

John Benjamins Publishing Company



This is a contribution from *International Journal of Corpus Linguistics* 12:3
© 2007. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

Lexical-grammatical patterns in spoken English

The case of the progressive with future time reference

Nadja Nesselhauf and Ute Römer
University of Heidelberg / University of Hanover

Based on a large set of data from one of the biggest available corpora of spoken British English (the 10-million word spoken component of the BNC), this article explores central lexical-grammatical aspects of progressive forms with future time reference. Among the phenomena investigated are verb preferences, adverbial co-selection, subject types, and negation. It is demonstrated that future time progressives in spoken British English are patterned to a considerable extent (for example that it is individual verbs, rather than semantic groups of verbs, that preferably occur in such constructions) and that actual language use often runs counter to claims that can be found in traditional grammatical descriptions of the construction. A number of general and often neglected issues in the analysis of lexical-grammatical patterns are also addressed, in particular the notion of pattern frequency.

Keywords: progressives, future time reference, lexical grammar, spoken English, corpus-driven analysis, pattern frequency

1. Introduction

While the lexical-grammatical patterns of the future time expressions *will*, *shall* and *going to* have received a considerable amount of attention, particularly in recent years (cf. for example Berglund (1997, 2000); Szmrecsanyi (2003); Krogvig and Johansson (1984); Fries (1925)), the patterns of the progressive with future time reference (as in *He's arriving tomorrow*) have hardly been empirically investigated to date.¹ Statements on the use of the progressive with future time reference are therefore frequently based on unsystematic observation and also often rather vague (see Section 3). In this paper, we would like

to contribute to filling this gap by providing a comprehensive corpus-driven analysis of lexical-grammatical patterns of the progressive with future time reference in spoken British English.

Our analysis is based on the spoken part of the British National Corpus (referred to as 'BNC_spoken' in what follows), which contains 10 million words of recorded and transcribed informal and formal conversations, radio- and TV-newscasts, lectures, interviews, business meetings etc. We will first investigate which verbs and groups of verbs are used, and which of these are predominantly used, in the progressive with future time reference. We will then consider other lexical-grammatical factors that have been shown or hypothesized to play a role in the choice of future time expressions and/or that might reveal co-selection phenomena in the use of this construction. The factors taken into account are the presence or absence of an adverbial, the type of subject occurring in the relevant clause, the form of TO BE, whether the construction is negated or not, and whether it occurs in a declarative or interrogative clause. We will also investigate possible interrelations of some of these factors.

2. Method

The analysis presented below is based on the 100 most frequent verbs in the spoken part of the BNC (according to the frequency list provided by BNCweb, which is based on the occurrences of verbal infinitives; for more details on the methodology cf. Römer (2005a: 47ff)).² For each of these verbs, a search for all *-ing* forms in BNC_spoken was performed. To obtain a manageable amount of data, a random sample of 200 was retrieved (from a maximum number of different BNC_spoken files) if there were more than 200 occurrences of a single verb in the *-ing* form; if there were 200 or less, all occurrences were retrieved. All the *-ing* forms were classified according to whether they constitute part of a progressive, and the progressives were then subclassified according to their time reference. As a considerable number of instances of progressives appear to refer to more than one time sphere at once and/or appear to have general validity, four types of time reference were distinguished: past time reference, present time reference, future time reference, and 'indeterminate'. An example of the latter type is *Erm but if you're dealing with issues about human relationship then they're very very important*. It has to be pointed out, however, that the decision of whether a given instance of the progressive refers to the future, to the present, or to several time spheres at once is not always easy to make (on this problem cf. also Mair and Hundt (1995)), in particular when investigating

spoken language with its false starts, interruptions etc. Nevertheless, a decision — the one which seemed to be most justified by the context — was made for all instances. In the present analysis, only those instances classified as having clear future time reference were included. To avoid distortion of the results by the random sampling procedure, for those verbs where the analysis of the random sample revealed that less than (an arbitrarily chosen threshold of) 1/30th of the progressives with future time reference in the whole of BNC_spoken had apparently been retrieved, the random sample was extended accordingly. This was necessary for two verb forms, *doing* and *going*.³ For *doing*, a random sample of 400 was thus retrieved in total, and for *going*, a random sample of 800. Instances of *going to* + verb were excluded from the investigation.

3. The progressive with future time reference

3.1 Verb preferences — earlier accounts

With respect to the verbs or types of verbs that can occur in the progressive with future time reference, verbs of motion or semantically related groups of verbs have traditionally been described as being the main ones which permit this use. H. E. Palmer, for example, states that “only a few verbs” are possible with the progressive with future time reference. The verbs he lists are COME, GO, START, LEAVE, and STAY (Palmer 1939:154, 281). F. R. Palmer writes that “with the progressives, reference to the future is very largely restricted to verbs of motion or verbs which at least imply motion”, including in this category, however, verbs such as MEET and JOIN (Palmer 1965:88). While stressing that the construction “is available for any verbs referring to action that people can plan”, Close (1977:150) claims that it is particularly frequent “with verbs referring to departure and arrival”. Jespersen also stresses the use of this construction with “verbs of movement”, in particular COME and GO (1909–49 Vol. IV:214), but also mentions DIE (and related verbs), DINE (and related verbs), and other verbs indicating “an agreement or appointment with regard to the future” (p. 218f). Zandvoort claims that the future use of the progressive mainly occurs with verbs denoting “a coming or leaving” (1972:76). More recent accounts often tend to describe the restriction, or preference, in more general terms. Leech, for example, states that the use of the construction is restricted “in the main to ‘doing’ verbs involving conscious human agency” (2004:63).

In what seems to be the first published corpus-based study on the present progressive (and present non-progressive) with future time reference (based

on some 270,000 words of written British English), Aarts finds that “[a]lmost 95% of the 50 verbs or so in the corpus that are so used [i.e. in the present with future time reference] are process verbs” (presumably Aarts is referring to types, not tokens, here). He also finds that “of these, verbs of coming and going form a clear minority” (Aarts 1969:578).

The only other corpus-based or corpus-driven studies on the use of the progressive with future time reference we are aware of are Mindt (1992), Wekker (1976), and Römer (2005b). Römer (2005b) is partly based on the same set of data as the present study (for differences cf. Note 2), but only looks at a few selected aspects of the use of the progressive with future time reference and has a different focus than the present study, namely the comparison of the use of the progressive in spoken language with the use in textbooks for foreign language learners. Mindt’s results are based on two corpora, a 170,000 word corpus of conversation and a 184,000 word corpus of plays. Wekker’s (1976) results are based on a corpus of 500,000 words of novels, newspapers, and periodicals, and 100,000 words of TV and radio recordings. Unfortunately, Wekker does not provide a more detailed analysis of the question of which verbs are possible in this construction. He finds, however, that the “construction is not limited to verbs denoting or implying motion, and some others (as has often been argued); it also occasionally occurs with verbs that are static” (Wekker 1976:110). His results also indicate that “[i]t is not true that all the verbs that can occur in the progressive can refer to the future” (cf. for example **It’s raining tomorrow*) and that “there are some verbs, e.g. *see* and *hear*, which do not normally combine with the progressive aspect, but which are quite common as progressives with future meaning” (ibid.).

Mindt, on the other hand, does not find any progressives with future time reference with SEE OR HEAR or any with other verbs of sensual perception (such as LOOK OR LISTEN), and neither does he find any with state verbs (such as BE OR BECOME), with the verb HAVE, with verbs of cognition such as KNOW, FORGET, LEARN OR UNDERSTAND, or with the ‘verbs of discovering’ FIND, DISCOVER, and DETECT (cf. Mindt 1992:72ff and 239). He does, however, find instances of the construction in the groups of verbs listed in Table 1.⁴

Verbs of position or motion clearly predominate in Mindt’s corpora, with over 66% of all tokens belonging to this category. Verbs of doing and the individual verb GET are also fairly frequent; other verbs are rare. Unfortunately, Mindt does not indicate with how many instances the individual verbs in the single groups are present in the corpus, and in most cases it is not fully clear which verbs exactly belong to the different groups.

Table 1. Tokens of progressives with future time reference in two corpora investigated by Mindt (1992: 72ff and 239)

Verbs or groups of verbs	CONV	PLAYS
Verbs of position or motion (e.g. COME, GO, LEAVE)	38	101
GET	10	5
Verbs of doing (e.g. MAKE, DO)	7	7
BRING, GIVE, TAKE	6	–
Verbs of verbal communication (e.g. TELL, TALK, SAY, SPEAK)	5	–

3.2. Verb preferences — the analyses

In our analysis, 1,117 instances of the progressive with future time reference were retrieved from BNC_spoken, with the methodology outlined above. For comparison, Mindt (1992) and Wekker (1976) investigated 209 and 90 instances, respectively (while the number of instances in the corpus investigated by Aarts (1969) remains unclear). The 100 verb types included in our study are represented with between zero tokens (for example *seeming* and *wondering*) and 57 tokens (*leaving*) of the progressive with future time reference (for all individual results, see the Appendix). As many of these numbers were obtained by examining only a random sample of the total number of occurrences of the *-ing* form in question (cf. Section 2), the results were extrapolated to give an indication of the total number of instances of the verb in the progressive with future time reference in BNC_spoken.⁵ The results, i.e. the approximate absolute number of occurrences of progressives with future time reference for each of the 100 most frequent verbs in BNC_spoken, are displayed in Figure 1.

These results show that there are considerable differences between the different verbs as regards their use in the progressive with future time reference. While *going* (excluding the *going to* future, cf. Section 2) and *coming* probably occur around 1,000 times in this function in the 10 million words of the whole corpus, and *doing*, *taking* and *getting* occur around 500 times, half of the 100 investigated verbs occur 20 times or less, 15 verbs perhaps not at all. Some examples of the most frequently occurring verb forms in context are provided in (1) to (8) below.

- (1) I'm *going* to the party and I'm gonna take ice. (fm7)
- (2) you can skip chapel this Sunday, and we'll go out er we're *going* into Derbyshire (fy4)
- (3) Erm what time are you *coming* out to play football? (fmf)
- (4) I'm not *coming* here any more. (h4h)
- (5) Ask Terry, what are we *doing* Er next (f7g)

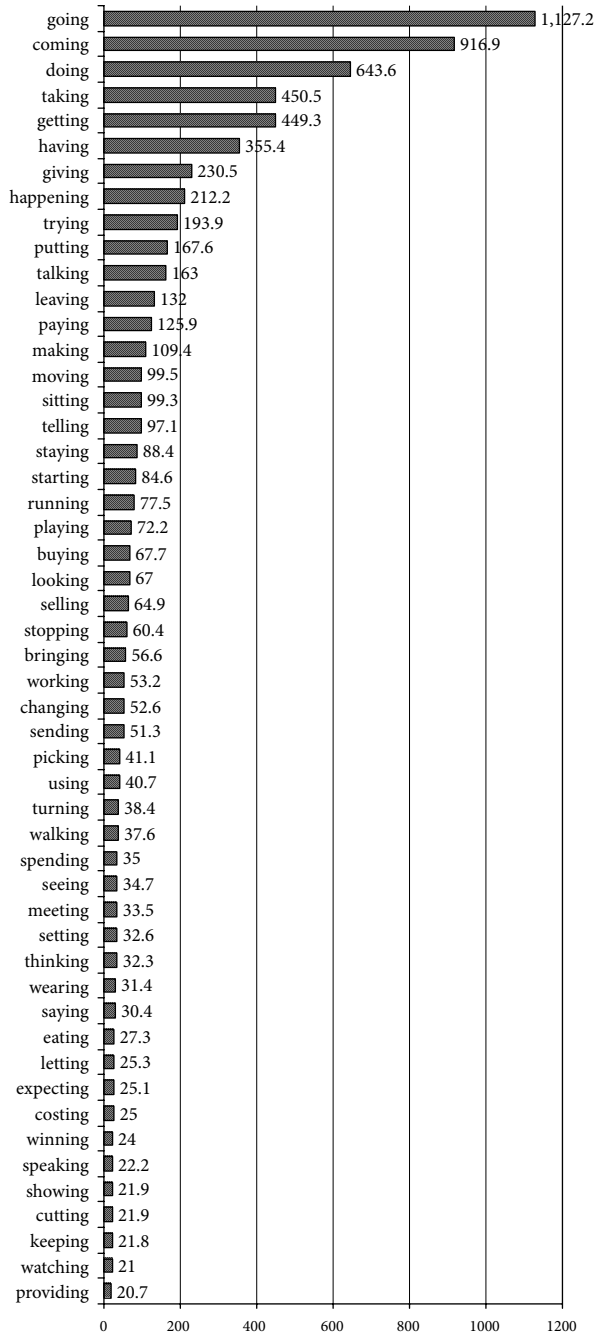
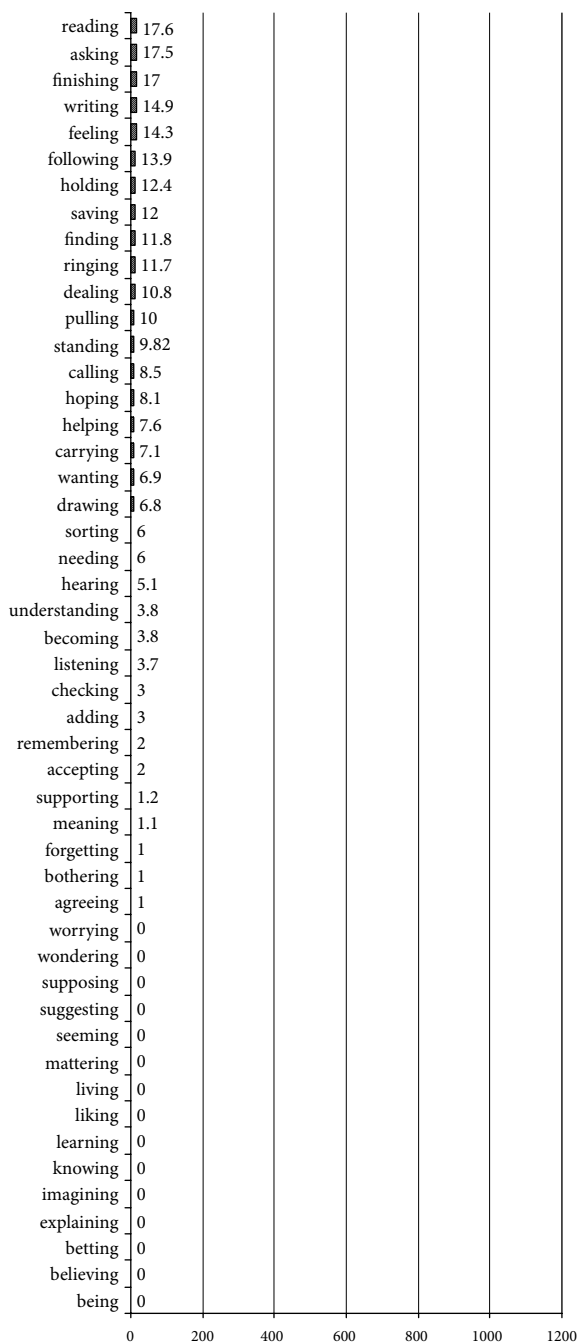


Figure 1. The distribution of progressives with future time reference across verbs in BNC_spoken (extrapolated absolute frequencies)



- (6) Okay erm I'm just trying to think what I'm *doing* later (fuh)
- (7) I won't be there Friday afternoon since I'm *taking* my son to the dentist (kd8)
- (8) Right ho hold on, so, so this is *taking* place within the village? (j1l)

On the basis of these results, H. E. Palmer's abovementioned statement that the progressive with future time reference is possible with only a few verbs, can be refuted. At least in spoken (contemporary British) English, the construction seems possible with a large number of verbs (at least 85% of those investigated here). It is true, however, that a few verbs account for the vast majority of occurrences of the construction. Verifying or refuting statements that the progressive with future time reference preferably or mainly occurs with verbs of motion, verbs of 'coming or leaving', 'coming or going' or 'departure and arrival' is more difficult, as these semantic groups are not usually defined more precisely in grammatical accounts of the construction. Four (reasonably frequent) verbs in particular leap to mind as belonging to the latter three groups, namely GO, COME, LEAVE, and ARRIVE. As *arriving* was not among the 100 verbs investigated, it was analysed separately, and was only found to occur 10 times with future time reference in BNC_spoken.

The statements found in traditional grammars (cf. Section 3.1) thus have to be made more precise in order to be correct. Of the 'verbs of coming or going' or similar, it is precisely the two verbs COME and GO that occur particularly often in the progressive with future time reference. LEAVE is among the (15 or so) verbs that occur often in this function, but this is not necessarily true for other verbs of the *coming/going/leaving/arriving/departing*-group, with *arriving*, for instance, being rather rare. Perhaps more importantly, it is by no means exclusively verbs belonging to this group that occur in this type of construction particularly often. Between *going* and *coming* on the one hand and *leaving* on the other, frequency-wise, we find *doing*, *taking*, *getting*, *having*, *giving*, *happening*, *trying*, *putting*, and *talking*. Whether these can all be considered 'verbs of motion' or 'verbs of movement' naturally depends on how these concepts are defined. A definition like F. R. Palmer's, which includes verbs that merely imply motion (cf. Section 3.1), might include all or most of them (although some, in particular *trying* and *talking* can hardly be subsumed even under such an extended definition). Such a wide definition makes the concept almost useless, however, at least for our purposes, as verbs occurring with much smaller frequencies could equally be considered as somehow implying motion, such as *speaking*, *eating*, *showing*, *writing*, *calling* etc. If motion verbs are narrowly defined, on the other hand — so that the class would be limited to *going*, *coming*,

leaving, travelling, riding, flying, driving, walking and a few other verbs — the statement is not confirmed by our results, as neither are all of these verbs frequent nor does this list cover more than a fraction of the most frequent verbs.

What can be confirmed on the basis of our results is Leech's claim that the progressive with future time reference occurs mainly with verbs involving conscious human agency (2004:63). Most of the occurrences with verbs that do not seem at first sight to involve conscious human agency, such as *forgetting*, nevertheless do so in the context in which they are used (in this case, the meaning is "forgetting on purpose" / "not thinking about something anymore": *So he is now in a position he can either take them to court or he's forgetting it* (j9x)). Contrary to Wekker (1976), we did not find any verbs in the progressive with future time reference that cannot also occur in the progressive with reference to the present. Of the examples cited by Wekker, *SEE* and *HEAR*, the latter occurs only rarely with future time reference in our corpus, and the former occurs with reference to the future mainly in the sense of "meet" (*I'm seeing John Saturday and I'll, I'll mention to him* (jt5)), in which it can also occur with reference to the present (as in *She is probably seeing him as we speak*), although this use is in all likelihood less frequent. As these examples demonstrate, valid statements about which verbs are often used and in particular which verbs cannot be used with the progressive with future time reference, can in many cases only be made with respect to particular senses of verbs.

Comparing our results to those of Mindt (1992) (as far as this is possible), we can observe that those groups of verbs that he finds to be particularly or fairly frequent in progressives with future time reference also occur frequently in this function in our data. Those for which he finds a few instances are fairly frequent or frequent in our data, which is most likely due to the difference in corpus size. Probably for the same reason, we have also found some and in some cases many instances for those verb groups for which Mindt does not list any occurrences, most notably for *HAVE* (as in *we're having a big party tomorrow night* (jp4)). In our corpus, there are also many occurrences of verbs which cannot readily be assigned to any of Mindt's groups of verbs (e.g. *HAPPEN, TRY, PUT*).

Dissatisfying as this may be, our analysis demonstrates that it is not possible to claim that a single (or a few single) reasonably narrowly defined semantic group(s) of verbs is (or are) preferably used with the progressive with future time reference. Rather, it is individual verbs for which this is true. Most of the verbs occurring frequently in this construction unsurprisingly are verbs that are particularly frequent overall. The question therefore arises whether, in order to describe the preferred patterns of a construction, considering absolute

frequencies is the best approach or whether some kind of relative frequency should be considered instead. Interestingly, in none of the previous claims concerning the verbal preferences of the construction any consideration seems to have been given to the question of what it means to state that it is mainly verbs a, b, and c or verb groups x and y that occur in the construction in question: are these verbs most frequent in absolute terms or in relative terms, and if in relative terms, relative to what?

Three possibilities of taking relative frequencies into account offer themselves. First, the frequencies of the individual verbs in the progressive with future time reference can be related to the frequencies of verbs in all other expressions of future time. Such an approach would tell us that, if such-and-such a verb was selected and to be used with reference to future time, which of the several future time expressions available in English would be most likely to be co-selected. Secondly, the frequencies of the individual verbs in the progressive with future time reference can be related to the overall frequencies of the individual verbs. This would tell us how likely the use of a certain verb in a progressive construction with future time reference is compared to the average likelihood of its overall occurrence. Thirdly, the relative frequency of progressives with future time reference as related to progressives in general could be considered, which would tell us how likely it is that a specific verb occurring in the progressive refers to future time. The first of these approaches, while probably being the most enlightening in several respects, is highly impractical if not to say impossible to take (even if we determine a number of expressions as constituting a complete set of expressions capable of expressing future time in English, an inordinate amount of – partly manual – analysis would be necessary). Therefore, approaches number two and three have been adopted here. The occurrences of progressives with and without (pure) future time reference have already been discussed in Römer (2005a:152ff), with a dataset partly overlapping with the present one (cf. Note 2).⁶ It is shown there that the degree to which verbs in the progressive are used with future time reference differs widely from verb to verb. The following verbs were identified as verbs which clearly favour future time reference when occurring in the progressive (in this order): *leaving*, *meeting*, *winning*, *stopping*, *coming*, *sending*, *staying*, and *setting*.⁷ If this result is compared to the absolute frequencies given above, it turns out that many of them are not among those occurring most frequently in the progressive with future time reference in absolute terms: the verbs occupy the positions 2 – *coming*, 12 – *leaving*, 18 – *staying*, 25 – *stopping*, 29 – *sending*, 36 – *meeting*, 37 – *setting*, 45 – *winning* (cf. Figure 1). What is interesting is that the verb *leaving*, which appears to be a stock example in discussions of

the progressive with future time reference (and which is usually assumed to be one of the verbs most often thus used), is the verb *which*, when it occurs in the progressive, displays the highest proportion of future time reference. This might contribute to the impression that *LEAVE* is one of the two or three most frequently used verbs in the construction, although it does not figure among the 5 or 10 most frequent verbs in absolute terms (cf. Figure 1).

The results for approach number two — the frequency of progressives with future time related to the overall occurrence of the individual verbs (i.e. verb lemmas) — are displayed in Figure 2.

A number of striking differences emerge when these results are compared to those of Figure 1, where absolute frequencies are displayed. While the verbs occurring most frequently in the progressive with future time reference in relative terms (i.e. independent of the overall frequencies of the individual verbs) are also generally frequent in absolute terms (the 10 most frequent verbs in relative terms are all among the 25 absolutely most frequent ones), the reverse is by no means true. The 10 verbs that were identified as the most frequent ones in the construction in question in absolute terms occupy the following positions in the list of relative frequency in Figure 2: 2 — *happening*, 3 — *coming*, 6 — *trying*, 7 — *taking*, 18 — *going*, 20 — *giving*, 36 — *putting*, 53 — *getting*, 56 — *doing*, 68 — *having*. Many of the 10 absolutely most frequent verbs are thus often frequent in progressives with future time reference because they are frequent in general. The only four verbs which can be said to be frequent in the construction both in absolute and relative terms are *happening*, *coming*, *trying*, and *taking*. The separate analysis of *arriving* indicates that this verb occurs in the progressive with future time reference 20 times per 1,000 instances of the verb. This means it would occupy a position between 13 and 14 in Figure 2 if it had been included in the overall analysis.

Examples (9) to (16) illustrate the use of four out of the five most frequent verbs according to this approach (*staying*, *happening*, *coming*, *selling*, *sitting*; for the use of *coming* cf. examples (3) and (4) above):

- (9) I think I'll stay here actually six I'm *staying* there (kc2)
- (10) Sunday I'm *staying* the night at house (kbn)
- (11) Not worrying about what's *happening* next week (kgl)
- (12) Well what's *happening* this morning Jean, where we going? (kdn)
- (13) So I'm *selling* these in a minute (kcu)
- (14) So are they *selling* his house? (kb8)
- (15) There's no way I'm *sitting* next to Oh have you got Fat face! (kd2)
- (16) Clint Eastwood nine forty. I'm not *sitting* up to that (kbb)

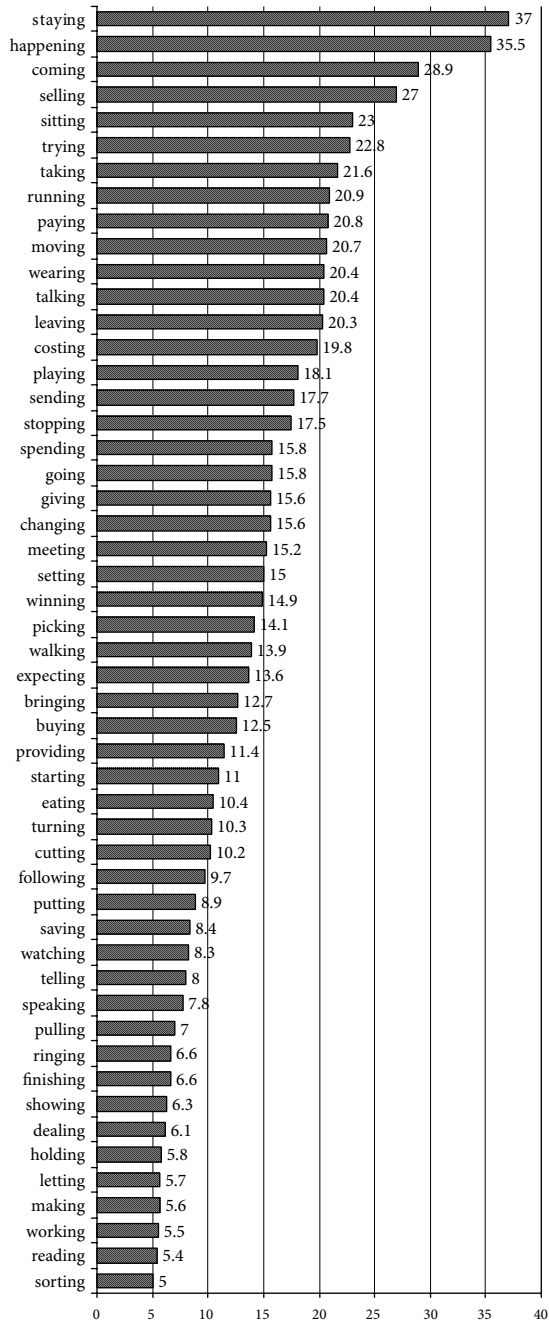
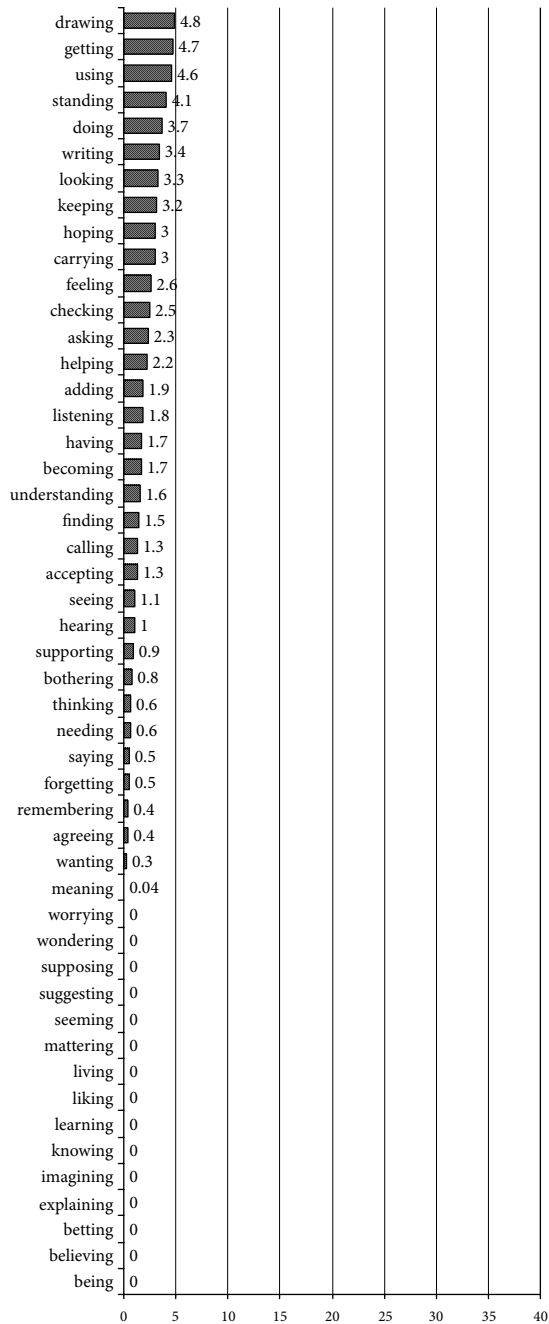


Figure 2. Frequencies of progressives with future time reference in BNC_spoken per 1,000 instances of each verb



If the frequencies of the verbs occurring in the progressive with future time reference are considered relative to their overall frequencies, the claim that it is mainly verbs of motion, or verbs of ‘coming and going’ or the like that occur in this construction is even less valid than when absolute frequencies are considered. While *coming* is highly frequent even in relative terms (the third most frequent, with about 29 instances per 1,000 occurrences of the verb), *going* is much rarer (the 18th position in the frequency list, with about 16 instances per 1,000 occurrences of the verb), and neither *leaving* nor *arriving* are (or would be, in the case of *arriving*) among the 10 most frequent verbs. Besides *coming*, the only motion verbs among the first 10 are *running* and *moving* (and perhaps some of the senses of *taking*). Again, no semantic group can be identified that would encompass the most frequent verbs according to this approach either; it again seems to be individual verbs that are preferably used in the construction.

If all three approaches to pattern frequency adopted here are considered, it turns out that there is only one verb which is among the 10 most frequent verbs in all of them, and this is *coming*. So whereas it is safe to claim that *coming* frequently occurs in the progressive with future time reference, such a claim has to be made more precise with any of the other verbs. This is now possible, on the basis of the results given above, which allow statements such as: *going* and *giving* belong, in absolute terms, to the verbs most frequently used in the progressive with future time reference; if the overall frequencies of the verbs are taken into account, however, they can only be said to occur fairly frequently in this construction.

3.3 Co-occurring adverbials

It has often been claimed that the progressive with future time reference frequently co-occurs with a temporal adverbial. Poutsma, for example, states that in this construction “some adverbial adjunct denoting future is often required” (1904–29 Part II.II:335). Similarly, F. R. Palmer claims that “future uses of both progressive and non-progressive have normally to be marked by an adverbial” (1987:64). It has even been claimed that the presence of an adverbial is obligatory (e.g. Crystal 1966:24f). This latter claim has already been disproved by a number of authors, for example by Visser, who, citing numerous counter-examples, states that it is incorrect to say that “the Expanded Form referring to the future is only possible when there is an indication of future time in the form of an adjunct” (1963–73 Vol. 3.2:1950).

Of the previous empirical studies on the progressive with future time reference, only Aarts (1969) and Wekker (1976) investigate the co-occurrence with adverbials referring to future time. Aarts (who does not differentiate between present progressive and non-progressive forms referring to the future) finds that in 42% of his results, an adverbial is present in the same sentence, and that in 45% there is an adverbial in “removed context, for example in a preceding sentence or in a subordinate clause” (1969:574f). The 13% of instances without an adverbial are almost exclusively instances of the progressive present with *GO*. Wekker, looking only at the occurrences in independent clauses and questions (67 in total), finds that 52% (35) of these are accompanied by a future time adverbial in the same sentence, and an additional 5 (7.5%) by a future time adverbial in “removed context”. He accordingly concludes that “with the progressive future present adverbial modification is optional” (Wekker 1976:104). Comparing his results to those of the simple present with future time reference, he finds that the non-use of an adverbial is much rarer with the present progressive than with the simple present referring to future time, where it is limited to “some rather special circumstances” (ibid.). Mindt (1991) arrives at a similar conclusion when investigating different expressions of future time and specification (though not limited to adverbials but also including, for example, conditional clauses). He finds that the simple present most frequently has an indication of future time in the context. This is followed by the progressive and (in that order) *will* and *going to*, which occur with such a specification less frequently. More precisely, Mindt finds that 86.4% and 79.1% of the progressives with future time reference in his corpora (one corpus of conversation and one of plays) have some kind of specification in the context.

Some authors have also made claims concerning those cases of progressives with future time reference that do not co-occur with a time adverbial. Leech, for example, claims that “[w]ithout an adverbial, a time in the near future rather than remoter future is generally intended” (2004:62); the same claim can be found in Close (1975:260). Leech additionally states that “[t]he future use of the Present Progressive without a time adverbial seems to be chiefly limited to verbs of motion and some other verbs signifying single events” (ibid.). Aarts’ results (cf. above) seem to point in the same direction.

In our analysis, 255 out of the 1,117 instances of the progressive with future time reference occur with a time adverbial either in the same sentence or elsewhere in nearby context, which is defined here as 100 characters to the left and right of the progressive form. (Note that the amount of context considered cannot merely be specified by sentence boundaries in spoken language, as these are often difficult, if not impossible to identify.) This means that 22.8%

of our progressives with future time reference occur with an adverbial. While this figure may not absolutely exactly reflect the state of affairs in the whole of BNC_spoken (considering the sampling methodology we employed), the sample of progressives with future time reference constituting the basis of our analysis should be large and varied enough to be a good approximation. So the ratio of progressives with future time reference with a time adverbial nearby is much lower in our analysis than commonly predicted. In fact, if this percentage is compared to the percentage of how often progressives in general co-occur with a temporal adverbial in BNC_spoken (Römer 2005a:75), it turns out that these are almost identical, with 23.6% of all progressives being thus modified. It is, therefore, by no means the case, as has often been assumed, that the progressive mainly occurs with adverbial specification if it refers to the future, at least not in contemporary spoken British English.

Part of the reason for the low percentage of adverbials co-occurring with progressives with future time reference in our corpus most likely is that the predictions found in the literature usually seem to be based either (exclusively or predominantly) on written language or on intuition-based examples. In Wekker (1976), spoken language is considered but makes up only 1/6th of the corpus, and consists of radio and TV recordings, whereas a large part of the data in BNC_spoken constitutes informal spontaneous conversation. Hence, it does not seem unlikely that in spoken language, and in particular in informal conversation, the degree of adverbial specification in the same sentence or in nearby context is fairly low, and that progressives with future time reference much more often than not occur without such specification. Whether written language differs in this respect only to a small or to a large degree, as is often claimed, will require further investigation. Our results may be an indication, however, that the co-occurrence of the construction with time adverbials in general terms has been overestimated, at least the co-occurrence in the same sentence or in the near vicinity.

How can future time reference of the present progressive be identified, then, if the co-occurrence with adverbials is so rare in spoken language? To answer this question, it is crucial to realise that this is not actually one question but two. For the researcher, the question is whether linguistic indicators can be found that allow the classification of a given instance of the progressive as having future time reference. For the listener, the situation is different, as he or she does not only have the linguistic context to rely on but also knowledge of the extralinguistic context in which the speech event in question takes place (and in many cases some knowledge of the interlocutor and perhaps some additional knowledge shared with the interlocutor concerning the topic being

discussed etc.). If somebody says, for example, *I'm putting on a dress* in the presence of the recipient, he or she does not need any additional linguistic context to decide whether this action refers to the present or the future. If the researcher encounters this form in a corpus, clues in the linguistic context have to be found that shed light on the time reference of this sentence.

If we exclusively consider the linguistic context, several ways of time specification other than through a time adverbial can be identified. Mindt (1991) mentions the presence of a conditional clause or a noun phrase indicating future time. What seems to be more frequent, at least in our data, is adverbials in 'removed context' (i.e. in our case more than 100 characters away from the progressive form), other future markers (such as *will* or *going to*), and negation used in such a way that it is clear that the negated progressive does not refer to the present. Co-occurring future markers can be found, for instance, in the examples quoted in (1) and (9) (*I'm going to the party and I'm gonna take ice; I think I'll stay here actually six I'm staying there*).⁸ Negation as an indicator of future time can be illustrated by utterances such as those in (17) to (19), which are not normally made with reference to the present:

- (17) There's no way I'm *sitting* next to Oh have you got Fat face! (kd2)
- (18) We're not *taking* you home! (kb1)
- (19) I'm not *reading* all that lot (kcu)

Adverbials in removed context are particularly frequent in our corpus (see for instance (20) and (21)). This context can be very far removed indeed, as example (21) illustrates.

- (20) if Clare comes **tomorrow** she's bringing her boyfriend. Is she? Mm. Right on. Long as he wears a kilt. I wanna meet him actually, I'm not *wearing* a kilt. (kce; our emphasis)
- (21) A: [...] and they said would there be any cakes available for **Saturday** and I, I think and I nearly said to erm, well last year you reckon you couldn't sell cakes, but anyway I helped her
 B: Nan, nan, nan
 A: so, an er I said do you sell butterfly ones, and she said I was going to have a go at making them, but she said I can't make them like you,
 B: Aha
 A: I thought well I ain't
 B: It involves, it involves too much work, that's why she can't do them like you.
 A: Christmas fillers you see, so I, she, cos I've, I'm going to make a couple of fruit cakes, I said well I was going to bring in a couple of fruit cakes and

said if you need them not, if not I said I'll take them back I thought no way is Sue gonna eat them, eat cakes like she did last year, or year before

B: Yeah.

A: she only had the bloody cheek to charge us last year for a turkey wrong, but I mean I did get sorted out afterwards.

B: So are you helping with the catering side or are you *running* a cake stall? (kc8; our emphasis)

The concept of 'context', and in particular 'removed context', then, is the aspect which makes statements of the co-occurrence of adverbials with the progressive with future time reference problematic and difficult to compare. While, for instance, both Wekker (1976) and Aarts (1969) use the concept of 'removed context', the difference in their results leads one to suspect that Aarts allows many more intervening words than Wekker (while the former finds about the same amount of adverbials in the same sentence and in removed context, Wekker finds only 1/7th in removed context as compared to those occurring in the same sentence). Mindt (1991) does not specify further his use of the word "context". While it is certainly difficult to define '(removed) context', the issue should certainly be addressed if the concept is used. Our suspicion is that if we consider as '(removed) context' the whole linguistic context of a text (which is of course a concept that is hard to define itself), adverbials will be found to co-occur fairly frequently with progressives with future time reference in spoken language. This does not mean, however, that it is this particular time adverbial which makes the recipient interpret a given progressive as referring to future time (the same applies to adverbials in close vicinity). For the recipient, the extralinguistic context may often suffice to interpret an utterance containing the construction in question correctly. (In example (21) above, for instance, the interlocutors — two housewives planning some kind of social event — in all likelihood do not need the mention of *Saturday* in this specific conversation to establish that the event they are talking about lies in the future.) If this is acknowledged, the occurrence of progressives with future time reference without any explicit marker of future time in the text is also easily explained, and, at least in spoken language, need not be infrequent, even if hard to identify for the researcher. On the other hand, it can also be the case that several indicators of future time are present in the context of a progressive, such as in example (21), where we find not only the adverbial, but (in the fourth turn of speaker A) an instance of the *going to*-future, which also indicates that the topic of discussion is an event taking place in the future.

To find out whether there are differences in the co-occurrence of progressives referring to future time and adverbials with respect to different individual verbs, we investigated the eight most frequently occurring verbs in this function (most frequent in absolute terms in BNC_spoken). The results are given in Table 2.

Table 2. Co-occurrence of eight selected verbs in the progressive referring to future time with temporal adverbials in BNC_spoken

	<i>coming</i>	<i>doing</i>	<i>getting</i>	<i>giving</i>	<i>going</i>	<i>happening</i>	<i>having</i>	<i>taking</i>
With	5	9	0	2	12	9	4	6
adverbial	(15.2%)	(34.6%)	(0%)	(5.1%)	(30%)	(20.9%)	(26.7%)	(15.8%)
Without	28	17	16	37	28	34	11	32
adverbial	(84.8%)	(65.4%)	(100%)	(94.9%)	(70%)	(79.1%)	(73.3%)	(84.2%)

While the numbers are too small for any firm conclusions, they do nevertheless indicate that the co-occurrence with temporal adverbials in the near vicinity is to some degree lexically dependent, with *doing* and *going* co-occurring with an adverbial at least 30% of the time, and *getting* and *giving* only very rarely. The results also indicate that the assumption that occurrences of the progressive with future time reference without a time adverbial are mainly cases with verbs of motion (Leech (2004:62), see above) is probably not accurate, at least not in spoken language and not if “without an adverbial” means without an adverbial nearby: *going* has a particularly high percentage of adverbials, and the low percentage of co-occurring adverbials seems to be true for a great number if not all verbs in our corpus.

Whether the claim that the future use of the progressive without an adverbial mostly refers to the near future (ibid. and Close (1975:260), cf. above) is correct, is more difficult to determine. Again, it is not clear what “without an adverbial” means exactly — no adverbial nearby or no adverbial in the whole text etc.; in addition it is unclear what “near future” refers to. In examples such as *I’m coming* when the coming probably will take place a few seconds or at most minutes after the utterance, one can certainly speak of “near future”. But do “tomorrow” or “next week” still fall into this category? An additional problem with supporting or refuting this claim is that, naturally, the exact time reference of those instances without an adverbial is often difficult to establish. In any case, there are at least some examples in our data without an adverbial nearby which nevertheless do not appear to refer to the next few days, such as *What’s father Christmas bringing you Ben?* (kb6), which was uttered about three weeks before Christmas, or *People don’t realise the lord is coming back and they*

[there]'ll be a day of judgement (gk5), so that the assumption that without an adverbial nearby the progressive usually refers to the near future is in all likelihood not accurate.

If we look at the individual adverbials co-occurring with the progressives with future time reference in our corpus, it turns out that there is one single adverbial that is considerably more common than any other individual adverbial. This adverbial, perhaps somewhat surprisingly, is *now*, which occurs 33 times (out of a total of 255 instances of co-occurring adverbials, i.e. in around 12% of the cases). The exchange in (22) serves to illustrate this.

- (22) A: I'd rather of had Christmas pudding
 B: Go and change it then
 A: No, I'm not *changing* it now. (kb2)

If all progressives regardless of their time reference are considered, *now* occurs in around 14% of the cases (Römer 2005a:77), which means that a progressive co-occurring with the adverbial *now* as likely as not refers to the future.

The only potential lexical verb-adverbial pattern that can be identified in the construction under investigation also consists of this adverb, plus the verb *happening*. The combination is present with three instances, and indicates that *happening now*, or probably *what's happening now?* (in the sense of "what are we going to do now?"), might be a frequent pattern in spoken British English.

3.4 Preferred subjects

While the question of subject distribution has been intensively discussed and studied for the future time expressions *shall* and *will* for decades (e.g. Fries 1925; Krogvig & Johansson 1984; Berglund 2000), it has received practically no attention for the progressive with future time reference, except by Mindt (1991 and 1992). Mindt (1991) compares the ratio of personal and non-personal subjects for the simple present and present progressive with future time reference, for *will*, and for *going to*. His analysis shows that of these four future time expressions, the progressive has the highest ratio of personal subjects (followed by *will*, *going to*, and the simple present, in that order). In his 1992 study (cf. above) he finds that the overwhelming majority of subjects with progressives with future time reference are intentional, with only 11 out of 209 instances occurring with non-intentional subjects (p. 105).

In our dataset, 891 out of 1,117 (79.8%) progressives with future time reference have a personal pronoun as subject, 217 (19.4%) have non-pronoun third person subjects, and in 9 (0.8%) instances the subject is unclear. By far

the majority of these constructions are therefore used with a personal pronoun as subject. For comparison, Römer (2005a:69) found that of her complete progressive sample from BNC_spoken, 76.1% are used with a pronoun, so that pronoun use with progressives with future time reference is slightly higher than pronoun use with progressives in general. If we look at Mindt's results, which are based on conversation and drama, we find that of his set of progressives with future time reference an even higher percentage, namely 91.9%, has a personal pronoun as subject. Berglund, who investigates *will*, *'ll*, *shall*, *going to*, and *gonna* in the spoken component of the BNC Sampler and the London Lund Corpus, finds that around 54% of *will* are used with personal pronouns, around 93% with *'ll*, 91% with *shall*, 67% with *going to*, and 74% with *gonna* (2000:42).⁹ Due to the differences in data bases and sampling, these results are not directly comparable. It does not seem unlikely, however, that the personal pronoun use with progressives lies somewhere between that with *will* on the one hand and with *'ll* and *shall* on the other. Our results also indicate that a high percentage of personal pronouns is not unusual for future time expressions in general and is by no means a special feature of the progressive with future time reference.

In Table 3, the results are given for each personal pronoun individually. The four third person pronouns *they*, and in particular *it*, *he*, and *she* are used with the progressive with future time reference fairly rarely (between 5 and 10 percent), whereas *we*, *you*, and in particular *I* are used fairly frequently (between 20 and 28 percent). To find out whether this state of affairs merely reflects the overall use of these pronouns in spoken language or whether or to what degree this pattern is specific to the construction, the total numbers of the respective pronouns in BNC_spoken were also examined (cf. Table 3). While the percentages of pronouns overall and with the progressive with future time reference do not greatly differ for most pronouns, a considerable difference can be observed for *it* and *we*. *It* occurs with the construction under investigation in only about 5% of the cases, whereas in BNC_spoken, *it* occurs in 19% of the cases when a personal pronoun is used. For *we*, the difference is the other way round: *we* occurs in almost 21% of the cases when a pronoun is used in the progressive with future time reference, but less than 10% overall. More generally, if *she* and *he* are considered together, all third person pronouns occur relatively more frequently in BNC_spoken as a whole, and all first and second person pronouns occur relatively more frequently with progressive constructions referring to the future.

Table 3. Personal pronouns with sample of progressives referring to future time and overall in BNC_spoken

	<i>I</i>	<i>you</i>	<i>he</i>	<i>she</i>	<i>it</i>	<i>we</i>	<i>they</i>
Progressives with future time ref.	248	220	47	54	46	184	92
Percentage	27.8%	24.7%	5.3%	6.1%	5.2%	20.7%	10.3%
BNC_spoken total	355,699	299,360	103,067	57,818	257,441	131,793	142,849
Percentage	26.4%	22.2%	7.6%	4.3%	19.1%	9.8%	10.6%

Compared to the use of pronouns with progressives overall in BNC_spoken (Römer 2005a:68), progressives with future time reference are relatively more often used with *we*, whereas for the other pronouns the percentages are similar (*I* 30.1%, *you* 23.5%, *he* 8.0%, *she* 4.4% — i.e. *he* + *she* 12.4% —, *it* 5.5%, and *they* 12.0%). This means that whereas the strikingly rare use of the progressive with future time reference with the pronoun *it* (in particular as compared to its overall number of occurrence) is a feature of the use of the progressive in general, the frequent use of *we* seems to be a feature that is at least partly special to the progressive with future time reference. A possible reason is that the apparently predominant meaning of the construction of “future arrangement” (cf. e.g. Leech 2004:61ff) often refers to an arrangement between the speaker and some other person. If the results by Berglund are taken as a basis for a rough comparison of the use of *it* and *we* with other future time expressions, it turns out that while *it* is used as a subject considerably more often with most other future time expressions (*will*, *’ll*, *going to*, and *gonna*, but not with *shall*), the use of *we* is probably relatively higher with *shall* and *’ll* than in our data, about the same with *going to* and *gonna*, and lower with *will* (cf. Berglund 2000:42). This again demonstrates that the question of whether a phenomenon can be considered frequent or not depends on what it is considered in relation to.

In order to find out whether the use of subjects is to some degree dependent on the co-occurring verb, the 8 most frequent verbs were again subjected to a close-up analysis. In Table 4, the frequency of co-occurrence of these verbs with personal pronouns and other 3rd person subjects is displayed, revealing considerable differences between the verbs. While the majority of the verbs investigated are predominantly used with personal pronouns (between about 80% and 95%), two verbs do not follow this pattern. *Coming* is used slightly more often than not with other 3rd person subjects, and for *happening*, the pattern is reversed, with a pronoun (*it*) as subject in only about 10% of the instances.

If the patterns that stick out are looked at more closely, it turns out that in the case of *happening*, 35 out of the 38 instances with “other 3rd person

Table 4. Types of subjects co-occurring with eight verbs in the progressive referring to future time in BNC_spoken

Subjects	<i>coming</i>	<i>doing</i>	<i>getting</i>	<i>giving</i>	<i>going</i>	<i>happening</i>	<i>having</i>	<i>taking</i>
Personal pronouns	15 (45.5%)	23 (88.5%)	14 (87.5%)	37 (94.9%)	24 (85%)	5 (11.6%)	12 (80.0%)	30 (78.9%)
Other 3rd ps. subjects	17 (51.5%)	3 (11.5%)	2 (12.5%)	2 (5.1%)	6 (15%)	38 (88.4%)	3 (20.2%)	8 (21.1%)
Subject unclear	1 (4.3%)	-	-	-	-	-	-	-

subjects” are instances of *what* (the other three being *everything*, *some of this lot*, and *another thing*). The pattern already identified above, *what’s happening now*, thus is a ‘subpattern’ of the more general pattern *what’s happening*. In the case of *coming*, there are several different types of non-pronoun 3rd person subjects co-occurring with the construction, of which the type *the* + NP predominates (with 10 instances). An example is *the financial speaker is really coming here in a commercial way* (fue), but a specific lexical pattern cannot be identified.

If the distribution of personal pronouns with the other verbs is examined, it also becomes evident that different verbs follow different patterns. While for some verbs, there is a fairly even distribution of the different pronouns (in our sample in particular for *coming* and *giving*), others seem to have certain preferences. *Going*, for example, co-occurs with first person pronouns in 22 out of the 34 instances; *having* co-occurs with first person pronouns even more frequently, in 10 out of the 12 instances. *Giving* co-occurs predominantly with *I* and *you* (25 out of 34 instances), *doing* with *you* and *we* (16 out of 23 instances). Some examples illustrating typical usage patterns are given in (23) to (26) below.

- (23) I’m *going* to the mountains (fmg)
 (24) we’re *having* a family fun day (jng)
 (25) I’m not *giving* them this one (kd0)
 (26) what are you *doing* this evening (hdj)

Of the instances of *you* + *doing*, all except one occur in the pattern *what are you doing*; the frame *what* + *doing* with future time reference is, however, apparently also sometimes filled with *I* and *we* (in four instances in our sample). From the concordance lines of our dataset, two additional patterns emerge, namely a predominance of first person singular with *leaving* (e.g. *I’m leaving at ten to eight* (kbn)), and with *telling*. In the case of *telling*, most of the instances referring to future time occur in the pattern *I’m not telling* (*you*).

3.5 Forms of TO BE

Two questions concerning the form of TO BE occurring in progressives with future time reference were investigated. First, how often does the construction occur in the past tense as compared to the present, and second, is the construction, or are certain types of the construction, particularly liable to be used with a contracted or a non-contracted form of TO BE?

That past progressive forms can refer to a future as seen from a vantage point in the past is usually mentioned in grammars. Leech, for instance, describes this construction as being “coloured by the notion of ‘intention’ or ‘imminence’ [so that they] do not guarantee that the event foreseen in the past actually did take place” (2004:52). Statements on how common the construction is seem to be non-existent, though. On the basis of our data, we are able to shed light on this question at least for spoken language. In our sample, only a mere 30 out of the 1,117 instances of the progressive occur in the past tense (i.e. with *was* or *were*). An example is *I thought we were walking?* (kb8). Whether this low ratio is due or partly due to the fact that in spoken language, or at least in the types of spoken language represented in BNC_spoken, the present tense is predominant in general was investigated by looking up all the (non-negated) forms of TO BE in the corpus (cf. Table 5).¹⁰ This analysis yielded slightly over 110,000 instances of *was* and *were*, and slightly over 425,000 instances of *is*, *'s*, *am*, *'m*, *are*, and *'re*. This means that the present tense can be assumed to be used about four times as much in BNC_spoken as the past tense. But even taking this fact into account, the ratio of past tense forms of the progressive with future time reference to present tense forms is still extremely low. Further research is required to determine whether this ratio is common to all future time expressions or whether it is specific to the progressive with future time reference, and if so, why.

Table 5 shows the distribution of the different forms of TO BE in our dataset for all non-past tense forms. What is striking is that negated contracted forms are very rare (in particular *aren't* and *isn't*, with a few occurrences of *ain't*) and that instead, negation is commonly expressed by a contracted form plus full *not* (cf. also Section 3.6): *'s not* occurs 16 times, *'re not* 44 times. For comparison, in

Table 5. Present tense forms of TO BE within the sample of progressives with future time reference

Form of TO BE	zero	'm	am	're	are	aren't	's	is	isn't	ain't
Instances	14	231	8	335	162	1	245	78	3	10

the whole of BNC_spoken, *'s not* occurs around 10,000 times and *isn't* around 8,400 times, and *'re not* occurs around 5,000 times and *aren't* about 3,000 times. So whereas the negated forms with contracted BE are somewhat more frequent than the negated forms with contracted *not* in the entire BNC_spoken corpus, this tendency is much more pronounced, in particular with *are + not*, in our dataset. Part of the reason for this seems to be that in the second person, negated forms in the progressive with future time reference often are orders, where the *not* is probably emphasized, as in (27) and (28).

(27) You're *not staying* in here all night. Come on. (kch)

(28) You're *not putting* it on under there (kbb)

If we only look at the non-negated forms, the following picture emerges: 811 are contracted forms and 248 non-contracted forms of TO BE, which means that 76.6% are contracted. In the only other existing study on forms of TO BE with progressives with future time reference we are aware of (Mindt 1992) the results are similar, though not stated explicitly (but deducible from Table 2.1.21 on page 230). Mindt finds 19 non-contracted forms (*is, are, am*) and 55 contracted forms (*'s, 're, 'm*) in his corpus of conversation, and 18 non-contracted and 96 contracted forms in his corpus of plays, which corresponds to a ratio of 74.3% and 84.2%, respectively, of contracted forms. In the whole of BNC_spoken, the ratio of contracted to non-contracted (*is, are, am*) present tense forms is 269,417 to 156,297 (which corresponds to 63.3% of contracted forms). Römer (2005a:66) performs a similar analysis for her sample of progressives in BNC_spoken and gives the following results: contracted 53.9%, non-contracted 45.3%, no form of TO BE 0.8%. These results strongly indicate that progressives with future time reference are particularly liable to occur with a contracted form of TO BE, both compared to the tendency for contraction in spoken language in general and compared to the tendency for contraction in progressives in spoken language. A possible reason for this is that the progressive with future time reference might be particularly often used in informal settings, both as compared to other future time expressions and to the progressive in general.

With respect to the distribution of forms of TO BE across the eight (absolutely) most frequent verbs, the ratios of contracted versus non-contracted forms are displayed in Table 6. The range of the percentage of contracted forms for these verbs is between roughly 60% and 95%, with *doing* and *coming* displaying the highest ratios of non-contracted forms and *getting* the highest ratio of contracted forms. While the high proportion of contracted forms with *getting* might have to do with the informal nature of this verb as such and not

Table 6. Forms of *to be* in the progressive referring to future time co-occurring with eight verbs in BNC_spoken

Forms of TO BE	<i>coming</i>	<i>doing</i>	<i>getting</i>	<i>giving</i>	<i>going</i>	<i>happening</i>	<i>having</i>	<i>taking</i>
Contracted	20 (60.6%)	15 (57.7%)	15 (93.8%)	33 (84.6%)	27 (67.5%)	36 (83.7%)	10 (66.7%)	25 (65.8%)
Non-con- tracted	13 (39.4%)	11 (42.3%)	1 (6.3%)	6 (15.4%)	13 (32.5%)	7 (16.3%)	5 (33.3%)	13 (34.2%)

be special to the construction under investigation, the high ratio of non-contracted forms with *doing* and *coming* cannot be explained with the degree of formality of these words. Instead, the high ratio of full forms with *coming* is a consequence of the low ratio of personal pronouns found with this verb (as contracted forms predominantly follow personal pronouns). The high ratio of full forms with *doing* is also connected to another finding, namely the tendency for the verb in this construction to be used in the pattern *what are you/we doing?* or, more generally, in questions (as in *which are we doing first* (h61), cf. below, Section 3.7), as the full form *are* is usually used (or at least written and thus so transcribed) after interrogative pronouns.

3.6 Negation

In our investigation of potential specific patterns of the progressive with future time reference with respect to negation, only those instances negated by *never*, *not* and its contracted form *n't* were included. Of the total of 1,117 instances of the construction, 147 (13.2%) are thus negated. This ratio of negation is considerably higher than the one Römer (2005a:73) reports for progressives in general (which is 8.6%). Mindt (1992:287) finds an even higher percentage of negation; 20.1% of all instances of this construction are negated. Szmrecsanyi (2003) provides an analysis of negation in the future time expressions *going to*, *gonna*, *will*, *'ll*, and *shall*, also based on BNC_spoken. He only considers negation with *not* and *n't*, but as these form by far the majority of instances (well over 95%) in our analysis as well, his results lend themselves quite well for comparison. Szmrecsanyi (2003:302f) finds 6,311 instances of *will not*, *'ll not* and *won't* as compared to 56,231 instances of *will*, which means that in BNC_spoken, *will* and *'ll* occur in negated forms 11.2% of the time. For *going to* and *gonna*, he finds 2,156 negated forms and 22,454 non-negated forms, which corresponds to 9.6% of negated forms. For *shall*, he reports 21 negated forms and 2,727 non-negated forms. If *shan't* is considered in addition (with

Table 7. Negation with eight verbs in the progressive referring to future time in BNC_spoken

	<i>coming</i>	<i>doing</i>	<i>getting</i>	<i>giving</i>	<i>going</i>	<i>happening</i>	<i>having</i>	<i>taking</i>
Negated	3 (9.1%)	1 (3.8%)	2 (12.5%)	10 (25.6%)	5 (12.5%)	4 (9.3%)	3 (20.0%)	3 (7.9%)
Non-negated	30 (90.9%)	25 (96.2%)	14 (87.5%)	29 (74.4%)	35 (87.5%)	39 (90.7%)	12 (80.0%)	35 (92.1%)

148 instances), *shall* has 6.2% of negated forms in BNC_spoken. If *will* and *shall* are considered together, they occur in negated form 11.0% of the time. These numbers indicate that the tendency of the progressive with future time reference to be used in negated form is somewhat stronger than for other future time markers (an analysis of the simple present with future time reference has, to our knowledge, not been performed yet).

Whether this is a general tendency or one brought about by the preference of negative patterns in the construction by certain verbs, was again examined by first analysing the behaviour of the eight most frequent verbs (cf. Table 7). As can be gleaned from Table 7, the range of the percentages of negated forms for the different verbs varies between about 4 and 25 percent, which points to verb-specific patterns regarding negation. The highest proportion of non-negated forms can be found with *doing*, which in all likelihood is due to the predominance of the pattern *what are you/we doing* identified above. The highest proportions of negated forms are found with *giving* and *having*. Whereas no individual pattern can be identified for negated *having* (which is only present with 3 instances in our dataset), for *giving* the pattern *I'm not giving* (sb. sth.) can be identified as being frequent. Also contributing to the high ratio of negated *giving* are orders of the type *you're not giving x to y*.

In order to get a fuller picture of verb-negation preferences in the progressive with future time reference, we also looked at each verb which occurred in negated form at least three times individually, and calculated its proportion of negated forms. Six verbs turned out to have a particularly high percentage of negated forms in the construction (at least 25% and occurring at least three times). These verbs are *buying*, *letting*, *paying*, *stopping*, *telling*, and *wearing*. Some examples are given in (29) to (34).

- (29) I'm not *buying* any more pencil cases (kcd)
- (30) I'm not *letting* you two in (kc9)
- (31) I'm not *paying* another month's rent (kb7)
- (32) Don't take your coat off cos we're not *stopping* (staying) (kcx)

- (33) I'm not *telling* you (e.g. kdv)
 (34) I'm not *wearing* them tight trousers again (kb3)

None of these verbs occurs relatively often in negative form in progressives in general (cf. Römer 2005a:140f). Negated *telling* even occurs in progressives less than the average ratio of negation (of 8.6%), and negated *buying* and *wearing* lie only slightly above this average.

A third and final approach that we took in order to identify potential patterns with negation and to attempt to explain the comparatively high ratio of negation with the progressive with future time reference was to look at the subjects co-occurring with negated constructions. A tendency already evident in the above examples is that, of the 147 negated constructions, a large proportion — almost 50% (72) — has *I* as their subject. These constructions, in turn, occur predominantly with *'m* (and 3 times with *am* and 6 times with *ain't*), so that there clearly exists a typical pattern: *I'm not* + future progressive. Some examples in addition to those above illustrating this pattern are given in (35) to (39) below.

- (35) I'm not *spending* that amount my poppet. (kcd)
 (36) I'm not *leaving* till I've solved this problem (kbb)
 (37) I'm not *sending* him tomorrow afternoon (kd6)
 (38) I'm not *giving* you no money (kbf)
 (39) I'm not *eating* [the] whole of it (kcv)

The six verbs identified above as occurring in negated form particularly often, with the exception of *stopping*, all occur predominantly with this pattern, but as the additional examples illustrate, the pattern is by no means limited to these verbs.

3.7 Questions

In this final investigation, we looked at potential patterns involving interrogative clauses. In total, 196 (17.5%) of the instances in our dataset occur in questions. The decisive criterion for labelling a BNC_spoken example as a question was not the presence of a question mark but the function of the stretch of text the progressive occurs in. Mindt (1992:296) finds ratios similar to ours in his analysis, namely 18.5% of interrogative clauses in his corpus of conversation, and 21.1% in his corpus of plays. Wekker (1976) finds only 8.9% of interrogative clauses. As his corpus consists mainly of written texts and no informal conversation, this might indicate that the percentage of questions

with the construction under investigation (or in general) is higher in spoken language and/or particularly in informal conversation than in written language or more formal spoken language. If we compare these results to those in Römer (2005a:72), it turns out that at least in the type of spoken language represented in BNC_spoken, progressives with future time reference occur more frequently in questions than progressives in general, which do so in 11.3% of all cases.

The individual results for the eight most frequent verbs are presented in Table 8. The highest percentage of interrogative clauses is found with the verbs *doing* and *happening*, which can be assumed to be due to the patterns already identified above. Fairly high percentages are also found with *taking* and *going*, but this does not seem to be tied to a particular lexical (or lexical-grammatical) pattern. If the ratio of interrogative clauses with these verbs in progressives in general is considered (Römer 2005a), the span is much narrower (with between 7% for *having* and 17.5% for *happening*), but the relative order of frequencies is very similar, i.e. *having*, *coming* and *getting* also occur in interrogatives considerably less frequently than *happening* and *doing*. The patterns emerging from Table 8 therefore do not seem to be limited to the progressive with future time reference, although they are more pronounced in this construction than in progressives in general.

Table 8. Interrogative clauses with eight verbs in the progressive referring to future time in BNC_spoken

Clause type	<i>coming</i>	<i>doing</i>	<i>getting</i>	<i>giving</i>	<i>going</i>	<i>happening</i>	<i>having</i>	<i>taking</i>
Interrogative	2 (6.1%)	12 (46.2%)	1 (6.3%)	5 (12.8%)	6 (15.0%)	9 (20.9%)	0 (0%)	6 (15.8%)
Declarative	31 (93.9%)	14 (53.8%)	15 (93.8%)	34 (87.2%)	34 (85.0%)	34 (79.1%)	15 (100%)	32 (84.2%)

Looking at all verbs individually again, we find that the following verbs occur in questions particularly often in the progressive with future time reference, namely at least 3 times and at least 25% of the time: *eating*, *finishing*, *leaving*, *letting*, *picking*, *playing*, *running*, *staring*, *trying*, *walking*, *watching*, *wearing*, and *winnig*. Some relevant examples are provided in (39) to (49).

- (39) Are you *eating* all that chocolate tonight then? (kbj)
 (40) You *finishing* off your muesli? (kbv)
 (41) Do we know exactly when Anne is *leaving*? (ja9)
 (42) are you *picking* us up (kbl)
 (43) Are you *playing* rugby this afternoon Kevin (kbc)
 (44) Mummy, when are we *starting*? (kbw)

- (45) on the tour well we're sort of trying an experiment Iris is *trying* an experiment (j8b)
 (46) are you *walking* that way (kb7)
 (47) What, to watch the football? Oh! You're *watching* it here? (kcw)
 (48) What you *wearing* tomorrow night? (kbw)
 (49) Who's *winning*? (kdn)

In progressives in general, many of these verbs also occur in questions considerably above average (Römer 2005a:144f). Some of them, however, occur in progressives in questions even below average, so that their high percentage of questions in progressives with future time reference seems to be a special feature of this construction; these are *trying*, *running*, *starting*, and *walking*. No typical individual patterns can be identified, however, for the use of these verbs in questions.

What the above examples seem to indicate, similar to what was observed for negation, is a more general pattern involving the subject *you*. To investigate whether such a pattern does in fact exist, the subjects of the progressives with future time reference in questions were examined. As expected, by far the most frequent subject in those cases is *you*, which occurs in about half of the questions (in 95 out of the 196 tokens). The most frequent pattern with *you* is *Are you ____?*, with *You're _____?* also occurring fairly frequently (and, as examples (40) and (48) above show, occasionally without a form of TO BE, plus a few instances of *You are _____?*). The pattern occurs with a wide variety of verbs, among which (besides *doing*) *leaving*, *staying*, *giving*, *finishing*, *stopping*, *picking (up)*, and *taking* belong to the most frequent ones. A few examples in addition to those already given above are provided in (50) to (55).

- (50) Are you *finishing* that (kbg)
 (51) Are you *giving* us a rendition? (kc3)
 (52) Are you *leaving* or aren't you? (kcx)
 (53) Are you *staying* at your mum's tonight? (kb8)
 (54) Are you *stopping* here or not? (kcx)
 (55) Are you *taking* Steve with you? (kcl)

In addition to this more general pattern and the individual patterns identified for *doing* and *happening* (*what are you/we doing?*, *what's happening?*), there is another individual pattern which occurs repeatedly, namely *who's winning?* (occurring 7 times in our dataset).

4. Conclusion

The great number of individual findings of our investigation warrants a summary. Our analysis has brought to light several general tendencies in the use of the progressive with future time reference, many of which run counter to claims found in descriptions of the construction or make a refinement of these claims necessary. Throughout the paper, our results refer to the use of the construction in spoken British English. The construction was found to be possible with a large number of verbs, rather than being restricted or largely restricted (either in absolute or in relative terms) to one semantic group, such as motion verbs. The claim that it mainly occurs with verbs involving conscious human agency was, however, found to be accurate. In absolute terms, *coming* and *going* are the verbs most frequently used; relative to the overall occurrence of individual verbs, *staying*, *happening*, *coming*, *selling*, and *sitting* are most frequent. Less than a quarter of the examined progressives were found to co-occur with an adverbial nearby (in a span of 200 characters), which is a much lower ratio than commonly assumed. It remains to be investigated to which degree this finding is dependent on the text type investigated (informal conversation). In removed context, adverbials were found to occur fairly often. Claims that occurrences without an adverbial nearby are largely restricted to motion verbs and usually express near future (i.e. the next few days) could not be confirmed. The single most frequent time adverbial near future time progressives is *now*. As for subjects co-occurring with the construction, a high percentage of *we* was found (though probably not higher than in the case of future time expressions in general) and a low percentage of *it* (which is, however, shared with progressives in general). Pronoun subjects were shown to display a high degree of inter-verb variation. The construction was found to occur in the past tense only extremely rarely, but to co-occur often with contracted forms of TO BE. A tendency (which is stronger than for other future time markers) to use negated forms has also been identified, in particular with the verbs *buying*, *letting*, *paying*, *stopping*, *telling*, and *wearing*. Finally, a few more general and several more verb-specific lexical-grammatical patterns emerged from our analysis, the more general ones being: *I'm not _____*, *Are you _____?*, *You're _____?* The most important specific patterns found are: *what's happening?* *what's happening now?*, *what are you/are we doing?*, *I'm not telling you*, *I'm not giving x to y*, *you're not giving x to y*, and *who's winning?*

The analysis has also provided a number of methodological insights. Probably the most important of these is that in accounts of lexical-grammatical patterns it is essential to make clear precisely what notion of frequency is being

used in each case. Does a given analysis discuss absolute or relative pattern frequency, and if the latter, what is the frequency considered in relation to? All instances of a lexical item under consideration in a given corpus? All instances of a grammatical construction under investigation? Or all functionally similar items? In other words, it is not sufficient to merely identify a pattern or tendency of some kind, but necessary to additionally examine whether this pattern is part of a more general one or specific to the construction under investigation. The most comprehensive analysis of patterns naturally is one where several kinds of pattern frequency are considered.

More specific methodological insights include the vagueness inherent in several terms commonly used in descriptions of patterns, in particular the terms 'co-occurring' and 'context' (and, for the construction under investigation, 'near future'). For a valid analysis of patterns, such terms thus have to be precisely defined. Finally, the double nature of what is sometimes referred to as 'specification' (of an indicator for the interpretation of the time reference of a construction) was uncovered, as one has to differentiate between indicators of time reference for the researcher and those for the recipient of an utterance.

Our investigation thus has demonstrated once more that a corpus approach to examining the usage of a grammatical construction (and probably of expressions of any kind) forces the linguist to be much more precise in her or his claims than a non-corpus approach. This precision that the approach demands then leads to improved descriptions of language usage, to which we hope to have contributed. For an even more comprehensive analysis of the progressive with future time reference, it would now be desirable to expand the investigation of the construction to large corpora of different, in particular written, text types and of further varieties of English, and to directly compare the lexical-grammatical patterns of the construction to those of other expressions of future time.

Notes

1. We would like to thank John Sinclair and an anonymous reviewer for their helpful comments on an earlier version of this paper.
2. The dataset of Römer (2005a and 2005b) partly overlaps with the present one, with the following differences: Römer uses data from BNC_spoken as well as from the Bank of English and also considers instances of *going to* + verb progressives. A further difference is that the progressives with *going* and *doing* are represented with a greater number of instances in the present data sample (cf. below).

3. In the case of *doing*, 17 instances with future time reference had been retrieved with the initial random sample of 200 instances of this item, while the entire BNC_spoken (which contains 9,901 occurrences of *doing*) by extrapolation can be expected to contain about 840 such instances. In the case of *going*, 16 instances with future time reference had originally been retrieved, with an expected 1,800 in the whole corpus (which contains 22,544 occurrences of *going*).
4. The groups distinguished by Mindt (1992:57) have the following German labels (with our translation in brackets): *Zustandsverben* (state verbs), *Verben der Lage oder Bewegung* (verbs of position or motion), *Verben des Machens, Tuns* (verbs of doing), GET, HAVE, *Verben der verbalen Kommunikation* (verbs of verbal communication), *Verben der sinnlichen Wahrnehmung* (verbs of sensual perception), BRING/GIVE/TAKE, *Verben der Kognition* (verbs of cognition), *Verben des Entdeckens* (verbs of discovering). Note that some of these groups consist only of one or a few specific verbs while others are not thus restricted.
5. As in some cases the number of occurrences investigated was fairly low, the extrapolated numbers can merely be seen as an approximation of the probable number of occurrences in the whole BNC_spoken. There were, however, only 12 cases of verbs with less than 20 occurrences in the construction in question, so that the extrapolated numbers should in most cases reflect the actual state of affairs quite accurately. An occurrence of zero instances of a verb, on the other hand, cannot, of course, be taken as conclusive evidence that this particular verb never occurs in the progressive with future time reference.
6. These differences should be kept in mind whenever the results are compared. As Römer (2005a) usually gives the results for the BNC_spoken and the Bank of English separately, most results can, however, be compared fairly directly.
7. Römer (2005a) also lists *going*, but if *going to* future constructions are excluded from the counts, this no longer holds true.
8. Of the examples given in Section 3.2, many more contain a future time adverbial than the actual ratio would warrant. This is because in most cases where no adverbial is present, a great amount of context is required to illustrate that a particular instance does in fact refer to the future (see the examples below).
9. Berglund does not give percentages but displays the results in a graph, so that the percentages given here are only approximations deduced from the graphs.
10. The negated forms were disregarded to avoid having to disambiguate the frequent instances of *ain't*. A further caveat is that a few contracted forms might be contracted past tense forms, but a cursory glance at 3 revealed no such instances, and a search for *I's* yielded only a total of four instances which can be assumed to stand for *I was*. As only a rough indication of past tense vs. present tense used was required for the present purposes, this rough approach was deemed sufficient.

References

- Aarts, F. G. A. M. (1969). On the use of the progressive and non-progressive with future reference in present-day English. *English Studies*, 50, 565–579.
- Berglund, Y. (2000). Utilising present-day English corpora: a case study concerning expressions of future. *ICAME Journal*, 24, 25–63.
- Berglund, Y. (1997). Future in present-day English: corpus-based evidence on the rivalry of expressions. *ICAME Journal*, 21, 7–20.
- Close, R. A. (1977). Some observations on the meaning and function of verb phrases having future reference. In W.-D. Bald & R. Ilson (Eds.), *Studies in English Usage: The Resources of a Present-Day English Corpus for Linguistic Analysis* (pp. 125–156). Frankfurt: Lang.
- Close, R. A. (1975). *A Reference Grammar for Students of English*. London: Longman.
- Crystal, D. (1966). Specification and English tenses. *Journal of Linguistics*, 2 (1), 1–34.
- Fries, C. C. (1925). The periphrastic future with *shall* and *will* in Modern English. *Proceedings of the Modern Language Association*, 40 (4), 963–1024.
- Jespersen, O. (1909–49). *Modern English Grammar on Historical Principles*. 7 Vols. London: Allen & Unwin/Copenhagen: Munksgaard.
- Krogvig, I. & Johansson, S. (1984). *Shall and will* in British and American English: A frequency study. *Studia Linguistica*, 38 (1), 70–87.
- Leech, G. (2004). *Meaning and the English Verb* (3rd ed.). London & New York: Longman.
- Mair, C. & Hundt, M. (1995). Why is the progressive becoming more frequent in English? A corpus-based investigation of language change in progress. *Zeitschrift für Anglistik und Amerikanistik*, 2, 111–122.
- Mindt, D. (1992). *Zeitbezug im Englischen: Eine didaktische Grammatik des englischen Futurs*. Tübingen: Narr.
- Mindt, D. (1991). Syntactic evidence for semantic distinctions in English. In K. Aijmer & B. Altenberg (Eds.), *English Corpus Linguistics. Studies in Honour of Jan Svartvik* (pp. 182–196). London & New York: Longman.
- Palmer, F. R. (1987). *The English Verb* (2nd ed.). London & New York: Longman.
- Palmer, F. R. (1965). *A Linguistic Study of the English Verb*. London: Longman.
- Palmer, H. E. (1939). *A Grammar of Spoken English on a Strictly Phonetic Basis* (2nd ed.). Cambridge: Heffer & Sons.
- Poutsma, H. (1904–1929). *A Grammar of Late Modern English*. 2 Vols. Groningen: Noordhoff.
- Römer, U. (2005a). *Progressives, Patterns, Pedagogy. A Corpus-Driven Approach to English Progressive Forms, Functions, Contexts and Didactics*. Amsterdam: John Benjamins.
- Römer, U. (2005b). Shifting foci in language description and instruction: Towards a lexical grammar of progressives. *Arbeiten aus Anglistik und Amerikanistik*, 30 (1–2), 127–142.
- Szmrecsanyi, B. (2003). *Be going to* versus *will/shall*: Does syntax matter? *Journal of English Linguistics*, 31 (4), 295–323.
- Visser, F. T. (1963–73). *An Historical Syntax of the English Language*. 4 Vols. Leiden: Brill.

- Wekker, H. Ch. (1976). *The Expression of Future Time in Contemporary British English: An Investigation into the Syntax and Semantics of Five Verbal Constructions Expressing Futurity*. Amsterdam: B.V. Noord-Hollandsche Uitgeversmaatschappij.
- Zandvoort, R. W. (1972). *A Handbook of English Grammar* (12th ed.). Groningen: Wolters-Noordhoff.

Corpus used

BNC_spoken = the spoken component of the British National Corpus (version 1.0). The British National Corpus is available on CD-ROM from the Oxford University Computing Services, 13 Banbury Road, Oxford OX2 6NN, UK. For further details see <http://www.nat-corp.ox.ac.uk> (accessed 2 January 2007).

Authors' addresses:

Nadja Nesselhauf
English Department
University of Heidelberg
Kettengasse 12
69117 Heidelberg
Germany

nadja.nesselhauf@urz.uni-heidelberg.de

Ute Römer
English Department
University of Hanover
Königsworther Platz 1
30167 Hannover
Germany

ute.roemer@engsem.uni-hannover.de

Appendix

Verb form	BNC_spoken				Verb form	BNC_spoken			
	BNC ran-dom set	Total freq in BNC	Future time ref.	Extra-pol. future time ref.		BNC ran-dom set	Total freq in BNC	Future time ref.	Extra-pol. future time ref.
<i>accepting</i>	94	94	2	2	<i>meaning</i>	200	217	1	1.1
<i>adding</i>	166	166	3	3	<i>meeting</i>	200	2,230	3	33.5
<i>agreeing</i>	62	62	1	1	<i>moving</i>	200	1,047	19	99.5
<i>asking</i>	200	1,164	3	17.5	<i>needing</i>	65	65	6	6
<i>being</i>	200	6,834	-	-	<i>paying</i>	200	1,007	25	125.9
<i>becoming</i>	200	253	3	3.8	<i>picking</i>	200	357	23	41.1
<i>believing</i>	49	49	-	-	<i>playing</i>	200	1,203	12	72.2
<i>betting</i>	34	34	-	-	<i>providing</i>	200	459	9	20.7
<i>bothering</i>	48	48	1	1	<i>pulling</i>	200	251	8	10.0
<i>bringing</i>	200	472	24	56.6	<i>putting</i>	200	1,596	21	167.6
<i>buying</i>	200	645	21	67.7	<i>reading</i>	200	881	4	17.6
<i>calling</i>	200	341	5	8.5	<i>remembering</i>	94	94	2	2
<i>carrying</i>	200	357	4	7.1	<i>ringing</i>	200	234	10	11.7
<i>changing</i>	200	554	19	52.6	<i>running</i>	200	1,409	11	77.5
<i>checking</i>	191	191	3	3	<i>saving</i>	200	240	10	12.0
<i>coming</i>	200	5,557	33	916.9	<i>saying</i>	200	6,072	1	30.4
<i>costing</i>	146	146	25	25	<i>seeing</i>	200	772	9	34.7
<i>cutting</i>	200	437	10	21.9	<i>seeming</i>	4	4	-	-
<i>dealing</i>	200	538	4	10.8	<i>selling</i>	200	618	21	64.9
<i>doing</i>	400	9,901	26	643.6	<i>sending</i>	200	285	36	51.3
<i>drawing</i>	200	340	4	6.8	<i>setting</i>	200	362	18	32.6
<i>eating</i>	200	546	10	27.3	<i>showing</i>	200	548	8	21.9
<i>expecting</i>	200	228	22	25.1	<i>sitting</i>	200	1,324	15	99.3
<i>explaining</i>	102	102	-	-	<i>sorting</i>	110	110	6	6
<i>feeling</i>	200	955	3	14.3	<i>speaking</i>	200	740	6	22.2
<i>finding</i>	200	393	6	11.8	<i>spending</i>	200	538	13	35.0
<i>finishing</i>	132	132	17	17	<i>standing</i>	200	982	2	9.82
<i>following</i>	200	697	4	13.9	<i>starting</i>	200	705	24	84.6
<i>forgetting</i>	85	85	1	1	<i>staying</i>	200	361	49	88.4
<i>getting</i>	200	5,616	16	449.3	<i>stopping</i>	200	257	47	60.4
<i>giving</i>	200	1,182	39	230.5	<i>suggesting</i>	200	276	-	-
<i>going</i>	800	22,544	40	1,127.2	<i>supporting</i>	200	234	1	1.2

<i>happening</i>	200	987	43	212.2	<i>supposing</i>	104	104	-	-
<i>having</i>	200	4,739	15	355.4	<i>taking</i>	200	2,371	38	450.5
<i>hearing</i>	200	257	4	5.1	<i>talking</i>	200	4,656	7	163.0
<i>helping</i>	200	382	4	7.6	<i>telling</i>	200	1,142	17	97.1
<i>holding</i>	200	354	7	12.4	<i>thinking</i>	200	2,152	3	32.3
<i>hoping</i>	200	404	4	8.1	<i>trying</i>	200	3,232	12	193.9
<i>imagining</i>	18	18	-	-	<i>turning</i>	200	404	19	38.4
<i>keeping</i>	200	546	8	21.8	<i>understanding</i>	200	380	2	3.8
<i>knowing</i>	200	360	-	-	<i>using</i>	200	1,628	5	40.7
<i>learning</i>	200	440	-	-	<i>walking</i>	200	836	9	37.6
<i>leaving</i>	200	463	57	132.0	<i>wanting</i>	200	343	4	6.9
<i>letting</i>	200	202	25	25.3	<i>watching</i>	200	701	6	21.0
<i>liking</i>	34	34	-	-	<i>wearing</i>	200	393	16	31.4
<i>listening</i>	200	749	1	3.7	<i>winning</i>	159	159	24	24
<i>living</i>	200	1,271	-	-	<i>wondering</i>	200	316	-	-
<i>looking</i>	200	4,464	3	67.0	<i>working</i>	200	3,549	3	53.2
<i>making</i>	200	2,430	9	109.4	<i>worrying</i>	169	169	-	-
<i>matter</i>	1	1	-	-	<i>writing</i>	200	992	3	14.9