

## Recent change in stative progressives: a collostructional investigation of British English in 1994 and 2014<sup>1</sup>

PAULA RAUTIONAHO  
*University of Eastern Finland*

and

ROBERT FUCHS  
*Universität Hamburg*

(Received 6 April 2019; revised 18 October 2019)

The spread of the progressive from dynamic to stative verbs started in the seventeenth century, and slowed down in the late twentieth century. The present study investigates recent change in the use of stative progressives in conversational British English from the early 1990s to the early 2010s. The analysis focuses on a total of 100 stative verb lemmata in the spoken, demographic sections of the original and new *British National Corpus*, restricted to a variable context where a progressive could potentially occur. Results indicate that overall, stative progressives have not become more frequent in the last twenty years, and that the group of stative verbs is highly heterogeneous. However, particular verbs, such as *EXPECT* and *THINK*, do indeed combine more frequently with the progressive *now*, which could be the cause of the popular impression of the continuing spread of stative progressives. In addition to a frequency-based analysis, a distinctive collexeme analysis offers a more fine-grained analysis of the collostructional preferences of individual verb lemmata and semantic classes of stative verbs. This analysis reveals that the stative verbs are heterogeneous and that the lemmata most distinctly associated with the progressive belong to the group of stance verbs.

**Keywords:** progressive, stative, British English, collostructional analysis, recent diachronic change

### 1 Introduction

We currently lack information on very recent developments in the increasing frequency of the progressive (*BE + V-ing*). The spread of the progressive has been convincingly documented to have started in the seventeenth century, escalated in the nineteenth century, and reached a plateau in the latter part of the twentieth century (see e.g. Kranich 2010: 95, and section 3 below). However, the factors propelling this change are controversial. Mair (2006: 88–9) offers the following three explanations:

<sup>1</sup> We would like to thank the two anonymous reviewers and the journal editor Laurel Brinton for their constructive feedback. All remaining errors are our own.

- (i) the increasing frequency of many already established uses of the progressive,
- (ii) the creation of new forms to fill in gaps in the progressive paradigm, and
- (iii) the greater tendency to use the progressive with stative verbs.

Indeed, according to Smith (2005), the progressive is spreading especially through the use of present tense forms, i.e. case (i), an already established form gaining even higher frequency. On the other hand, Smitterberg (2005: 123) points out that the emergence of the passive progressive (e.g. *The house is being built*) in the nineteenth century filled a gap in the progressive paradigm, i.e. case (ii). With regard to the spread from the prototypical domain of the progressive, dynamic verbs, to new domains, such as stative verbs (i.e. case (iii)), there is convincing real-time evidence of the increasing use of stative progressives beginning in the seventeenth century and spreading rapidly in the nineteenth century, thus coinciding with the overall increase of the progressive construction (see e.g. Kranich 2010: 95, 193). However, the spread of stative progressives may, in fact, have halted in the twentieth century, with just a small increase in frequency in written British and American English in the 1960s to 1990s (Smith & Leech 2013: 89), which did not contribute to the increasing use of the progressive in general in this period (Leech, Hundt, Mair & Smith 2009: 130).

Beyond the 1990s, we currently lack evidence on the further development of stative progressives. It remains unclear whether the spread of stative progressives has picked up again, or indeed if the trend has been reversed. The present study will address this question with the help of the recently published Spoken BNC2014 corpus and will investigate the frequency of the progressive with 100 stative verbs. We focus on developments in the use of individual lemmata, as well as groups of stative verbs, based on the analysis of relative frequency within a variable context and with the help of distinctive collexeme analysis (see Gries & Stefanowitsch 2004). This method enables us to assess degrees of association between individual lexical verbs and the two constructions, the progressive and the non-progressive.

Sections 2 and 3 survey previous research on the use of stative verbs in the progressive, providing the rationale for the approach chosen in the present study. The data and their statistical treatment are introduced in section 4, and the results of the analyses are presented in section 5. Finally, section 6 discusses the findings of the study.

## 2 The progressive aspect and the stative verb: poles apart?

Conceptually, stative verbs are incompatible with the progressive aspect because of a semantic conflict: states lack, by definition, the element of progression that is crucial to the progressive (Quirk, Greenbaum, Leech & Svartvik 1985: 198; see also Declerck 1991: 167). On one hand, the progressive aspect is used to portray an action or event as ongoing at event time and as temporally bounded (Biber, Johansson, Leech, Conrad & Finegan 1999: 470–1). Stative verbs, on the other hand, are characterized by the absence of change and by temporal unboundedness (Brinton 1988: 24). We can thus clearly see the inherent semantic mismatch: the progressive implies ongoing

change and temporal boundedness, while stative verbs are temporally unbounded and lack change. It should be pointed out that, in fact, it is not verbs that are stative, but rather the situations that the verbs denote (see e.g. Quirk *et al.* 1985: 200). Thus, while it is a commonly accepted approximation to address the categorization of verbs into different semantic domains, one should keep in mind that the stativity arises from the predication made, and the situation denoted, rather than from stative verbs per se. In the present study, we acknowledge this issue by considering stativeness as a gradable property and by distinguishing 'stative' and 'dynamic' readings of 'stative' verbs (see below).

Despite the semantic mismatch, stative verbs in English do occur in the progressive. As Leech *et al.* (2009: 129; see also Declerck 1991: 167–73; Biber *et al.* 1999: 471) point out, verbs that are normally stative can occur in the progressive when they portray temporary states (as in (1)); states changing by degrees (as in (2)); and agentive uses of the verb BE (as in (3)). Biber *et al.* (1999: 472) provide information on the frequency of different lexical verbs occurring with the progressive; stative verbs vary to a great extent, with some listed as occurring rarely with the progressive (e.g. BELIEVE, KNOW) and others frequently (e.g. HOPE, THINK, stance verbs).

- (1) Mary's **living** in a flat in London.<sup>2</sup>
- (2) The baby's **resembling** his father more and more every day.
- (3) John's **being** silly.

In each sentence, the situation is no longer purely static, but the progressive renders the situation dynamic. Smith (1991: 251) discusses an example similar to (3) in more detail and shows how the subject's active engagement in the situation allows for a dynamic reading, and how, simultaneously, the use of the progressive portrays the situation as temporary; the present simple *John's silly* would suggest a permanent trait. Similarly, Granath & Wherrity (2014), discussing the use of the verbs LOVE and KNOW in particular, propose that the use of stative verbs in the progressive construction simply results in an aspectual change, with 'heightened agentivity and greater vividness' (2014: 12), as in the much-discussed McDonald's slogan (in (4)).

- (4) I'm **loving** it.

While progressive use of the verb LOVE has increased slightly in frequency in recent British English (BrE) (see e.g. Freund 2016; Martínez-Vásquez 2018), this might not be the case for other stative verbs, and many researchers prefer to consider stativity as a cline rather than a dichotomy (see e.g. Paulasto 2014: 252). To make better sense of the behavior of different stative verbs with the progressive, it is useful to distinguish different groups, ordered 'in a decreasing order of resistance to the progressive' (Freund 2016: 53, following Aarts, Close & Wallis 2010: 163), as shown in table 1 (based on

<sup>2</sup> Examples (1) to (3) are from Leech *et al.* (2009: 129).

Table 1. *Semantic categories of stative verbs, from most to least 'resistant' to the progressive*

| Category   | Example verbs                 |
|------------|-------------------------------|
| Relational | BE, CONTAIN, HAVE, RESEMBLE   |
| Cognitive  | KNOW, THINK, UNDERSTAND       |
| Affective  | DISAGREE, LOVE, PREFER, WANT  |
| Perception | SEE, HEAR, SMELL, HURT, TASTE |
| Stance     | LIE, LIVE, SIT, STAND         |

Greenbaum & Quirk 1990 and Huddleston & Pullum 2002). It should be noted, however, that different verbs allocated to the same category may occur in different frequencies with the progressive (see e.g. Biber *et al.* 1999: 471–2).

Relational verbs (as in (5)) refer to permanent situations, which do not normally have a dynamic reading, and are thus incompatible with the progressive (Declerck 1991: 169).<sup>3</sup> Similarly, cognitive verbs (as in (6)) usually refer to permanent qualities, and thus are not normally used with the progressive. Affective verbs (as in (7)) may occur in the progressive when they express temporariness or tentativeness (Quirk *et al.* 1985: 203), or for 'greater vividness', as discussed above. Verbs of perception form a more complex group in which some verbs are compatible with the progressive while others are not. Verbs referring to involuntary reaction of the senses (e.g. SEE, HEAR; as in (8a)) are not usually used in the progressive (Declerck 1991: 167), while with verbs of bodily sensation (e.g. HURT, ITCH, as in (8b)), the non-progressive and the progressive are practically interchangeable (Declerck 1991: 169).<sup>4</sup> Finally, stance verbs (as in (9)) are 'intermediate between the stative and dynamic categories' (Quirk *et al.* 1985: 205), and can be used to describe the state of affairs even if there is no element of dynamism involved (Kranich 2010: 50).

- (5) He **has** a nice wife. (\*is having)<sup>5</sup>  
 (6) I **know** Cantonese. (\*am knowing)  
 (7) I **love** you. (?am loving)  
 (8a) I **noticed** her leave the shop. (\*was noticing)  
 (8b) He said his stomach **hurt**. (was hurting)  
 (9) John **sat** in the chair. (was sitting)

The interested reader will find a more detailed characterization of stative progressives in numerous previous studies approaching the topic from different perspectives; theoretical

<sup>3</sup> However, the agentive use of BE, for instance, warrants progressive usage when the attribute is under the control of the subject, cf. *John's being silly* vs. \**John's being tall* (see Smith 1991: 251).

<sup>4</sup> The two groups of verbs are sometimes dealt with separately; e.g. Quirk *et al.* (1985: 203) discuss 'states of perception' and 'states of bodily sensation'.

<sup>5</sup> Examples (5) and (7)–(9) come from Declerck (1991: 167–9).

(e.g. Smith 1991), diachronic (e.g. Levin 2013), regional variation (e.g. Rautionaho 2014), or second-language acquisition (e.g. Fuchs & Werner 2018). In the following section, we focus on the spread of stative progressives, on the basis of previous studies.

### 3 The spread of stative progressives before the 1990s

The spread from the prototypical domain, dynamic verbs, to stative verbs is commonly regarded as one of the catalysts of the spread of the progressive construction overall, despite the mismatch of the semantic properties of the two (e.g. Mair 2006: 92–4; Leech *et al.* 2009: 129–30). The rise of the progressive (i.e. not only stative uses) started slowly during the seventeenth and eighteenth centuries (see figure 1), with a rapid increase in the nineteenth century, slowing down in the twentieth century. This development follows the typical S-shaped curve, and might not yet have reached its final stages, as the last period for which data are available still shows a relative rise instead of leveling out into a plateau (Kranich 2010: 107fn.).

The rapid increase of stative progressives (see figure 2) actually starts slightly earlier than that of the progressive overall, thus lending support to the role of stative progressives as a catalyst for the overall increase of the progressive construction. The slight slowdown in the rise of stative progressives seen around the turn of the twentieth century in figure 2 is likewise reflected in a slightly later slowdown in the rise of progressives in general, during the mid twentieth century (figure 1).

Overall, however, the frequency of stative progressives remains rather low throughout history; Kranich's (2010) data show that during the twentieth century their normalized frequency (per 100,000 words) rises from *c.* 32 to 50. However, Leech *et al.* (2009), focusing on developments in the use of stative progressives between the 1960s and the 1990s, find only a slight increase in the frequency and number of verb types in the given period (2009: 130).<sup>6</sup> Based on their own results, Leech *et al.* do not see a clear connection between the use of stative progressives and the overall increase of the progressive construction, although they do admit that their dataset is perhaps too small 'to enable firm conclusions to be drawn about diachronic change' (2009: 130). Smith & Leech (2013: 89) confirm that the increase is not statistically significant.

Freund (2016) is the study most similar to the present one. Freund compares the frequency of 25 stative verbs in BNC1994DS and a corpus including speech-like computer-mediated communication data collected in the years 2012–14; she finds no increase in the frequency of stative progressives overall, or in the proportion of the different semantic categories (see section 2), but shows that some verbs (such as FEEL, LOVE) do increase. An acceptability survey and a survey of BrE student grammars' handling of the stative progressive indicates that 'there is increasing tolerance of stative-progressive combinations', and Freund goes on to predict that the frequency of stative progressives is likely to 'expand over time' (2016: 59). However, the small size

<sup>6</sup> Leech *et al.* (2009: 292) provide raw frequencies of stative progressives: RF=83 for the *Lancaster-Oslo/Bergen* corpus (LOB; 1961) and RF=102 for the *Freiburg-Lancaster-Oslo/Bergen* corpus (F-LOB; 1991).

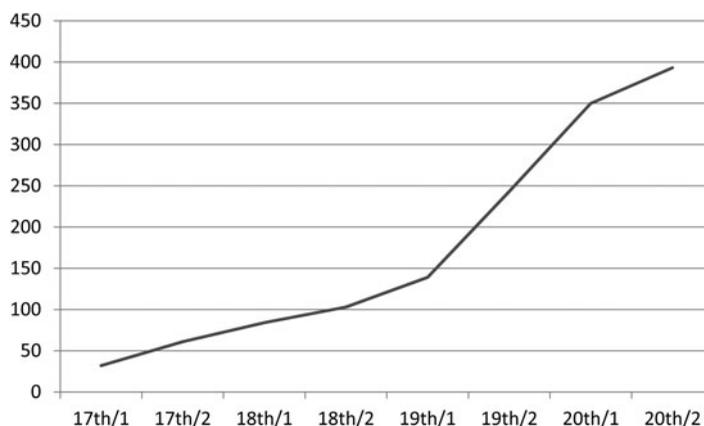


Figure 1. Increase in the frequency of the progressive construction (normalized per 100,000 words; based on Kranich 2010: 95 table 3a)

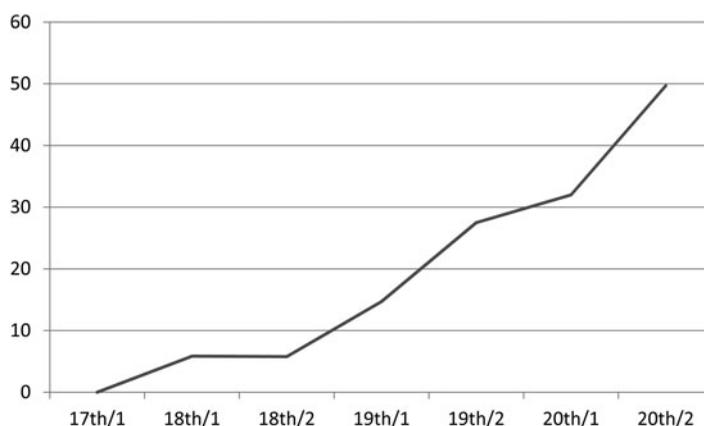


Figure 2. Increase in the frequency of stative progressives (normalized per 100,000 words; based on Kranich 2010: 153 table 16)

of Freund's dataset again allows only limited conclusions – the present study, by contrast, contributes a larger dataset that will reveal more robust results.

With regard to developments in the semantic categories of stative verbs, Leech *et al.* (2009: 130) find a small increase in the number of progressives used with verbs of perception and relational verbs. They, however, go on to point out that the dataset is once again too small to enable firm conclusions to be drawn about diachronic change regarding different types of stative verbs. Moreover, Leech *et al.*'s findings are not always in line with Freund's (2016) assertion that relational or cognitive verbs would be more resistant to the progressive than stance verbs, for instance; thus, following Freund (2016), we would expect to see increasing use of the progressive with stance or affective verbs. In a similar vein, Freund's (2016) recent data show only two stative verbs

increasing their progressive usage in statistically significant numbers: BE (relational) and THINK (cognitive). Finally, on a methodological level, measuring the frequency of stative progressives by comparison to the frequency of all progressives may yield counter-intuitive results if the frequency of the progressive with non-stative verbs varies over time. We address this issue by adopting an alternative measure of progressive frequency.

It is against this background that we set out to investigate recent change in stative progressives. Previous studies have laid out an increasing trend in the use of the progressive with stative verbs in the past centuries, with a leveling-out of the increase in the latter part of the twentieth century, but we currently lack convincing evidence on the further development of stative progressives since the 1990s as well as their usage in conversations, where they are likely to be more frequent than in formal written registers. More specifically, our study seeks to answer the following research questions:

- (i) Are stative progressives more frequent in the present day, compared to the early 1990s?
- (ii) Is the progressive applied to a broader range of stative verbs in the present day, compared to the early 1990s?
- (iii) Is the progressive applied to a broader range of semantic domains of stative verbs in the present day, compared to the early 1990s?
- (iv) Has there been any change in the collostructional preferences of stative verbs, and stative categories between the early 1990s and the present day?

## 4 Data and methods

### 4.1 BNC1994 and BNC2014

The present study relies on data from the *British National Corpus*; we compare data from the original 1994 version to data from the new 2014 version. The spoken component of the BNC1994 includes data from the late 1980s and early 1990s and has two sections: the demographically sampled part and the context-governed part (Love, Dembry, Hardie, Brezina & McEnery 2017). We use only the demographically sampled part (BNC1994DS), which has approximately 4.2 million words from over 1,000 speakers, who were ‘selected by age group, sex, social class and geographic region’ (Aston & Burnard 1998: 31). The Spoken BNC2014, with material recorded between the years 2012 and 2016 (Love *et al.* 2017), provides a point of comparison from the present day. Here we use the sample released in 2016 (BNC2014S), which contains approximately 4.8 million words (Love *et al.* 2017: 9–10). For future sociolinguistic analyses of the data, the queries were restricted to tokens with full metadata available, i.e. if any of the sociolinguistic variables of age, sex and social class were unknown, the occurrence was excluded. The number of words included in the two datasets after the unknown metadata was excluded is shown in table 2.

Table 2. *Number of words in the datasets*

|   | BNC1994DS | BNC2014S  |
|---|-----------|-----------|
| No. of words                            | 4,233,962 | 4,788,887 |
| No. of words, unknown metadata excluded | 2,617,294 | 3,538,652 |

#### 4.2 *The variable context: progressive versus non-progressive*

We searched the two corpora, BNC1994DS and Spoken BNC2014S,<sup>7</sup> for the lemmata of 100 stative verbs (e.g. ‘{ache/V}’; see appendix 1 for the full list<sup>8</sup>). The verbs were collected from earlier studies on stative progressives (Leech *et al.* 2009; Freund 2016), and from grammars of English (Quirk *et al.* 1985; Biber *et al.* 1999; Huddleston & Pullum 2002). The extracted tokens were then subjected to a manual editing process in order to restrict the analysis to a variable context, i.e. to only those instances where the progressive could potentially occur. In most cases (89 out of 100 lemmata), we included all of the tokens, whereas in the case of very frequent verbs, occurring more than 5,000 times in the two corpora combined, we included a random sample of 1,000 tokens per corpus (these verbs are marked with an asterisk in appendix 1). In determining where a progressive form can potentially occur, we followed the criteria laid out by Fuchs & Werner (2018: 206): we excluded any catenative verb constructions (10), any infinitival (11) or imperative (12) forms, and where the token was not a verb at all (13). We decided to exclude instances that involved a modal auxiliary (as in (14)),<sup>9</sup> in order to keep the amount of data manageable. Further, we excluded a number of idiomatic expressions where the alternation is not possible (as in (15)),<sup>10</sup> and any repeated (16) or unclear (17) occurrences of the target verb.

- (10) I keep **forgetting** her name. (BNC1994DS, KE4)  
 (11) and it was going **to cost** him more than ... (BNC2014S, STB7)  
 (12) Mhm. Right, **stand up**. (BNC1994DS, KDE)  
 (13) because seeing is **believing** (BNC2014S, SGSY)  
 (14) **May I see** him? (BNC1994DS, KCK)  
 (15) Oh **it doesn’t matter**. (BNC1994DS, KBG)

<sup>7</sup> Data cited herein have been extracted from the *British National Corpus* online service, managed by Oxford University Computing Services on behalf of the BNC Consortium. All rights in the texts cited are reserved. We used BNCWeb (bncweb.lancs.ac.uk) for extracting the data from BNC1994DS, and CQPweb (Hardie 2012) for extracting the data from BNC2014S.

<sup>8</sup> The appendices are available online at <https://osf.io/eqamv/>

<sup>9</sup> The progressive does occur with some, but not all, modal auxiliaries (see e.g. Rautionaho 2014: 106–20).

<sup>10</sup> Such idiomatic expressions include *you know*; *it doesn’t matter*; *mind you*; *I/you see* (intransitive); *was sat/stood*; *BE supposed to*. Similarly, *I mean* could have been excluded from the data, but since the progressive is actually possible (1 instance found in BNC1994, 3 in BNC2014), the alternation does exist, and *I mean* was included in the data.

(16) S0254: did you say she **was aching** the other day

S0253: she **was aching** but that's only a good thing isn't it (BNC2014S, S4YQ)<sup>11</sup>

(17) it's like I'm **appreciate** it a bit more (BNC2014S, SEDS)

Next, all dynamic uses of the stative verbs were identified and eventually excluded from the data. These instances include, for instance, *SEE* in the sense of 'to date/meet with someone', *HANG* in the sense of 'to fasten', and *HAVE* in senses other than 'possess'. The dynamic uses were excluded in order to, first and foremost, focus on actual stative uses with the progressive; for instance, *HAVE* is frequently used as part of a more or less fixed expression in combinations such as *having problems*, *having a bad day*, or *having a party*, in which the meaning of *HAVE* is far removed from the stative meaning, and rather refers to a dynamic situation (see Kranich 2010: 155).

Finally, all remaining instances were annotated as either 'progressive' or 'non-progressive', depending on the form of the verb construction. This manual editing of the extracted tokens left us with altogether 44,016 occurrences of the 100 lexical verbs, either in the progressive or the simple. These instances were further analyzed with regard to the semantic category, following Greenbaum & Quirk (1990), among others (see section 2 above). It should be noted that, in order to account for potential polysemy, each token was analyzed separately, rather than assigning all instances of a lemma into one, most prominent, semantic category. For instance, sentence (18) was analyzed as containing a relational verb, while (19) contains an affective verb.

(18) See how much it's **costing**? Two hundred and fifty pounds a night! (BNC1994DS, KBC)

(19) I **hate** it when people do that (BNC2014S, SLDD)

Once all the tokens were analyzed, we normalized the number of progressives in two ways: first, since the overall number of words in our two datasets are different, we include normalization per million words to allow the comparison of the number of stative progressives in each dataset. Perhaps more interestingly, we calculate the proportion of progressive instances of all occurrences of the lemma in order to investigate the variable context, i.e. actual versus potential occurrences. The proportion of progressives was calculated as follows: (number of progressives/number of all occurrences of lemma)\*100. This type of normalization takes into account the potential increase in the frequency of the lemma on the whole. Finally, log-likelihood (LL) tests were conducted using the Lancaster University LL Wizard<sup>12</sup> to test the statistical significance of the corpus data. The next section presents the statistical approach that allows us to empirically validate the associations between (i) individual lemmata and (ii) semantic categories, and the two grammatical constructions in question.

<sup>11</sup> Only the first instance was included in the dataset; the repeated occurrence was excluded to avoid skewing the data.

<sup>12</sup> Accessible at <http://ucrel.lancs.ac.uk/llwizard.html>

### 4.3 Statistical approach

We use distinctive collexeme analysis (DCA) to measure associations between the progressive and the non-progressive and (i) individual lexical verbs and (ii) semantic categories. DCA is a type of collostructional analysis, and takes into account the fact that, while the linguistic context in which a word occurs gives us important information on the semantic and syntactic properties of that word, we also need to take into account the interaction of words and grammatical constructions (Gries & Stefanowitsch 2004: 100; for a critical discussion of collostructional analysis, see Schmid & Küchenhoff 2013). In brief, the DCA ‘identifies lexemes that exhibit a strong preference for one member of the pair [of grammatical constructions] as opposed to the other, and thus makes it possible to identify subtle distributional differences between the members of such a pair’ (Gries & Stefanowitsch 2004: 97). The identification of distinct collexemes<sup>13</sup> is done by assessing the observed and expected frequencies of the co-occurring lexical verbs in relation to the absolute frequencies of both constructions (Hilpert 2006: 244). In practice, the DCA requires a file containing the raw number of all collexemes in each grammatical construction. The data are then subjected to Fisher-Yates tests, which yield the distinctiveness value for each collexeme (i.e. the collostructional strength); the higher the collostructional strength, the more distinct the collexeme. We used Gries’ (2014) *Coll.Analysis 3.5 R script*<sup>14</sup> for the DCA, with the aim of finding out

- (i) which construction each lexical verb is attracted to,
- (ii) what the degrees of association between the verbs and the two constructions are, and
- (iii) to what extent those associations vary across the two datasets (BNC1994 vs. BNC2014), with regard to both attraction and degrees of association.

We perform separate DCAs for each period of time under investigation, the early 1990s and the early 2010s; separate DCAs provide a more accurate picture of the different relative frequencies of lexical verbs, and of the collostructional preferences of the two constructions in question, the progressive and the non-progressive.

Previous research suggests that a DCA can be a useful tool for the analysis of change in the use of stative progressives. Deshors & Rautionaho (2018) perform a DCA of the progressive versus non-progressive alternation in native and non-native varieties of English and are able to empirically validate that Activity situations<sup>15</sup> are strongly attracted to the progressive, while States are strongly attracted to the non-progressive. Interestingly, they find evidence supporting the possible extension of the progressive to non-delimited stative verbs<sup>16</sup> in the second-language varieties studied: while the

<sup>13</sup> Collexemes are the words appearing in the constructions under investigation (Gries & Stefanowitsch 2004: 101).

<sup>14</sup> The script and more information on it are available online at [www.stgries.info/teaching/groningen/index.html](http://www.stgries.info/teaching/groningen/index.html)

<sup>15</sup> Activities are situations that are dynamic, have prolonged duration and do not have a natural end-point (see e.g. Brinton 1988: 27–9).

<sup>16</sup> Non-delimited stative progressives refer to permanent, rather than temporary, situations, and are thus considered non-standard (e.g. *Then what they’ll feel is like, we are knowing each other*; see Sharma 2009).

lemmata HAVE and KNOW are associated with the non-progressive in the native and learner varieties studied, that association does not arise in Indian and Singapore English. This is a good example of how the DCA can reveal fine-grained and subtle distributional differences in the progressive vs. non-progressive alternation, among other alternations of (near-synonymous) grammatical constructions.

Before we move on to the results of our analyses, a potential caveat in our dataset should be discussed. As mentioned in section 3.2, we decided to include all instances of lemmata occurring less than 5,000 times in the two corpora combined, and for verbs with higher frequency, we resorted to a random sample of 1,000 instances per lemma (i.e. 1,000 instances per corpus). This approach leads to samples of varying sizes and proportions of progressives: for some lemmata we have a representation of 100 percent of all instances, while for others (marked with an asterisk in appendix 1) we may have, say, 10 or 80 percent of all instances. For these lemmata, the results remain tentative until all instances extracted from the corpora are edited (a total of 353,745 tokens). Also, the varying sizes of the samples may be problematic for some statistical analyses. However, the Fischer-Yates test does not require any particular sample size (Stefanowitsch & Gries 2003: 101), and thus the distributional differences across our samples are unlikely to affect the rankings of our collexemes (see also Deshors & Rautionaho 2018: 315).

## 5 Results

This section presents the results of our analyses: section 5.1 focuses on the frequency of stative progressives in the 1990s and 2010s, while section 5.2 presents the collostructional preferences of stative verbs. Section 5.3 focuses on developments in the semantic categories of stative verbs.

### 5.1 *Frequency of stative progressives*

Overall, the frequency of stative progressives has not increased in the period between the early 1990s and early 2010s; the proportion of progressive instances of all stative verbs is 6.43 percent in BNC1994DS, and 6.50 percent in BNC2014S (see table 3). Although the raw number of stative progressives has actually increased from 1,204 to 1,645, the number of stative verbs overall has increased, too, thus leveling out the development; this is clearly seen in the normalized figures which are very close to one another (460 and 465, respectively). The frequency of stative verbs at the turn of the twenty-first century is in line with Kranich's (2010) findings (see figure 2).

Our results indicate that the group of stative verbs is highly heterogeneous. Out of the 100 stative verbs investigated, as many as 27 do not occur in the progressive at all (e.g. BELIEVE, EXIST, HATE; see appendix 2 for the full list). Of the verbs that do take the progressive in the 1994 dataset, 20 do so rarely (proportion of progressive instances of all occurrences is less than 3 percent, e.g. BE + adj), 10 verbs moderately frequently, 3 to 15 percent (e.g. IMAGINE), 20 verbs fairly often, 15 to 50 percent (e.g. HOPE), and 9

Table 3. *Number of progressives and non-progressives with stative verbs in BNC1994DS and BNC2014S*

| Corpus    | Progs (pmw) | Non-progs      | Total          | % of prog |
|-----------|-------------|----------------|----------------|-----------|
| BNC1994DS | 1,204 (460) | 17,508 (6,689) | 18,712 (7,149) | 6.43      |
| BNC2014S  | 1,645 (465) | 23,659 (6,686) | 25,304 (7,151) | 6.50      |
| Total     | 2,849 (463) | 41,167 (6,687) | 44,016 (7,150) | 6.47      |

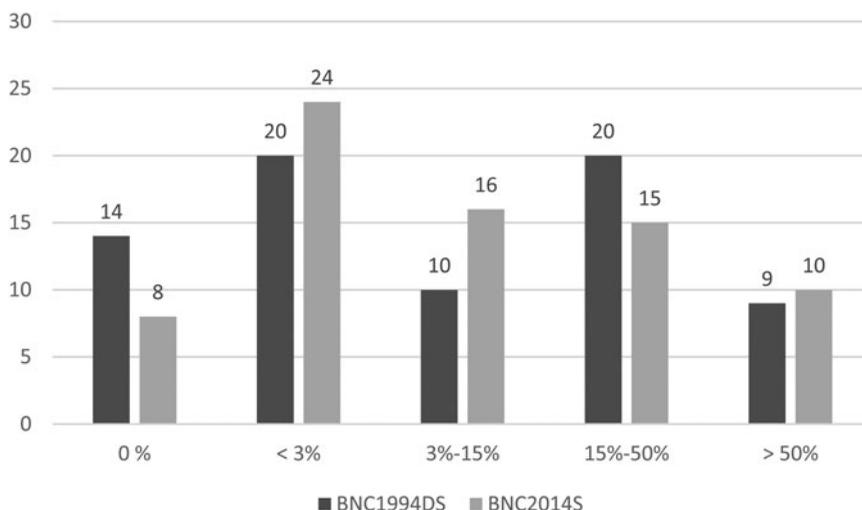


Figure 3. Number of stative lemmata by proportion of progressives

verbs very frequently, above 50 percent (e.g. *WAIT*; see [figure 3](#)). In terms of diachronic change, we find that the number of lemmata occurring in the progressive construction has increased from 59 in 1994 to 65 in 2014.

Looking at developments in the use of individual lemmata, we see change in many directions. Overall, 22 individual lemmata show decreasing proportions of progressive instances, while 26 lemmata show increases. The 14 lemmata which do not occur in the progressive in BNC1994DS have progressive proportions ranging from very low (e.g. *MIND* 0.34 percent) to very high (e.g. *BEAR* 60 percent) in BNC2014S - in most cases, however, the raw number of progressive instances remains very low (*MIND* 0 to 1; *BEAR* 0 to 3). Thus, to test the statistical significance of the developments in individual lemmata, we calculated the log-likelihood (LL) values based on the proportion of progressive instances of all occurrences of an individual lemma; these results show that the increase or decrease of progressive instances between the 1994 and 2014 datasets is statistically significant for 12 individual lemmata (see [figure 4](#); the numerical results are available in [appendix 3a](#)). [Figure 4](#) shows the proportional

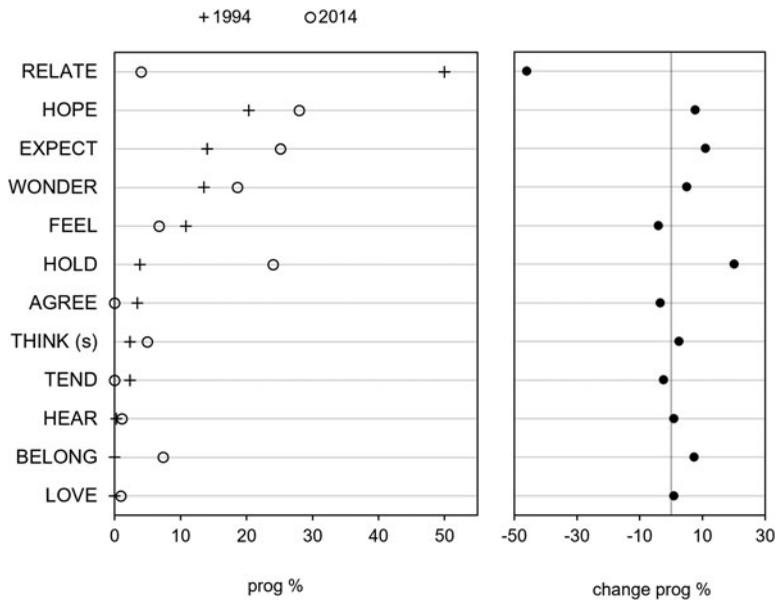


Figure 4. Proportion of progressive usage of all occurrences of a lemma (left-hand panel) and change between BNC1994DS and BNC2014S (right-hand panel)<sup>17</sup>

frequency of progressive instances of all occurrences of the individual lemma with statistically significant changes between the two datasets. However, in some cases the relative frequency per million words is very low, which may affect the proportion of progressive usage. For instance, the diachronic change portrayed in the right-hand panel seems to be rather drastic for *RELATE* and *HOLD*, possibly also *BELONG*, but these lemmata turn out to have very low relative frequencies (*RELATE* decreasing from 0.8pmw to 0.3pmw, *HOLD* increasing from 0.4pmw to 1.7pmw, and *BELONG* from 0pmw to 0.9pmw; see appendix 3b). The lemmata included in figure 4 that actually have a relative frequency higher than 5pmw are *HOPE*, *EXPECT*, *WONDER*, *FEEL* and *THINK* – these are the only stative verbs, out of the 100 investigated, whose relative frequency is high enough to allow detailed analysis, and whose proportional use of progressives has increased or decreased statistically significantly between the 1990s and the 2010s.

As briefly discussed in section 2, *LOVE* has gained attention in the last fifteen years because of McDonald's adopting the progressive *I'm lovin' it* as their slogan in 2003. The present data show that progressive use of *LOVE* has increased from 0 instances in 1994 (before the McDonald's slogan) to 9 (0.92 percent of all instances of *LOVE*) in 2014. Although the proportional frequency remains very low, the semantic restrictions applying to the progressive use of this stative lemma have clearly

<sup>17</sup> This figure and other dotplots were generated with a template provided by Sönning (2016). The 's' after *THINK* refers to 'sample' (see section 3.2).

loosened. In BNC2014S, the nine instances come from eight different speakers in nine different texts and may be described as intensifying an emotion (see Freund 2016: 51–2), as in (20).

- (20) S0012: anyhow now --ANONnameF --ANONnameF is a full-time student  
S0013: yeah (.) she's **loving** it (BNC2014S, SMF3)

A group of lemmata in figure 4 is related to one of the so-called subjective meanings of the progressive, i.e. the tentative use of the progressive (see e.g. Levin 2013: 203–5; Rautionaho 2014: 196). Verbs such as HOPE, WONDER and also THINK may be used to render, for instance, a request more polite by adding an element of tentativeness (as in (21)). The proportion of progressive usage for HOPE and WONDER has increased by 38 and 37 percent respectively,<sup>18</sup> and for THINK (based on a sample of 1,000 instances) by as much as 112 percent.

- (21) I **was** just **wondering** if you could email me and I can come and collect it or something (BNC2014S, SR96)

Besides the tentative use, HOPE, THINK and WONDER, together with EXPECT, which also shows a statistically significant increase by 79 percent, portray a mental process that has duration, which is a meaning compatible with the semantic properties of the progressive construction (as in (22)).

- (22) Is she, is she gonna move or does sh--, cos she **was thinking** of moving (BNC1994DS, KBF)

In the case of THINK, the increasing progressive usage is due to the increasing use of so-called quotative uses (as in (23)), which have risen from 16 to 44 percent. Simultaneously, the proportion of the prototypical meaning 'cogitate' (as in (22)) has decreased from 74 to 47 percent. Levin (2013: 209–11) reports similar findings for American English and ties this development to the process of colloquialization, or drift towards the norms of spoken rather than written language (Mair 2006: 187).

- (23) next minute kitchen door goes we're **thinking** oh my god we're all gonna die (BNC2014S, SAR5)

This section has shown that while the overall frequency of stative progressives has not changed in the period of time between 1994 and 2014, there are still changes in the use of individual lemmata. Some stative verbs investigated were not found in the progressive construction in either dataset, while a few others were found in one dataset but not the other. A number of other lemmata showed either increasing or decreasing use, to varying degrees, and eventually only five lemmata turned out to be both frequent enough and showing statistically significant development between the two datasets (progressive usage of EXPECT, HOPE, THINK and WONDER has increased, while progressive

<sup>18</sup> The change in the percentage between the two datasets was calculated as follows:  $((\text{percentage}_{2014} - \text{percentage}_{1994}) / \text{percentage}_{1994})$ .

usage of FEEL has decreased). What these results, however, are unable to tell us is whether each stative verb is more attracted to the progressive or the non-progressive construction. This aspect is the focus of the next section.

### 5.2 Collostructional preferences: I remember, but I'm wondering

The distinctive collexeme analysis provides a ranking of lemmata that associate most strongly with a grammatical construction: the association is calculated on the basis of the differences between the observed and expected values of the collexemes occurring in the two constructions. Figure 5a portrays the 16 most distinct collexemes for the progressive construction (cut at  $p < 0.01$  in BNC1994DS), and figure 5b the 16 most distinct collexemes for the non-progressive construction.<sup>19</sup> Overall, BNC1994DS contains 27 non-progressive and 21 progressive collexemes with statistically significant collostructional strengths, while the corresponding numbers for BNC2014S are 29 and 22, respectively (see appendix 4). According to the collostructional strengths shown in appendix 4, individual lemmata are associated more strongly with the progressive construction than with the non-progressive; the collostructional strengths are clearly higher for WAIT and SIT associating with the progressive (as in (24) and (25)), than they are for MEAN and HEAR associating with the non-progressive (as in (26) and (27)).<sup>20</sup> In other words, WAIT and SIT are associated with the progressive more strongly than MEAN or HEAR is with the non-progressive. This finding concerns both datasets: the association strengths seem to be slightly higher in the 2014 data overall, indicating that the associations between lemmata and the two constructions have become even stronger, overall.

(24) We're just **waiting** for you grandma. (BNC1994DS, KBW)

(25) I remember he **was sitting** in front of me (BNC2014S, S7K2)

(26) lol do you know what I **mean**? (BNC2014S, S9E6)

(27) **Did** you **hear** that? (BNC1994DS, KDN)

As shown earlier in figure 4, the frequencies and proportional use of the progressive for the lemmata FEEL, HOPE, WONDER, EXPECT and THINK have changed significantly between the two datasets. Figure 5a (esp. the right-hand panel) validates the increase in BE + *hoping*, *wondering* and *expecting* by showing how the association strengths have become stronger in the 2014 dataset. THINK, on the other hand, associates more strongly with the non-progressive, but the association strength has weakened from 10.79 in 1994 to 3.08 in 2014 (see figure 5b). In other words, THINK is still associated with the non-progressive, but the association is weaker in BNC2014S. Regarding FEEL, the lemma associates with the progressive in BNC1994DS (coll.strength 3.25; see appendix 4a), but in BNC2014S the association has turned to the non-progressive,

<sup>19</sup> Coll.strength higher than 3 equals  $p < 0.001$ ; coll.strength  $> 2 \Rightarrow p < 0.01$ ; coll.strength  $> 1.30103 \Rightarrow p < 0.05$ .

<sup>20</sup> For instance, the collostructional strength for WAIT is 128.149 in BNC1994DS and 108.2626 in BNC2014S; for MEAN it is 29.6204 in BNC1994DS and 28.8914 in BNC2014S.

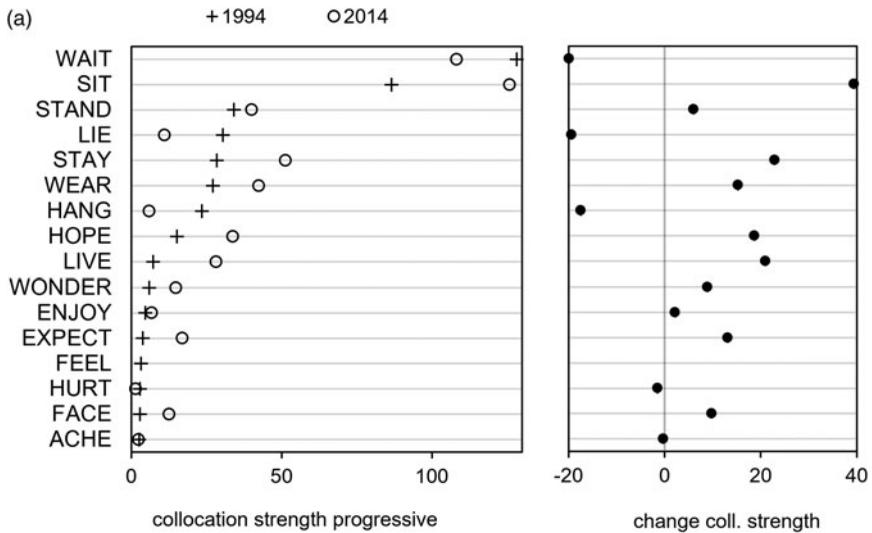


Figure 5a. The most distinctive collexemes for *progressive* (left panel), and the change between BNC1994DS (the vertical line) and BNC2014S (the dot) (right panel)

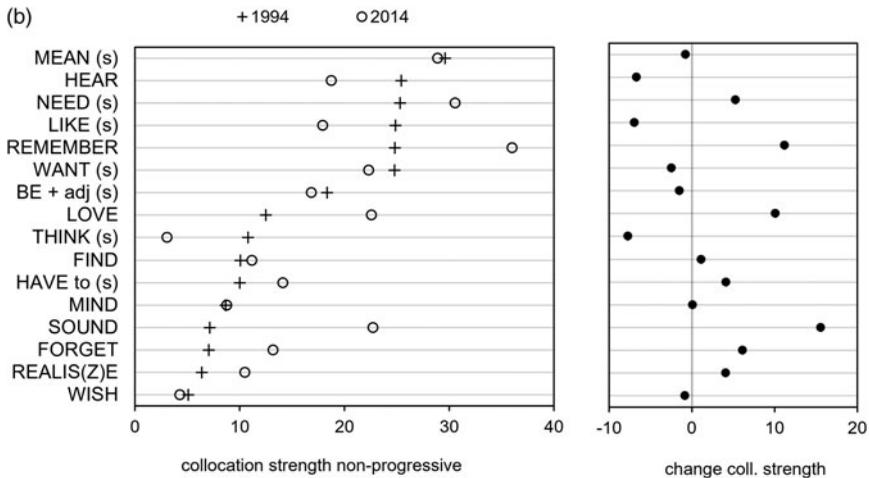


Figure 5b. The most distinctive collexemes for *non-progressive* (left panel), and the change between BNC1994DS (the vertical line) and BNC2014S (the dot) (right panel)

albeit not statistically significantly (coll.strength 0.91). This validates the decrease in the proportion of progressive usage reported in [figure 4](#).

Regarding *LOVE*, which showed a subtle increase in [figure 4](#) (by 0.92 percent), the DCA actually shows that the increase in frequency is not validated; *LOVE* associates strongly with the non-progressive and the association becomes even stronger in the 2014 dataset (from

12.49 to 22.59; see [figure 5b](#)). Thus, the nine occurrences of *loving* in BNC2014S remain individual occurrences rather than signs of LOVE becoming integrated in the progressive paradigm; Mair (2006: 93) calls such occurrences ‘occasional instances of contextually licenced rule-breaking’.

Finally, let us see what our results tell us about BE (followed immediately by an adjective) and HAVE (in the sense ‘possess’). These two lemmata are interesting in the sense that their inclusion in the progressive paradigm has been argued to have affected the general development and rise in frequency of the progressive construction by allowing the loosening of the traditional semantic constraints regarding the progressive and stative verbs (Smitherberg 2005: 156, also Kranich 2010). In the case of BE + adj, the proportion of progressives of all occurrences remains low: 0.18 percent in 1994, and 0.62 percent in 2014,<sup>21</sup> and the DCA shows that BE + adj is associated statistically significantly with the non-progressive. The collostructional strength decreases only slightly, from 18.35 in 1994 to 16.86 in 2014, indicating that BE + adj remains inimical to the progressive. In the present data, BE *being* + adj expresses negative evaluation of another person’s behavior (as in (28); other collocates *boring*, *critical*, *mean* and *nasty*) – by virtue of expressing a temporary state, the progressive renders the criticism less face-threatening (see Levin 2013: 195–9).

(28) cos I’ve gotta stay in with my class cos they’re **being horrible** (BNC2014S, SJNB)

HAVE (in the sense ‘possess’) has slightly higher proportions of progressives, 2.80 and 2.41 percent, and the association is with the non-progressive. However, compared to BE + adj, the collostructional strength is clearly lower at 1.51 for BNC1994DS ( $p < 0.05$ ) and 2.43 for BNC2014S ( $p < 0.01$ ), which means that the progressive is more likely with HAVE than it is with BE + adj.<sup>22</sup>

So far, we have considered developments in individual stative verbs: next, we turn our attention to groups of stative verbs.

### 5.3 *Types of stative verb: spread to new semantic domains?*

The lemmata attracted to the non-progressive are semantically different from those attracted to the progressive, as the list of lemmata statistically associated with either grammatical construction reveals (see [figures 5a](#) and [5b](#)). Lemmata attracted to the non-progressive, e.g. MEAN, LIKE, and BE + adj., belong to the semantic categories of cognitive, affective and relational verbs (respectively). In total, 23 of 29 of the lemmata associated with the non-progressive in BNC2014S, and 22 of 27 in BNC1994DS, belong to one of these categories. On the other hand, lemmata attracted to the

<sup>21</sup> It must be noted that both BE+adj and HAVE are included in the set of highly frequent lemmata for which we included only a sample of 1,000 tokens. However, we did edit all progressive instances of these verbs; these results show that the number of progressive instances of BE+adj decreases from 39pmw to 29pmw, while for HAVE we find an increase from 1pmw to 11pmw. This shows that the results based on samples of lemmata should be considered tentative.

<sup>22</sup> HAVE occurs more frequently in the progressive when a dynamic reading is valid (as in *having a snack/a cup of tea/a garden party*); these, however, were not included in the present dataset.

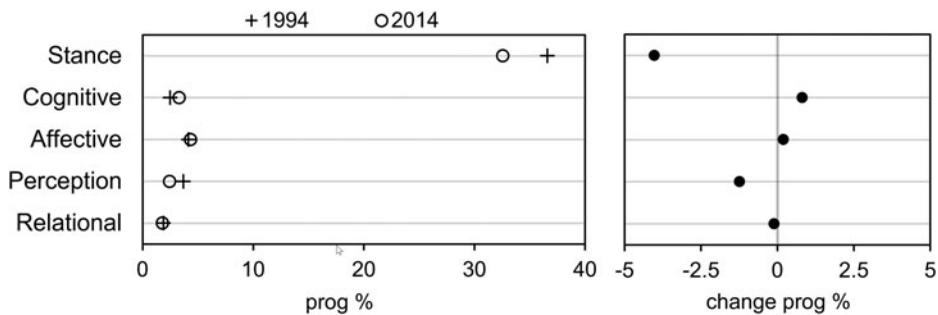


Figure 6. Proportion of progressive usage in semantic categories (left-hand panel), and change between BNC1994DS (the vertical line) and BNC2014S (the dot) (right-hand panel)

progressive, e.g. *WAIT*, *SIT* or *STAND*, are usually stance verbs, and in fact many of the lemmata statistically significantly associated with the progressive are indeed stance (-like) (8 of 21 in BNC1994DS and 9 of 22 in BNC2014S). This agrees well with their characterization as being intermediate between stative and dynamic meaning and overall most likely to combine with the progressive (see section 2 and Quirk *et al.* 1985: 205). As figure 6 shows, approximately a third of all occurrences of stance verbs occur in the progressive (37 percent for BNC1994DS and 33 percent for BNC2014S), while the proportion of progressive usage is rather low for cognitive, relational, perception and affective verbs (below 4.5 percent).

However, in relative terms, stance verbs as a group have seen a decrease in the proportion of progressive instances, as have perception verbs (see the right-hand panel in figure 6, and appendix 5). By contrast, cognitive verbs have seen an uptake in progressive usage. Finally, relational and affective verbs underwent only small, statistically not significant changes.

The uptake in progressive usage with cognitive verbs (LL value 8.44,  $p < 0.01$ ) might indicate a loosening of the semantic restrictions applying to cognitive states (on such semantic restrictions, see Declerck 1991: 168). Within this group of (in the present data) 24 lemmata, there is considerable variation. Some cognitive verbs do not occur in the progressive at all (e.g. *KNOW*, *SUPPOSE*; as in (29)), others occur in the progressive moderately frequently (e.g. *FIND*, *THINK*), and yet others very frequently (e.g. *EXPECT*, *WONDER*; as in (30)). Further, the uptake in progressive usage seems to be associated with a few individual lemmata: *EXPECT* and *THINK* are used in the progressive clearly more frequently in the 2014 dataset (+79 percent and +112 percent, respectively, both  $p < 0.01$ ). Rather than an uptake in progressive usage among cognitive verbs in general, the increase concerns only a few individual lemmata.

(29) I **don't know** what it does actually. (BNC1994DS, KCD)

(30) well I'**m expecting** an inspection soon (BNC2014S, SQPN)

Similarly, perception verbs vary in their propensity to take the progressive (see section 2), with some not readily occurring in the progressive (e.g. *SEE*, *HEAR*), while others occur as

often in the progressive as in the non-progressive (e.g. *ITCH*, *HURT*; Declerck 1991: 167). This distinction is very clear in the present data, too: verbs of bodily sensation (*ACHE*, *HURT*, *FEEL*, *ITCH* and *TICKLE*; as in (31)) have clearly higher proportion of progressive usage (ranging from 12 to 60 percent) than verbs referring to involuntary reaction (e.g. *SOUND*, *TASTE*; 0 to 4 percent; as in (32)). Overall, the proportion of progressive usage with perception verbs has fallen from 3.7 to 2.4 percent (LL value 9.54,  $p < 0.01$ ).

(31) yeah it's a hard lesson my brain's **hurting** (BNC2014S, SJ5D)

(32) **Did** that **taste** alright, that lasagne? (BNC1994S, KD3)

The third group of verbs with statistically significant changes between the two datasets is stance verbs, for which the proportion of progressive usage has decreased from 37 to 33 percent (LL value 5.48,  $p < 0.05$ ). Although the non-progressive (as in (33)) is common with stance verbs, corresponding progressive instances (as in (34)) seem to be more frequent. Semantically, the two sentences are similar, and in each case the grammatical construction could be replaced with the other possibility without major change in the meaning.

(33) we just both **sat** there thinking exactly the same thing (BNC2014S, S38F)

(34) we **were** all **sitting** together talking (BNC2014S, SKPP)

Let us now turn to the results of the distinctive collexeme analysis of the progressive and non-progressive constructions and the five semantic categories, to see whether they confirm the findings of the analysis of relative frequencies above. Again, stance verbs are found to behave differently from the other groups as they are the only group that is statistically significantly attracted to the progressive construction (see appendix 6). For all other groups, the association is with the non-progressive, and in all cases the association strength is statistically very highly significant (coll.strength is higher than 3). As figure 7 shows, the association has become even stronger for relational, perception and affective verbs, whereas it has weakened slightly for cognitive verbs. This validates the proportional increase of progressive use reported earlier; at least some cognitive verbs now operate with looser semantic constraints.

Interestingly, some stative verbs are polysemous, which influences to what degree they are associated with the progressive. For instance, *FEEL* has meanings that can be categorized as cognitive (as in (35)), perception (as in (36)) and affective (as in (37)).

(35) I'm **feeling** I should cancel my class tomorrow (BNC2014S, SMCW)

(36) **Are** you **feeling** alright? Do you **feel** blurgh blurgh blurgh? (BNC1994DS, KPX)

(37) I **feel** so sorry for the kids nowadays (BNC2014S, SD2R)

The relative frequencies indicate that the cognitive use has increased (from 2 percent of all progressive occurrences of *FEEL* in 1994 to 23 percent in 2014), and the affective use more moderately from 19 to 28 percent, whereas the perception use has decreased (from 78 to 50 percent). A DCA measuring the association between the three meanings of *FEEL* reveals that while cognitive uses are statistically significantly associated with the non-progressive ( $p < 0.001$ ), the perceptive use is drawn towards

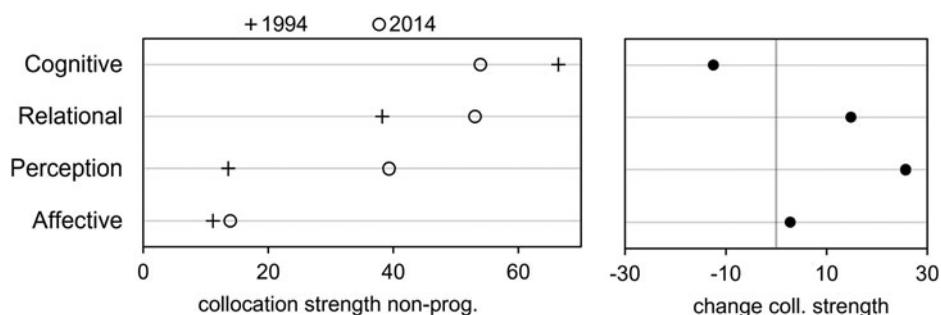


Figure 7. Diachronic change in collocations strength for the non-progressive by semantic category

the progressive ( $p < 0.001$ ).<sup>23</sup> Affective uses of *FEEL* are associated with the non-progressive, but this association is not statistically significant. Thus, the polysemous meaning of *FEEL* explains the fact that association with the progressive in BNC1994DS (see appendix 4a) turns into a non-significant association with the non-progressive (coll.strength 0.91) in BNC2014S; the rise of the cognitive meaning (associated with the non-progressive) affects the overall collostructional behavior of the lemma. This rise can be considered an example of the generalization of the progressive (see Levin 2013: 215–16).

## 6 Discussion

This study set out to investigate recent changes in stative progressives from the point of view of the progressive versus non-progressive alternation, by means of measuring relative frequency by text size (per million words), in the variable context (percentage) and its collostructional preferences. Thus, we combined frequency measures with a distinctive collexeme analysis that helped us to identify statistically significant associations between the two grammatical constructions and individual lexical verbs, as well as types of stative verbs, and to validate the findings of the first part of the study. Specifically, we have addressed four research questions:

- (i) Are stative progressives more frequent in the present day, compared to the early 1990s?
- (ii) Is the progressive applied to a broader range of stative verbs in the present day, compared to the early 1990s?
- (iii) Is the progressive applied to a broader range of semantic domains of stative verbs in the present day, compared to the early 1990s?
- (iv) Has there been change to the collostructional preferences of stative verbs, and stative categories between the early 1990s and the present day?

With regard to (i), we find no substantial change in the frequency of stative progressives in the last twenty-five years, with just a small increase from a proportion of progressive uses

<sup>23</sup> The association strengths are as follows: for cognitive uses, BNC1994DS 11.39, BNC1994S 9.98; for affective uses BNC1994DS 0.29, BNC2014S 0.60; for perceptive uses BNC1994DS 7.91, BNC2014S 17.17.

of all stative verbs of 6.43 to 6.50 percent. Regarding (ii), we do see statistically significant changes in 12 individual stative verbs; four lemmata saw a drop in the frequency of progressive usage (RELATE, FEEL, AGREE and TEND), while eight showed an increase (HOPE, EXPECT, WONDER, HOLD, THINK, HEAR, BELONG and LOVE). Furthermore, the number of lemmata not occurring in the progressive has fallen from 14 in BNC1994DS to 8 in BNC2014S, indicating that a broader range of individual lemmata is used in the progressive now, compared to the 1990s. Regarding (iii), the analysis revealed that cognitive verbs have seen an uptake in progressive usage, while stance and perception verbs have seen a fall in the proportion of progressive use. The increasing frequency of cognitive verbs in the progressive may indicate a loosening of the semantic constraints. However, the increase is mostly propelled by the increasing progressive use of individual lemmata (esp. WONDER and EXPECT) rather than the group as a whole.

Finally, with regard to (iv), the collostructional analysis of individual lemmata revealed that association strengths are higher for the progressive construction than the non-progressive in both datasets, indicating that lemmata associated with the progressive are much more strongly attracted to that grammatical construction than lemmata associated with the non-progressive. The DCA validated the relative changes detected in the frequency-based analysis regarding individual lemmata such as HOPE, EXPECT and THINK. Regarding groups of semantic verbs, the DCA clearly showed that stance verbs differ from all other groups by being the only one to be statistically significantly associated with the progressive. Coupled with the high proportional usage of progressives, these findings suggest that stance verbs should be considered distinct from stative verbs in general.

On a methodological note, the benefit of combining the DCA with the more traditional relative frequency data becomes evident in figures 8a and 8b, which combine the results from the two analyses. In both corpora, the majority of collexemes have low relative frequency and low collostructional strength, which explains the concentration of data points in the bottom left corner. The collexemes that have higher than usual relative frequency or collostructional strength, or both, are shown to diverge upwards and towards the right-hand side of the figure.<sup>24</sup>

In BNC1994DS, as shown in figure 8a, BE *hanging* and BE *lying* both have high relative frequencies (c. 89 and 77 percent, respectively), but their collostructional strengths are not exceptionally high (23.46 and 30.44, respectively). BE *waiting*, on the hand, has both a high relative frequency (c. 78 percent) and the highest association strength of all collexemes (128.15). Further, BE *sitting* has the second highest association strength (86.49), while the relative frequency is more moderate at c. 40 percent. These results are explained by the subtle differences in the observed and expected values of the collexemes (provided by the DCA with the Coll.Analysis 3.5 script) in relation to the absolute frequencies of the two grammatical constructions. For instance, despite the very high relative frequency for BE *hanging* seen in figure 8a, the difference between the observed frequency of progressive instances (24) and the expected frequency (2.08)

<sup>24</sup> Only these lemmata are labeled in figures 8a and 8b, in order to make the labels readable.

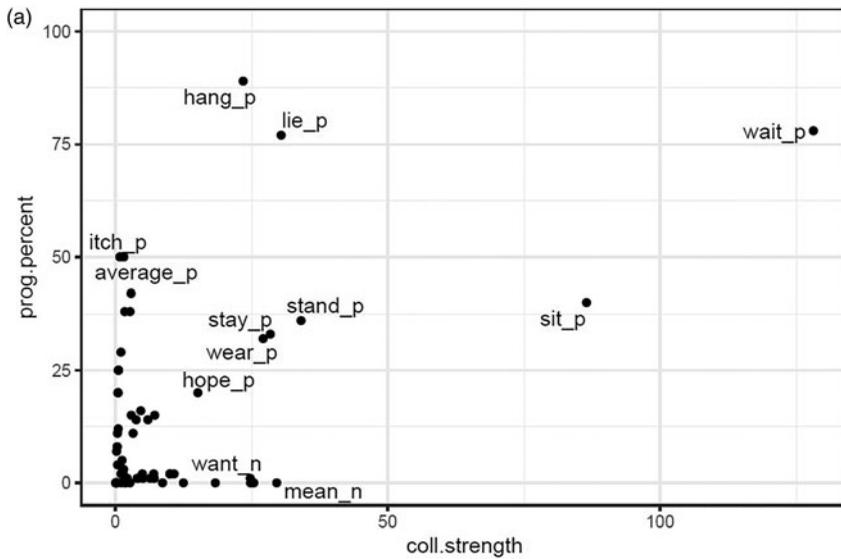


Figure 8a. Comparison of relative frequencies of progressive usage and collostructional strengths in BNC1994DS

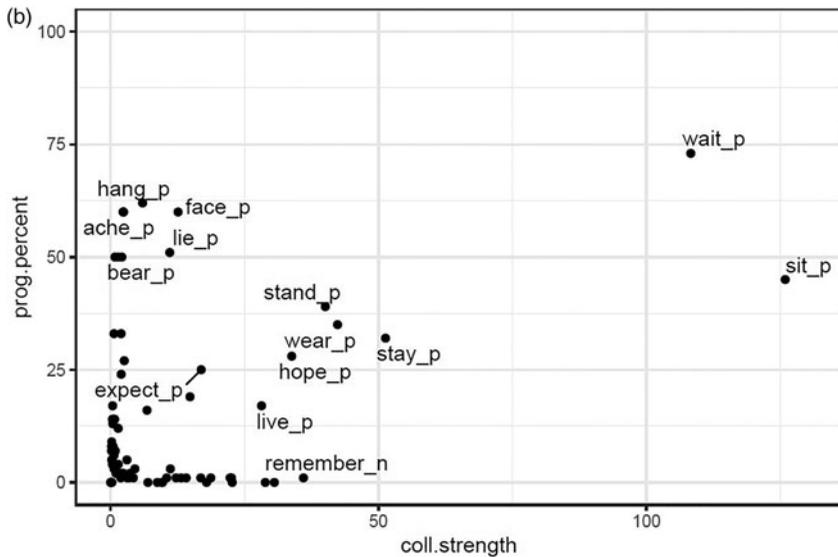


Figure 8b. Comparison of relative frequencies of progressive usage and collostructional strengths in BNC2014S

is smaller than the respective values for BE *waiting* (observed 148, expected 14.54). Comparing the two datasets, there are more collexemes diverting from the mass in the bottom left-hand corner in BNC2014S (figure 8b), indicating that a higher number of

collexemes are statistically significantly attracted to the progressive in the latter dataset. Furthermore, what is notable in both figures is the fact that the collexemes that are distinct from the others are all ones attracted to the progressive. This suggests that progressive usage of stative verbs is restricted to some lemmata, and further, that these lemmata are outliers that are strongly attracted to the progressive. In a sense, then, lemmata associated with the progressive could be seen as unusual compared to the mass of stative verbs; many stative verbs do occasionally occur in the progressive, but those that are statistically associated with the construction are very much attracted to it. Thus, more broadly, the present analysis confirms that stative verbs as a group fall outside the domain of progressive use and that it is mainly a few exceptional stative verbs that are frequently used with the progressive.

As mentioned briefly in section 5.1, at least some developments in the progressive usage of stative verbs may be connected to colloquialization (see e.g. Smith 2005; Mair 2006; Leech *et al.* 2009; Levin 2013). The quotative use of BE *thinking* (as in (23) and (38)) is an example of the increasing use of informal and interpersonal language connected to colloquialization (Levin 2013: 210), while other developments in the use of stative progressives may be better explained through the related process of democratization of language use, or ‘the tendency to avoid unequal and face threatening modes of interaction’ (Farrelly & Seoane 2012: 393; see also Leech *et al.* 2009: 259). The process of democratization is relevant for the progressive in so far as it is increasingly used to express tentativeness and politeness, which in turn can be seen as a sign of attuning to the ‘changing norms in society towards less emphasis on power and social distance’ (Levin 2013: 215). Lemmata such as HOPE and WONDER are specifically suited for expressing tentativeness (as in (39) and (40)); progressive use of both was noted to have increased in the present study, from 20 to 28 percent in the case of HOPE, and from 14 to 19 percent in the case of WONDER (see figure 4, and appendix 3). Further, both HOPE and WONDER are statistically significantly associated with the progressive, and the association has strengthened between the 1990s and 2010s (see figures 5a and 5b). Another collexeme used to express politeness, by avoiding face-threatening expressions, is BE *being* + adj (as in (28) repeated here as (41)). According to Levin (2013: 193), the use of the progressive rather than the simple construction distances the speaker from the potential threat to the addressee’s face by rendering the portrayed characteristic temporary rather than permanent. Thus, democratization might be one of the factors leading to the recent increase of stative progressives in BrE, as the lemmata whose progressive usage has increased significantly between the 1990s and the 2010s, and whose relative frequency is higher than 5pmw (i.e. EXPECT, HOPE, THINK and WONDER), may all be used to express tentativeness and politeness and avoid potentially face-threatening acts and signs of power disparities between participants.

- (38) they need more models and case studies uhu so I **was thinking** why not give it a go  
(BNC2014S, SP9F)
- (39) oh I **was hoping** that you’d come and see us (BNC1994DS, KC0)

- (40) er I'm **wondering** if did you guys get invited to – ANONnameF's wedding?  
(BNC2014S, SP2X)
- (41) cos I've gotta stay in with my class cos they're **being horrible** (BNC2014S, SJNB)

Finally, the spread of the progressive use of stative verbs has been considered one of the main explanations for the overall increase of the progressive. However, the present study reveals that the deceleration of the increase detected in the twentieth century (e.g. Leech *et al.* 2009; Smith & Leech 2013) has turned into a complete halt: stative progressives have stopped their spread in recent conversational BrE. What does this suggest for the progressive construction overall – has it, too, reached the end of the spread? To investigate the matter, we searched the same datasets, BNC1994DS and BNC2014S, for the combination of any form of the auxiliary BE followed immediately by the present participle of the main verb.<sup>25</sup> This tentative comparison suggests that the increase in the frequency of the progressive in BrE has indeed halted: the frequency in 1994 is 6,278pmw, while in 2014 it is 6,124pmw, which means a decrease by 2.5 percent (LL value 5.79,  $p < 0.05$ ). Although tentative, this result is interesting since it indicates, for the first time, that the overall increase of the progressive in spoken English may indeed have reached the plateau at the end of the S-curve, while in written English, its spread may continue for some time due to ongoing colloquialization. While our results suggest that the rise of stative progressives, too, has been halted and is confined to particular lemmata, this increase in the few particular lemmata might, however, suffice to sustain the popular impression of a general, continuing spread.

*Authors' addresses:*

*English Language and Culture, School of Humanities*  
*University of Eastern Finland*  
*PO Box 111*  
*FIN-80101 University of Eastern Finland*  
*Finland*  
[paula.rautionaho@uef.fi](mailto:paula.rautionaho@uef.fi)

*Institute of English and American Studies*  
*Faculty of Humanities*  
*Universität Hamburg*  
*Überseering 35*  
*22297 Hamburg*  
*Germany*  
[robert.fuchs@uni-hamburg.de](mailto:robert.fuchs@uni-hamburg.de)

<sup>25</sup> Search phrase was '{be/V} \*ing V\*'. This analysis is preliminary; the data are likely to include occurrences that are not actually progressives and the analysis disregards occurrences where BE and *-ing* are separated by intervening words. Future studies will need to determine whether our preliminary finding of a complete halt can be confirmed.

## Supplementary material

The supplementary material for this article can be found at <https://osf.io/eqamv/>

## References

- Aarts, Bas, Joanne Close & Sean Wallis. 2010. Recent changes in the use of the progressive construction in English. In Bert Cappelle & Naoaki Wada (eds.), *Distinctions in English grammar; offered to Renaat Declerck*, 148–67. Tokyo: Kaitakusha.
- Aarts, Bas, Joanne Close, Geoffrey Leech & Sean Wallis (eds.). 2013. *The verb phrase in English: Investigating recent language change with corpora*. Cambridge: Cambridge University Press.
- Aston, Guy & Lou Burnard. 1998. The BNC handbook: Exploring the British National Corpus with SARA. <http://corpus.leeds.ac.uk/teaching/aston-burnard-bnc.pdf> (accessed 5 December 2018).
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Brinton, Laurel J. 1988. *The development of English aspectual systems: Aspectualizers and post-verbal particles*. Cambridge: Cambridge University Press.
- Declerck, Renaat. 1991. *Tense in English: Its structure and use in discourse*. London: Routledge.
- Deshors, Sandra C. & Paula Rautionaho. 2018. The progressive versus non-progressive alternation: A semantic exploration across World Englishes. *English World-Wide* 39(3), 309–37.
- Farrelly, Michael & Elena Seoane. 2012. Democratisation. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford handbook of the history of English*, 392–401. Oxford: Oxford University Press.
- Freund, Nina. 2016. Recent change in the use of stative verbs in the progressive form in British English: *I'm loving it*. *University of Reading Language Studies Working Papers* 7, 50–61.
- Fuchs, Robert & Valentin Werner. 2018. The use of stative progressives by school-age learners of English and the importance of the variable context: Myth vs. (corpus) reality. *International Journal of Learner Corpus Research* 4(2), 195–224.
- Granath, Solveig & Michael Wherrity. 2014. *I'm loving you – and knowing it too: Aspect and so-called stative verbs*. *Rhesis: Linguistics and Philology* 4(1), 2–22.
- Greenbaum, Sidney & Randolph Quirk. 1990. *A student's grammar of the English language*. London: Pearson English Language Teaching.
- Gries, Stefan Th. 2014. *Coll.analysis 3.5: A script for R to compute perform collocation analyses*.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collocation analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9(1), 97–129.
- Hardie, Andrew. 2012. CQPweb – combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics* 17(3), 380–409.
- Hilpert, Martin. 2006. Distinctive collexeme analysis and diachrony. *Corpus Linguistics and Linguistic Theory* 2(2), 243–56.
- Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Kranich, Svenja. 2010. *Progressive in Modern English: A corpus-based study of grammaticalization and related changes*. Amsterdam: Rodopi.
- Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicholas Smith. 2009. *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press.
- Levin, Magnus. 2013. The progressive verb in modern American English. In Aarts *et al.* (eds.), 187–216.

- Love, Robbie, Claire Dembry, Andrew Hardie, Vaclav Brezina & Tony McEnery. 2017. The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics* 22(3), 319–44.
- Mair, Christian. 2006. *Twentieth-century English: History, variation and standardization*. Cambridge: Cambridge University Press.
- Martínez-Vásquez, Montserrat. 2018. *I'm loving it!* A corpus-based study of the progress of love. *Journal of English Linguistics* 46(2), 140–66.
- Paulasto, Heli. 2014. Extended uses of the progressive form in L1 and L2 Englishes. *English World-Wide* 35, 247–76.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Rautionaho, Paula. 2014. *Variation in the progressive: A corpus-based study into World Englishes*. Tampere: Tampere University Press.
- Sharma, Devyani. 2009. Typological diversity in New Englishes. *English World-Wide* 30, 170–95.
- Schmid, Hans-Jörg & Helmut Küchenhoff. 2013. Collostructional analysis and other ways of measuring lexicogrammatical attraction: Theoretical premises, practical problems and cognitive underpinnings. *Cognitive Linguistics* 24(3), 531–77.
- Smith, Carlota. 1991. *The parameter of aspect*. Dordrecht: Kluwer.
- Smith, Nicholas Ian. 2005. A corpus-based investigation of recent change in the use of the progressive in British English. PhD dissertation, Lancaster University.
- Smith, Nicholas & Geoffrey Leech. 2013. Verb structures in twentieth-century British English. In Aarts *et al.* (eds.), 68–98.
- Smitterberg, Erik. 2005. *The progressive in 19th-century English: A process of integration*. Amsterdam: Rodopi.
- Sönning, Lukas. 2016. The dot plot: A fine tool for data visualization. In Hanna Christ, Daniel Klenovšak, Lukas Sönning & Valentin Werner (eds.), *A blend of MaLT: Selected contributions from the Methods and Linguistic Theories Symposium 2015*, 101–32. Bamberg: University of Bamberg Press.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8, 209–43.