177–200.
Rusiecki, Jan. 1985. *Adjectives and Comparison in English. A Semantic Study*. London & New York: Longman.
Säily, Tanja. 2011. "Variation in Morphological Productivity in the BNC: Sociolinguistic and Methodological Considerations." *Corpus Linguistics and Linguistic Theory* 7(1):119–41.
Säily, Tanja. 2014. *Sociolinguistic Variation in English Derivational Productivity: Studies and Methods in Diachronic Corpus Linguistics*. Helsinki: Société Néophilologique.

Säily, Tanja, and Jukka Suomela. 2009. “Comparing Type Counts: The Case of Women, Men and *-ity* in Early English Letters.” In *Corpus Linguistics: Refinements and Reassessments*, edited by Antoinette Renouf and Andrew Kehoe, 87–109. Amsterdam: Rodopi.

Säily, Tanja, and Jukka Suomela. Forthcoming. “*types*2: Exploring Word-Frequency Differences in Corpora.” In *Big and Rich Data in English Corpus Linguistics: Methods and Explorations*, edited by Turo Hiltunen, Joe McVeigh, and Tanja Säily. Helsinki: VARIENG. https://users.ics.aalto.fi/suomela/types2-d2e/.

Schröder, Anne. 2008. “On the Productivity of Verbal Prefixation in English.” Habilitationsschrift, Martin-Luther-Universität Halle-Wittenberg.

Schwegler, Armin. 1994. “Analysis and Synthesis.” In *The Encyclopedia of Language and Linguistics*, edited by R. E. Asher and J. M. Y. Simpson, vol. I, 111–14. Oxford: Pergamon Press.

Štekauer, Pavol, Don Chapman, Slávka Tomaščíková, and Štefan Franko. 2005. “Word-Formation as Creativity within Productivity Constraints: Sociolinguistic Evidence.” *Onomasiology Online* 6:1–55.

Strang, Barbara. 1970. *A History of English*. London: Methuen.

Stump, Gregory T. 1998. “Inflection.” In *The Handbook of Morphology*, edited by Andrew Spencer and Arnold M. Zwicky, 13–43. Oxford: Blackwell.

Suomela, Jukka. 2016. “*types*2: Type and Hapax Accumulation Curves.” Computer program. http://users.ics.aalto.fi/suomela/types2/.

Szmrecsanyi, Benedikt. 2005. “Language Users as Creatures of Habit: A Corpus-based Analysis of Persistence in Spoken English.” *Corpus Linguistics and Linguistic Theory* 1(1):113–50.

Szmrecsanyi, Benedikt. 2012. “Analyticity and Syntheticity in the History of English.” In *The Oxford Handbook of the History of English*, edited by Terttu Nevalainen and Elizabeth Closs Traugott, 654–65. Oxford: Oxford University Press.

Tagliamonte, Sali A. 2008. “So Different and Pretty Cool! Recycling Intensifiers in Toronto, Canada.” *English Language and Linguistics* 12:361–94.

Tagliamonte, Sali A., and Derek Denis. 2014. “Expanding the Transmission/Diffusion Dichotomy: Evidence from Canada.” *Language* 90(1):90–136.

van Rooij, Robert. 2010. “Explicit versus Implicit Comparatives.” In *Vagueness and Language Use*, edited by Paul Égré and Nathan Klinedinst, 51–72. Houndsmills: Palgrave Macmillan.

van Rooij, Robert. 2011. “Measurement and Interadjective Comparisons.” *Journal of Semantics* 28:335–58.

Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. “Empirical Foundations for a Theory of Language Change.” In *Directions for Historical Linguistics: A Symposium*, edited by Winfred P. Lehmann and Yakov Malkiel, 95–188. Austin: University of Texas Press.

Zeldes, Amir. 2009. “Quantifying Constructional Productivity with Unseen Slot Members.” In *CALC ’09: Proceedings of the NAACL HLT Workshop on Computational Approaches to*

Linguistic Creativity, edited by Anna Feldman and Birte Lönneker-Rodman, 47–54. Stroudsburg, Pennsylvania: Association for Computational Linguistics.

Zeldes, Amir. 2013. “Productive Argument Selection: Is Lexical Semantics Enough?” *Corpus Linguistics and Linguistic Theory* 9(2):263–91.