Recent Changes in the Use of the Progressive Construction in English*

Bas Aarts,a Joanne Close b and Sean Wallis a

a University College London and b University of Leeds

1. Introduction

A classical distinction has entrenched itself in linguistics, namely the diachronic and synchronic ways of studying a language. The first considers language in its stages of development, whereas the latter looks at languages viewed from the present moment. This old Saussurean dichotomy has recently been called into question, and it has been argued that the distinction is artificial (see, for example, Labov (1972)). Instead, it is argued that languages change all the time, even within the synchronic phases. As a result of these new attitudes to language development there has emerged a new research impetus in linguistics which concerns itself with what has been called recent change or current change (see Mair (1995, 1997), Denison (1998, 2001, 2004), Hundt (2004), Smitterberg (2000, 2005), Núñez-Pertejo (2007), and Aarts, López-Couso and Méndez-Naya (forthcoming)).

Christian Mair at Freiburg was the first to construct parallel corpora of written British and American English spanning four decades in the twentieth century (the LOB/FLOB and Brown/Frown corpora). These are excellent resources enabling linguists to research changes in written English over 30 years. Manual searches are still unavoidable, however, as these corpora have not been parsed.

At the Survey of English Usage (UCL) we have taken Mair’s initiative further by constructing a corpus of British English comprising selections of largely spontaneous spoken English from the London-Lund Corpus (dating from the late 1950s to early 1970s) and from the British Component of the International Corpus of English (ICE-GB; dating from the 1990s). This corpus, which we have called the Diachronic Corpus of Present-Day Spoken English (DCPSE; Aarts and Wallis (2006)), allows researchers to investigate recent changes in the grammar and usage of Present-Day English over a period of 25–35 years. DCPSE differs from FLOB and Frown in a number of important ways. Firstly, the corpus is unique in containing exclusively spoken English. We opted for a corpus of spoken English because it is generally recognised that spoken language is primary, and the first locus of changes in lexis and grammar. Secondly, the corpus is parsed, which will permit research into synchronic and diachronic grammatical variation. Thirdly, the corpus is fully searchable using the International Corpus of English Corpus Utility Program (ICECUP), the corpus exploration software that we developed for ICE-GB. DCPSE is already being used as a major new resource complementing the Freiburg corpora.

In this paper we will look at the changing use of a particular grammatical construction in English, namely the progressive, which has recently been receiving a lot of attention. Our data are derived from DCPSE. We will show how it can be used to perform grammatical searches in spoken English.

2. Changes in the Use of the English Progressive: Previous Studies

It is commonly accepted that the progressive increased in frequency during the nineteenth century (see e.g. Denison (1998), Hundt (2004), Smitterberg (2000, 2005), Núñez-Pertejo (2007), and Aarts, López-Couso and Méndez-Naya (forthcoming)). Recent research has shown that the nineteenth century trend of an increase in the frequency of use of the pro-

* This research was carried out as part of the project The changing verb phrase in present-day British English at the Survey of English Usage, UCL, funded by the UK Arts and Humanities Research Council (grant number AH/E006299/1). We gratefully acknowledge the AHRC’s support.
Recent Changes in the Progressive

The progressive has persisted into the twentieth century. Hundt (2004) uses Mossé’s (1938) M-coefficient, which normalises the frequency of the progressive to occurrences per 100,000 words, to track the frequency of the progressive from 1650 to 1990 in ARCHER (A Representative Corpus of Historical English Registers). Her results indicate a rise in the frequency of the progressive in the twentieth century (the lower line in Figure 1).

Mair and Leech (2006: 323) investigate the increased use of the progressive using the Brown quartet of corpora. These are four matching corpora of one million words each from 1961 and 1991/92. Table 1 shows that for Press subcorpora in British English the use of the progressive seems to be advancing more quickly than in American English. British English has a higher frequency of progressives than American English in 1961 and the use of the progressive increases by a larger percentage between 1961 and 1991/92.

Table 1: Progressive Forms in the Press Sections (A-C) of Four Equal Size Reference Corpora (LOB, FLOB, Brown and Frown).

<table>
<thead>
<tr>
<th>Press</th>
<th>1961</th>
<th>1991/92</th>
<th>% rise from 1961</th>
</tr>
</thead>
<tbody>
<tr>
<td>British English (LOB/ FLOB)</td>
<td>606</td>
<td>716</td>
<td>+18.2%</td>
</tr>
<tr>
<td>American English (Brown/Frown)</td>
<td>593</td>
<td>663</td>
<td>+11.8%</td>
</tr>
</tbody>
</table>

Overall this research points to an increased use of the progressive in Present-Day English, a trend that has continued from the nineteenth century. Our aim in this paper is to contribute to work on the progressive by investigating the construction systematically in spoken English, using the Diachronic Corpus of Present-Day Spoken English.

1 ARCHER consists of written British and American English across a range of genres dating between 1650 and 1990. It was first constructed by Douglas Biber and Edward Finegan in the 1990s, and the latest release, ARCHER 3.1, was completed in 2006.

2 The upper line in Figure 1 shows absolute frequencies rather than M-coefficients, but Kranich asserts that, aside from the first half-century, the number of words per half-century do not differ greatly.

3 Press reportage consists of 44 texts, editorials 27 texts, and reviews 17 texts.

4 Mair and Leech (2006) claim that this trend is statistically significant. It is certainly large, but the data presented does not allow us to meaningfully draw this conclusion. Significance must be considered in relation to a choice between alternate forms. See section 3 for discussion.
3. Exploring DCPSE to Research the Use of the English Progressive

DCPSE allows us to monitor the use of the English progressive in spoken English over a number of decades. Before showing how this can be done we will introduce a few general features of the corpus, and discuss more precisely how the relative frequency of the progressive should be defined.

3.1. The Diachronic Corpus of Present-Day Spoken English

DCPSE was released by the Survey of English Usage (SEU) in 2006. It contains 400,000 words of 1960s spoken material from the London-Lund Corpus (LLC), and 400,000 words of 1990s spoken material from ICE-GB in matching text categories. DCPSE includes a wide range of spoken English, such as face-to-face conversations, telephone conversations, various types of discussions and debates, legal cross-examinations, business transactions, speeches and interviews. Much of DCPSE is spontaneous, which is important because changes in English propagate themselves in the first instance through spontaneous discourse. It is possible for researchers to listen to the spoken material. As in the previously released ICE-GB, every sentence in DCPSE is syntactically annotated with a phrase structure tree diagram, like the one shown in Figure 2 below.

---

5 The LLC is the spoken part of the Survey of English Usage Corpus, founded by Randolph Quirk in 1959. It contains 510,576 words of 1960s spoken English, is prosodically annotated, and has been used—and continues to be used—by many scholars for their research. ICE-GB is composed of both spoken and written material from the 1990s. It contains textual markup, and is fully grammatically annotated.

---

6 Gloss (features are in italics): PU=parse unit, CL=clause, main=main, intr=intransitive; ingp=ing participle; SU=subject, NP=noun phrase, NPHD=NP head, PRON=pronoun, pers=personal, plu=plural; VB=verbal, VP=verb phrase, OP=operator; AUX=auxiliary; MVB=main verb, V=verb, prog=progressive; encl=enclitic; pres=present; A=adjunct; AVP=adverb phrase; AVHD=adverb phrase head; ADV=adverb; ge=general.
Recent Changes in the Progressive

3.2. Investigating How the Use of the Progressive Varies

There are a number of ways in which the frequency of the progressive can be calculated in a diachronic study such as this one. A variationist study of the progressive would measure its frequency of use against the frequency of any grammatical variants. The obvious variant for the progressive is, of course, the simple form, although it is clear that the progressive cannot replace any verb phrase in the simple form.

In corpus studies like those by Nehls (1988) and Smith (2002), the progressive is calculated per 100,000 words (using the ‘M-coefficient’ (Mossé (1938)). Calculating the frequency of the progressive in this way is simple, but it has a major flaw: it does not take into account the possibility that the number of verb phrases per 100,000 words may not be stable diachronically. Nor does it guarantee that the opportunity for a progressive to be used is uniform, i.e. as compared to a baseline of ‘progressives plus alternate variants,’ as noted in Wallis (2003). We need to treat the graph in Figure 1 with some caution because it shows the variation of the total number of progressives in each sample. Strictly speaking, in these studies, we do not know whether the total number increases over the time period because (a) there are actually more progressives as a proportion of all alternate forms, or (b) there are simply more circumstances where progressives could be used.

Smitterberg (2005) discusses this issue and a range of others in his study of the progressive in nineteenth century English, and he compares the frequency of progressive use using the M-coefficient and his own ‘S-coefficient.’ This S-coefficient is a formula which calculates the number of finite progressives as a proportion of finite verb phrases (excluding the be going to future and what he refers to as ‘knock-out’ factors, i.e. contexts where the progressive cannot appear; see below).

Smitterberg demonstrates how the chosen methodology affects the results: the S-coefficient shows that the progressive has increased by 81% between periods 1 and 3 (1800–1900), whereas using the M-coefficient puts the increase at 71% (Smitterberg (2005: 62)). The difference between these percentage figures means that in this particular dataset, the opportunity to use the progressive (measured as the number of finite verb phrases, excluding ‘knock out’ cases) has actually fallen over time, and the S-coefficient can be defended as a more precise measure. Smitterberg also found that if the M-coefficient is used, the progressive is most common in Drama, followed by Letters, Trials, Fiction, History, Debates and Science; but if the S-coefficient is used, the progressive is more common in Letters than in Drama, and more frequent in History than in Fiction (Smitterberg (2005: 77–78)). In conclusion, simple normalised frequencies are potentially misleading, particularly in the case of differing genres (where the opportunity of progressives being used may indeed vary) and small samples (see also Nelson et al. (2002: 260)).

Smitterberg (2005: 46) lists a number of ‘knock-out factors’ in calculating the progressive. These include ‘demonstrations’ (I take this hat), per-
formatives (I name this ship Elizabeth), simple imperatives, non-finite verb phrases and stative situations. As discussed by Smitterberg, some of these factors are easier to exclude than others. Imperatives, for example, can easily be removed from any corpus which is tagged, whereas removing stative verb phrases requires manual checking of each example, a time-consuming process.

In calculating the use of the progressive in DCPSE, we follow Smitterberg (2005) in measuring its use against the number of verb phrases, taking knock-out factors into account. As Smitterberg’s study was based on nineteenth century English, some modifications are made. Firstly, we have not excluded stative verbs from the study; Mair and Leech (2006: 324) point out that in twentieth century English the progressive may occur with stative verbs, although occurrences are too infrequent to account for the statistically significant overall increase of the progressive with such verbs. Secondly, in order to exclude demonstrations and performatives, as Smitterberg does, each example would need to be manually checked. As they are rare and unlikely to affect the results, they have not been removed. 8

3.3. The Progressive in DCPSE

We used FTFs to look for progressives in DCPSE. The FTF below instructs the search engine to search for a progressive VP (note the feature ‘progressive’ in the bottom section of the node).

---

8 Smitterberg (2005: 47) also excludes non-finite VPs (progressive and non-progressive) on the grounds that in terms of complementation the choice seems to be between a to-infinitive and present participle (e.g. She continued to read vs. She continued reading), rather than between a non-progressive and progressive to-infinitive (e.g. She continued to read vs. She continued to be reading). It seems to us that this depends on the verb in the superordinate clause. In PDE with a verb like pretend, for example, the choice does seem to be between non-progressive and progressive to-infinitive (cf. She pretended to read vs. She pretended to be reading). In order to exclude non-finite verb phrases, each example would have to be checked. This was beyond the scope of this paper.

---

Fig. 4: An FTF Used to Search for Progressive VPs

Studies using corpora that are not grammatically parsed must, at this stage, remove instances of the be going to future (see, for example, Smitterberg (2005)). However, the be going to future is not marked as progressive in DCPSE and so is automatically excluded. Our results are summarised in Table 2.

<table>
<thead>
<tr>
<th>(spoken)</th>
<th>progressive</th>
<th>not progressive</th>
<th>Total</th>
<th>χ²(prog)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LLC (1960s–70s)</td>
<td>2,973 (4.73%)</td>
<td>59,906</td>
<td>62,879</td>
<td>31.23</td>
</tr>
<tr>
<td>ICE-GB (1991/92)</td>
<td>3,294 (5.72%)</td>
<td>54,305</td>
<td>57,599</td>
<td>28.61</td>
</tr>
<tr>
<td>Total</td>
<td>6,267</td>
<td>114,211</td>
<td>120,478</td>
<td>59.84s</td>
</tr>
</tbody>
</table>

Table 2: Changes in the Proportion of Progressive VPs in the LLC and ICE-GB components of DCPSE. The result is significant (‘s’) for p<0.01.

The table shows that in the LLC portion of DCPSE, out of a total of 62,879 verb phrases that could have been ‘progressivised,’ 2,973 were progressive (4.73%), while in the ICE-GB part, out of 57,599 verb phrases, 3,294 were progressive (5.71%). Formally, we can say that we have refuted the null hypothesis that speakers did not change their behaviour regarding the use of the progressive between the two periods. 9

9 Strictly speaking, we should add ‘for language data consistent with the sampling of the corpus.’ One possibility is that the way that samples were collected by Quirk et al. (1985) differs from the methodology of Sidney Greenbaum’s team in the 1990s, and this explains the result.
Comparing the spoken content in the ICE-GB and LLC corpora with the written LOB and FLOB corpora, Smith (2005) found that progressives were almost twice as frequent in spoken rather than written English over the same period (Table 3).

Using DCPSE we can obtain the incidence of the progressive over time. The result is shown in Table 4. Since the total number of cases per year is relatively low, to be on the safe side we also calculate the binomial confidence interval (‘error’ column). The best way of understanding this error estimate is by looking at a row in the table. In the first sample from 1958, of the 1,731 cases of plausible progressive VPs, 52 (2.94%) were progressive with an error of ±0.79% (i.e. from 2.15% to 3.73%). As can be seen by looking at the next row for 1959, the margin of error is greater.

As a general rule, the fewer the total number of cases for any year, the larger the margin of error and the greater the likelihood that our progressive percentage is not representative of the population. To make sense of the trend it is helpful to plot this distribution on a timeline. The graph in Figure 5 shows annual data points with the 95% confidence interval expressed as a T-shaped error bar. The graph illustrates the fact that DCPSE does not include data for the period between 1978 and 1989, but this does not prevent us estimating a trend (dotted line). We have also added the centre points (indicated by the ‘X’ symbols) for the LLC and ICE-GB subcorpora from Table 4.

### Table 3: Spoken and Written Language Compared (from Smith (2005))

<table>
<thead>
<tr>
<th></th>
<th>Spoken progressive</th>
<th>per million words</th>
</tr>
</thead>
<tbody>
<tr>
<td>LLC (1960s-70s)</td>
<td>2,396</td>
<td>5,990</td>
</tr>
<tr>
<td>ICE-GB (1990s)</td>
<td>3,153</td>
<td>7,882</td>
</tr>
<tr>
<td>Written</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOB (1961)</td>
<td>2,932</td>
<td>2,916</td>
</tr>
<tr>
<td>FLOB (1990s)</td>
<td>3,202</td>
<td>3,176</td>
</tr>
</tbody>
</table>

### Table 4: Use of the Progressive in DCPSE over Time

<table>
<thead>
<tr>
<th>Year</th>
<th>Progressive</th>
<th>Error (%)</th>
<th>Not Progressive</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1958</td>
<td>52 (2.94%)</td>
<td>±0.79%</td>
<td>1,717</td>
<td>1,769</td>
</tr>
<tr>
<td>1959</td>
<td>47 (6.02%)</td>
<td>±1.67%</td>
<td>734</td>
<td>781</td>
</tr>
<tr>
<td>1960</td>
<td>138 (6.02%)</td>
<td>±0.97%</td>
<td>2,155</td>
<td>2,293</td>
</tr>
<tr>
<td>1961</td>
<td>239 (5.11%)</td>
<td>±0.63%</td>
<td>4,437</td>
<td>4,676</td>
</tr>
<tr>
<td>1963</td>
<td>26 (5.04%)</td>
<td>±1.89%</td>
<td>490</td>
<td>516</td>
</tr>
<tr>
<td>1964</td>
<td>102 (4.49%)</td>
<td>±0.85%</td>
<td>2,172</td>
<td>2,274</td>
</tr>
<tr>
<td>1965</td>
<td>129 (4.68%)</td>
<td>±0.79%</td>
<td>2,629</td>
<td>2,758</td>
</tr>
<tr>
<td>1966</td>
<td>153 (4.65%)</td>
<td>±0.72%</td>
<td>3,139</td>
<td>3,292</td>
</tr>
<tr>
<td>1967</td>
<td>132 (5.63%)</td>
<td>±0.93%</td>
<td>2,214</td>
<td>2,346</td>
</tr>
<tr>
<td>1969</td>
<td>122 (3.89%)</td>
<td>±0.68%</td>
<td>3,017</td>
<td>3,139</td>
</tr>
<tr>
<td>1970</td>
<td>112 (7.04%)</td>
<td>±1.26%</td>
<td>1,478</td>
<td>1,590</td>
</tr>
<tr>
<td>1971</td>
<td>280 (4.83%)</td>
<td>±0.55%</td>
<td>5,518</td>
<td>5,798</td>
</tr>
<tr>
<td>1972</td>
<td>99 (3.23%)</td>
<td>±0.63%</td>
<td>2,968</td>
<td>3,067</td>
</tr>
<tr>
<td>1973</td>
<td>46 (6.43%)</td>
<td>±1.80%</td>
<td>669</td>
<td>715</td>
</tr>
<tr>
<td>1974</td>
<td>338 (4.68%)</td>
<td>±0.49%</td>
<td>6,884</td>
<td>7,222</td>
</tr>
<tr>
<td>1975</td>
<td>657 (4.63%)</td>
<td>±0.35%</td>
<td>13,544</td>
<td>14,201</td>
</tr>
<tr>
<td>1976</td>
<td>249 (4.52%)</td>
<td>±0.55%</td>
<td>5,256</td>
<td>5,505</td>
</tr>
<tr>
<td>1977</td>
<td>52 (5.55%)</td>
<td>±1.47%</td>
<td>885</td>
<td>937</td>
</tr>
<tr>
<td>1990</td>
<td>272 (4.88%)</td>
<td>±0.57%</td>
<td>5,306</td>
<td>5,578</td>
</tr>
<tr>
<td>1991</td>
<td>2,283 (5.66%)</td>
<td>±0.23%</td>
<td>38,028</td>
<td>40,311</td>
</tr>
<tr>
<td>1992</td>
<td>739 (6.31%)</td>
<td>±0.44%</td>
<td>10,971</td>
<td>11,710</td>
</tr>
</tbody>
</table>

---

Note that the figures in Table 2 differ from those in Table 3 reported by Smith (2005). They are LLC: 2,973 cases (6,406 per million words) and ICE-GB: 3,294 cases (7,824 per million words). Smith’s observation of a higher usage in speech remains sound.

The simple binomial confidence interval for a probability (or percentage) is calculated by the following formula.

\[ e = z_{crit} \sqrt{p \cdot (1-p) / N} \]

where \( z_{crit} \) is the critical value of \( z \) for a given confidence level, \( p \) is the probability of the event occurring (in this case, that the VP is progressive) and \( N \) is the total number of cases (i.e. applicable VPs). \((1-p)\) is the probability that the VP is not progressive.

Note that for a 95% confidence interval, which we adopted for our data, \( z_{crit} \) is approximately 1.96.
4. Why Has the Progressive Increased in Use?

Mair (2006: 88–89) comments that there are three types of changes affecting the progressive:

(i) many uses which were fully established around 1900 have increased in frequency since then;
(ii) new forms have been created; and
(iii) there is a tendency to use the progressive with stative verbs such as understand (see also above and below).

Smith (2005) suggests the following factors as probable causes of the increase in the use of the progressive in recent times.

(i) Contact—the progressive is more common in American English than in British English (Biber et al. (1999: 462)) and the growing contact between the two countries may have contributed to the increased usage in British English.
(ii) Increased functional load—‘[T]he progressive has evolved historically such as to convey a rather complex meaning, or set of meanings’ and ‘probably as a result of the varied and developing nature of its meanings, the progressive has enjoyed a meteoric increase in frequency of use.’ (Smith (2005: 2))

Regarding an increase in functional load, Nesselhauf (2007) studied the ‘progressive futurate’ (e.g. She is graduating next week), and observed that it tripled in usage between 1750 and 1990. Other linguists, Wright (1994, 1995), Smith (2005) and Smitterberg (2005) among them, have suggested that the so-called ‘interpretive,’ ‘explanatory’ or ‘modal’ progressive, as in (1) and (2) below, has also led to an increase in frequency of the progressive construction in British English, particularly the present progressive.

(1) If John says that, he’s lying.
(2) When I said the ‘boss’, I was referring to you.

(Huddleston and Pullum et al. (2002: 165))

This use of the progressive “interprets the speaker’s attitude and perspective of the situation; and, in so doing, conveys her epistemic stance at a particular moment in the context of utterance” (Wright (1995: 157)). As Smith (2005: 166) puts it: “Interpretatives are often considered to signal a higher degree of pragmatic meaning and/or subjectivity on the part of the speaker than regular uses of the progressive.” For Quirk et al. (1985: 198 fn. b) “the event described has an interrelationship or identity with another simultaneous event,” and Leech (2004b: 22) observes that “it is as if we are seeing the speech act ‘from the inside,’ not in a temporal sense, but in the sense of discovering its underlying interpretation.” Huddleston and Pullum et al. (2002: 165) note that ‘regular’ progressives in English trigger what they call a ‘mid-interval implicature.’ Thus if I say I was reading a book in answer to the question What were you doing last night?, there is an implicature that my reading of the book was not coextensive with the beginning and end of the evening. In other words, with such progressives there is an implicature that the time referred to by the progressive is part of a larger situation. This inference is not an entailment,
Recent Changes in the Progressive  

because it is cancellable. Huddleston and Pullum et al. (2002) observe that when the interpretive progressive is used, the mid-interval implicature is always cancelled.

Smitterberg (2005: 222) records 364 examples of interpretive progressives over the 19th century as a whole in his corpus of 1 million words, with absolute frequencies steadily increasing over his three sub-periods. The interpretive use has also increased during the twentieth century according to Smith (2005), as Table 5 shows.

<table>
<thead>
<tr>
<th>Genre</th>
<th>LOB (1961) frequency per million</th>
<th>FLOB (1991) frequency per million</th>
</tr>
</thead>
<tbody>
<tr>
<td>Press</td>
<td>9</td>
<td>19</td>
</tr>
<tr>
<td>General prose</td>
<td>23</td>
<td>40</td>
</tr>
<tr>
<td>Learned</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Fiction</td>
<td>20</td>
<td>31</td>
</tr>
<tr>
<td>Total</td>
<td>52 (52)</td>
<td>97 (96)</td>
</tr>
</tbody>
</table>

Table 5: Genre Distribution of the Interpretive Use of the Progressive (Present Tense), in LOB and FLOB, based on the Clearest Cases (Smith (2005: 196)). Numbers in brackets are means.

A word of caution is in order at this point. In our view it would be prudent not to draw too firm conclusions from the data in the various sources we discussed regarding the increased use of the interpretive progressive, given the differences between scholars’ definitions of the concept and the difficulty of detecting an interpretive meaning in individual examples.

As noted above in discussing Mair’s work, another use of the progressive construction that may have contributed to its increased frequency can be found in the examples below.

(3) I’m lovin’ it! (McDonald’s slogan)
(4) I’m loving every moment with you.
(5) Who’re you wanting to seduce? (DCPSE:DI-C01 #0211:1:A)

Stative verbs like love and want do occur in the progressive, although for many speakers the simple present is still the expected form. This usage is not new. Denison (1998: 146) records some examples from as early as 1803 and 1820, and an example with the verb love from 1917. Nevertheless, it would be fair to say that until quite recently utterances like (3) and (4) would have been less marked if they contained a verb in the simple present tense. Mufwene (1984: 36) presents a ‘scale of stativity,’ from punctual (‘least stative’) verbs to the ‘highest stative’ verbs as shown in (6) below (simplified).

(6) High: e.g. contain, know, belong to, consist of, need, concern, matter, owe
Intermediate: e.g. love, hate, depend, want, intend, wish
Neutral: e.g. enjoy, wait, stay, stand, lie; revolve, turn, work, run, read, write; call, claim, speak, say
Punctual/low: e.g. kick, reach, crack, die, break, hit, etc.

The reports on the progressive discussed in section 2 above suggest that the progressive is spreading up the scale in Present-Day English; currently it is often found with stative verbs such as love, wish and want, and perhaps in the future we will see an increase in use with verbs such as know, need, etc. Interestingly, in DCPSE there is an example of know in the progressive from 1961.

(7) We will compare a play written in the Restoration Period with something that happened in Elizabethan times and we assume that our students are knowing what we are talking about you see. (DCPSE:DL-A01#0512)

5. Conclusions

In this paper we have shown how the Diachronic Corpus of Present-day Spoken English can be used to track short-term changes in the use of the progressive. The dataset shows an increased use of the construction in recent times, possibly due to a wider range of uses. In order to investigate this it would be necessary to explore the data further to differentiate these uses.

We have additionally demonstrated why it is very important to consider
the frequency of the progressive relative to the possibility of it being used in the first place. Many studies have considered progressive use in terms of normalised absolute frequencies, such as the M-coefficient, i.e. frequencies considered in proportion to the total number of words in a particular dataset. However, the opportunity to use any linguistic construction, including the progressive, may not be constant between different time periods or genres. The danger is that we end up measuring two things at the same time—(i) the opportunity to use the progressive combined with (ii) the decision to use the progressive, once the opportunity has arisen. Since we are interested in whether people increasingly choose to use the progressive, we must measure usage relative to opportunity.

A big advantage of using a parsed corpus like DCPSE is that in many cases the detailed grammatical analysis makes it easier to identify the set of cases where the opportunity for an event to occur arises. Ideally, we would wish to count the set of true alternates, i.e. those cases where we can say that the speaker could have chosen to use the progressive, but did not. This could be done by checking every VP in the corpus, or by estimating the total number of true alternates by inspecting a random subsample. Both options would be time-consuming. What we have done instead is assume that by discounting known VP forms which cannot take the progressive (i.e. taking into account so-called ‘knock-out’ factors, pace Smitterberg), we arrive at the set of contexts where the option of using the progressive arises, thus obtaining a clearer picture of the increase in the use of the progressive in English.

References
Aarts, Bas, Joanne Close, Geoffrey Leech and Sean Wallis (editors) (forthcoming) Current Change in the English Verb Phrase.
Leech, Geoffrey (2004a) “Recent Grammatical Change in English: Data, Description, Theory,” Advances in Corpus Linguistics, ed. by Karin Aijmer and Bengt Altenberg, 61–81, Rodopi, Amsterdam.
Leech, Geoffrey and Nicholas Smith (2009) “Change and Constancy in Linguistic
Nelson, Gerald, Sean Wallis and Bas Aarts (2002)
Mossé, Ferdinand (1938)
Mair, Christian (2006) Twentieth-Century English. History, Variation and
More Frequent in English? A Corpus-based Investigation of Language
Change in Progress,” Zeitschrift für Anglistik und Amerikanistik 43.2, 111–122.
Language Change in Progress,” The Handbook of English Linguistics, ed. by Bas Aarts and April McMahon, 318–342, Blackwell, Malden, MA.
Mossé, Ferdinand (1938) Histoire de la Forme périphrastique être + participe
Club, Bloomington.
Verbal Aspect in English,” Essays on the English Language and Applied
Linguistics: On the Occasion of Gerhard Nickel’s 60th Birthday, ed. by Josef
Klegraf and Dietrich Nehls, 173–198, Julius Groos Verlag, Heidelberg.
Nelson, Gerald, Sean Wallis and Bas Aarts (2002) Exploring Natural Language:
On Going

Raphael Salkie

University of Brighton

Mmm, folks won’t find us now because
Mister Satch and Mister Cros
We gone fishin’ instead of just a-wishin’
Bah-boo-bah-boo-bah-bay-mmm-bo-bay
Oh yeah!

Gone Fishin’ (Lyrics and music by Nick and Charles Kenny, 1951)

1. To Go Drinking vs. ??to Go Eating

In a squib in Linguistic Inquiry, Arlene Berman (1973) pointed out some interesting restrictions on expressions like We(‘ve) gone fishing, which we shall refer to as ‘expeditionary go.’ As well as fishing, Louis Armstrong and Bing Crosby could with full grammaticality have gone camping, shopping, visiting or travelling, but probably not *working, *dining or *smoking. To relax afterwards there was no linguistic reason for them not to go drinking, but they are unlikely to have ??gone eating.

Typical corpus examples are:

(1) So it looks as though we’ll be able to go shopping tomorrow by the weather forecast.

(2) I might be going swimming at lunchtime today so I could probably run up there or put it in the post if I go.

(3) One day the young lord went hunting with his hound in a densely thicketed part of the forest.