Shifting foci in language description and instruction: Towards a lexical grammar of progressives

Ute Römer

This paper[*] refers to selected results from a large-scale corpus-driven analysis of progressives in spoken British English and in the language of teaching materials used in the EFL classroom. It starts from the observation that existing accounts of the progressive either lack a broader empirical basis or mainly focus on written language data. While the larger study deals with a wide range of context and function features of progressive forms in spoken English and so-called 'school' English, the analytic focus in the present paper is on only one particular co-selection phenomenon: progressives and future time reference. Patterns of progressive use will be isolated and the adequacy of existing linguistic and pedagogical descriptions of this construction will be discussed. Of particular interest in this context is the role that the choice of different lexical items plays in a grammatical construction. It will be explored whether co-selection patterns are strongly lexically determined or valid for progressives in general.

1. Introduction

Studies on the English progressive, i.e. the combination of a form of the verb TO BE and the present participle of a verb (e.g. *'m having*, *were wearing*), are numerous but also rather diverse. Existing accounts provide very different information on the functions which progressive forms mainly express and the contexts in which they tend to occur (cf. e.g. Comrie 1976, Declerck 1991, Landman 1992, Ljung 1980, Recktenwald 1975, Scheffer 1975, Williams 2002). What can be considered problematic in this context is the general lack of a broader empirical basis in most accounts. Theoretical statements are often derived from a limited collection of

example sentences, usually non-attested examples, which the authors selected from earlier studies or which they invented themselves for illustratory purposes.

In the present paper it is assumed that large-scale empirical analyses of language phenomena form a much sounder basis for linguistic statements than predominantly intuitive approaches. I suggest that, prior to theorising, highest possible adequacy ought to be reached at a descriptive level, i.e. in the systematic description of real (or in Brazil's 1995 terms "used") language. A few empirical studies on the progressive in contemporary English have already been made available by corpus researchers (cf. Biber et al. 1999, Legenhausen 1985, Levickij & Romanova 1997, Mindt 2000, Smith 2002, Virtanen 1997). With the exception of Mindt (2000), however, these studies are either based on comparatively small corpora, which might cast doubt on the validity and representativeness of their results, or they only examine rather restricted sets of context and function features of progressives. Besides, and this is a problem that also applies to Mindt's analysis, the general focus is on writing. Spoken English is not covered as extensively as its written counterpart.[1] In order to at least partially close this gap, I carried out a systematic analysis of more than 10,000 progressive form tokens of 100 high-frequency verbs from contemporary spoken British English (cf. Römer Forthcoming; see also section 2 of this article). The present paper refers back to some selected results from this large-scale analysis and complements them with a number of additional empirical findings on the use of progressives in spoken English. One of my aims is to discuss how appropriate existing descriptions of the English progressive actually are in the light of massive corpus evidence.

In connection with questions about the appropriateness of linguistic descriptions it is of major importance to find out whether the progressive forms of a larger number of verbs all behave the same (or at least in similar ways), i.e. whether their preferred contexts and functions are roughly comparable so that it is justified to talk about '**the** progressive' as an underlying grammatical construction which allows for many different verbal realisations or surface forms but carries a meaning (or

---

[1]   In his comprehensive analysis of the English verb system, Mindt (2000) retrieves data from a large meta-corpus merging both spoken and written text collections. Unfortunately, it does not become clear which corpora (or which subsets from which corpora) were used during which stage of his research. From the list of corpora given at the end of the book, however, we can deduce that spoken English only formed a comparatively small part of Mindt's meta-corpus.

meanings) of its own, more or less independent of the respective surface form. The alternative option would be that the contexts and functions of progressives are strongly lexically determined, meaning that they differ significantly from verb to verb. Hence, another question to be dealt with is 'Is there a **purely grammatical** or rather a **lexical-grammatical** progressive?'

Adequate descriptions are not only an important basis of scholarly work in linguistics, they are also essential in language pedagogy. It seems reasonable to claim that in a language teaching context learners ought to be presented with 'true facts' about the language they have chosen to study, i.e. they should be provided with descriptions which do not deviate too much from actual language use. Since the progressive is known to be a constant source of problems for learners, even on an advanced level (cf. e.g. Hahn et al. 2000, Johansson & Stavestrand 1987, Mindt 1997, Williams 2002, Zydatiß 1976), it may be interesting to investigate why this is the case. Is the progressive in itself particularly difficult to handle? Are its meanings so difficult to grasp? Or do problems originate from inadequate (or even faulty) presentations of progressives in teaching materials? To address these questions, I have taken an approach to progressives which compares their use in natural spoken English with their presentation in spoken-type texts included in EFL coursebooks (e.g. dialogues, interviews). Before I turn to the discussion of some representative results of this comparative analysis, I shall now give a brief account of the approach taken and describe what data my investigations are based on.

2. Data and method: A corpus-driven approach to progressives in spoken English and 'school' English

In researching the above-formulated questions related to issues of adequacy in (pedagogical) language description I took a corpus-driven approach. Following scholars in the Birmingham corpus linguistics tradition (cf. Hunston & Francis 2000, Sinclair 1991 and 2004, Tognini-Bonelli 1996 and 2001), I started from a detailed examination of large amounts of corpus data and tried to keep the amount of theory-

reliance to a minimum. Instead of using corpora to merely quantify an existing set of categories taken from a theoretical study or grammar book, categories were formulated in the presence of massive corpus evidence, readjusted from time to time in the light of new data, and quantified subsequently.

Also corresponding to researchers who stress the importance of contextual approaches to the study of language along the lines of J. R. Firth and J. McH. Sinclair, I argue that the careful examination of verbal contexts (or co-texts) of language items in use is absolutely central in linguistic analysis. Embedded in the British contextualist research tradition, my analysis attempts to give a holistic account of progressives. I do not treat the examined structure as an isolated syntactic phenomenon in its own right but as a contextualised form which can come in different shapes and which co-occurs with a number of different lexical-grammatical features. Covered in the larger study (Römer Forthcoming) is a set of some twenty context and function features. Analysed types of co-selection are, for instance, progressives and negation, progressives and adverbials, progressives and time reference, and progressives and central and additional functions.

As mentioned before, the corpus-driven analysis is based on a large collection of more than 10,000 progressives in context. These examples were retrieved from the two largest available corpora of spoken British English, the 10-million word spoken part of the British National Corpus (BNC_spoken) and the 20-million word spoken British subcomponent of the Bank of English (BoE_brspok), and from two small corpora (a little more than 50,000 tokens each) of spoken-type texts from two best-selling German EFL textbook series, *Green Line New* (GLN) and *English G 2000 A* (EG 2000).[2] For the retrieval of corpus examples I used two corpus-analytic software packages: *WordSmith Tools* (with BNC_spoken, GLN, and EG 2000) and *Lookup* (with BoE_brspok). Each of the 10,000+ progressive concordance lines from the four different corpora was imported into an Access database and carefully annotated according to the set of 20+ co-selection features. Figures could then be retrieved from the database semi-automatically for a wide range of feature- and verbform-combinations. The remaining part of this paper will focus on just one of the examined co-selection phenomena: progressives and future time reference.

---

[2]   Details about the composition and compilation of the textbook corpora can be found in Römer (2004) and Römer (Forthcoming).

3. Progressives and future time reference

One of the features I examined in the function analysis of progressives was 'time reference'. For each example retrieved from the selected spoken English and 'school' English corpora, I determined whether the verbform in question referred to an action or event in the past, present, or future, as illustrated in (1) (past time reference), (2) (present time reference), and (3) (future time reference). Let us now just concentrate on one of these types: future time reference.

(1)   I think basically he *was checking* up on me to see what I was doing (BoE_brspok)

(2)   I*'m* just *ringing* up to say congratulations to Birmingham City on making the final at Wembley (BoE_brspok)

(3)   I won't be there Friday afternoon since I*'m taking* my son to the dentist (BNC_spoken)

The fact that we can use progressive tense forms to refer to the future is usually mentioned in every grammar book and in every theoretical study of the progressive. What we find in most grammars are comments on "the progressive with future time reference", either in chapters which deal with progressive aspect (e.g. in Quirk et al. 1985: 210 and Huddleston & Pullum 2002: 171) or in those sections which discuss different means of expressing futurity in English (e.g. in Greenbaum & Quirk 1990: 58 and Greenbaum 1996: 259; see also Quirk et al. 1985: 215). 'Future' is usually regarded as a separate meaning of the progressive, though as a rather minor one. In their 1,800-page grammar, Huddleston and Pullum (2002: 171) only dedicate a short paragraph to "[t]he progressive futurate". Quirk et al. (1985: 210) label the reference to the future as one of the "special uses" of the progressive, and Biber et al. (1999: 470) only briefly mention that present progressives can "describe events that […] are about to take place in the near future" but do not go into further detail or provide any related frequency information. On the whole, this function of progressive forms is to a large extent marginalised and treated as being rare and not particularly noteworthy.

This marginalisation is empirically justified by Smith (2002) in a corpus-based study of progressives in written English. On the basis of two comparatively small corpora of British English writing from the 1960s (LOB, the Lancaster-Oslo/Bergen Corpus) and the 1990s (FLOB, the Freiburg version of LOB), he finds very low shares of future time expression (4.5% for LOB and 3.9% for FLOB; cf. Smith 2002: 324). Basing his research on a considerably larger collection of spoken and (mainly) written corpus data, Mindt (2000: 253) also finds that, with roughly 5% of all progressive instances, future time reference is rather infrequent. Mindt's and Smith's studies hence indicate that, at least with respect to written English, the co-selection of progressives and future time reference is not a particularly common feature.

I shall now discuss what the situation is like in natural spoken language data and in the language of EFL teaching materials. A first focus (in 3.1) will be on average distributions, covering a group of 100 different verb types. Sections 3.2 and 3.3 will deal with two important co-occurrence phenomena within the group of future time progressives and determine the shares of negation and the distribution of different subject types across the progressive concordance datasets. With the aim of throwing light on the question about a purely grammatical vs. a lexical-grammatical progressive, I will then (in 3.4) take a closer look at the behaviour of selected individual verb forms.

3.1 Average distributions in different corpora

To retrieve frequencies of progressives with future time reference in the four selected corpora of spoken English and 'school' English, the annotated concordance examples in the database (10,171 altogether) were filtered for the feature value 'future' in the 'time reference' feature box and for the values 'BNC_spoken', 'BoE_brspok', 'GLN', and 'EG 2000' in the 'corpus' category. Relative frequencies were calculated to make the values more easily comparable. Figure 1 displays the results of this data retrieval and visualises for each of the four corpora the percentages of progressives that express futurity.
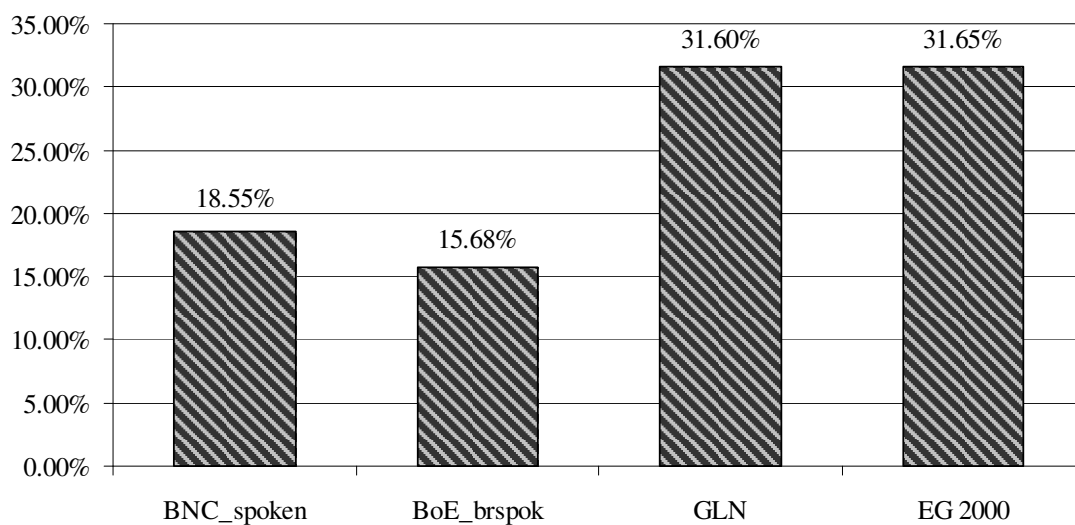
Figure 1: Average shares of progressives with future time reference in different corpora

We see at first glance that progressives with future time reference are significantly more frequent in EFL coursebook language than in natural spoken English.[3] Whereas we determined shares of 15.68% and 18.55% for BoE_brspok and BNC_spoken, the percentages which go back to GLN (31.60%) and EG 2000 data (31.65%) are almost twice as high. Two examples of such progressives, one from 'used' English, one from the coursebook corpus, are given in (4) and (5). We also note that the values for the respective sets of comparable language types, i.e. spoken English on the one hand and 'school' English on the other, are rather similar – a finding which hints at a certain degree of representativeness of the results concerning spoken British English and German EFL textbook English in general.

(4)   *Are* you *eating* all that chocolate tonight then? (BNC_spoken)

(5)   'Let's go!' 'Not now, Dick,' said Grandma. 'The Manaxis monsters *are coming* and I have to save the world. Bye bye. Enjoy your trip!' (EG 2000)

---

[3]   In order to check on the statistical significance of the results, I carried out the chi-square test for each comparison of figures from different sets of data. By comparing actually observed with probabilistically expected frequencies (always raw frequencies, no percentages), chi-square serves to determine whether found differences are significant or not. All chi-square tests were performed using the Georgetown University "Web Chi Square Calculator" (http://www.georgetown.edu/ faculty/ballc/webtools/web_chi.html; consulted 01.01.2005).

Compared to the statements in different reference grammars and to the results of Mindt's and Smith's studies, my rather high percentages of future time progressives may appear surprising. Being faced with attested shares of 18.55 and 15.68 per cent in two large collections of spoken British English data, it seems unjustified to go on marginalising this particular time orientation type and to treat 'future reference' as a special or untypical function of the progressive. My corpus-driven findings hence call for some adjustments of grammatical descriptions, at least of those descriptions that do not exclusively aim at accounting for written language. The considerable mismatch between my findings and the much lower shares determined by Mindt (2000) and Smith (2002) can be explained on the basis of the different types of corpora used in these studies. Smith exclusively refers to written English sources and Mindt's large corpus collection also contains a lot more written than spoken data.[4] Apparently, as my research indicates, the expression of future time by means of progressive forms is much more a feature of speech than of writing. Perhaps we ought to consider this in the compilation of grammars that distinguish between different registers. Biber et al. (1999) in their entirely corpus-based grammar include plenty of valuable information on register variation and on the distribution of linguistic features across written and spoken genres, but, as mentioned above, they just very briefly mention present progressives with future time orientation and, to my knowledge, do not account for register differences in this particular context.

Pedagogical descriptions present a picture which is in complete opposition to what we find in general grammatical accounts. Instead of more or less neglecting progressives with future time reference, as linguistic grammars often do, school grammars and coursebooks clearly overemphasise this function and provide too many examples of progressive forms which refer to actions or events in the future. According to my investigations (cf. Römer Forthcoming), this happens at the expense of underusing progressives with past time reference and, especially of underusing 'indeterminate' progressives which refer to generally valid actions or events, i.e. events expressing present and/or future time reference, as illustrated in (6).

[4]   Spoken data only make up about 4% of the total meta-corpus.

(6)   but people *are drawing* their pensions and they get their lump sums at fifty (BNC_spoken)

Having discussed the general distribution of future progressives in spoken English and 'school' English, I shall now look further into the isolated datasets of progressive form tokens with future time orientation and investigate two co-occurrence phenomena (progressives and negation, and progressives and subjects) in order to trace any possibly existing co-selectional patterns and to see whether similar patterns and shares of these patterns can be found in spoken English and in 'school' English. Since the results for BNC_spoken and BoE_brspok were always comparable in different sorts of co-occurrence feature analysis, they will from now on be summarised and given as 'BNC/BoE'.

3.2 Future time reference and negation

A general analysis of progressives and negation, based on datasets from spoken British English and from my EFL textbook corpus, has rendered the following results (cf. Römer Forthcoming): of all progressives from BNC/BoE, 8.39% are negated, whereas in GLN and EG 2000 we only find shares of negation of 4.56% and 3.80% respectively.

	non-negated	negated
▫ BNC/BoE	86.55%	13.45%
▨ EG 2000	94.40%	5.60%
▢ GLN	93.90%	6.10%
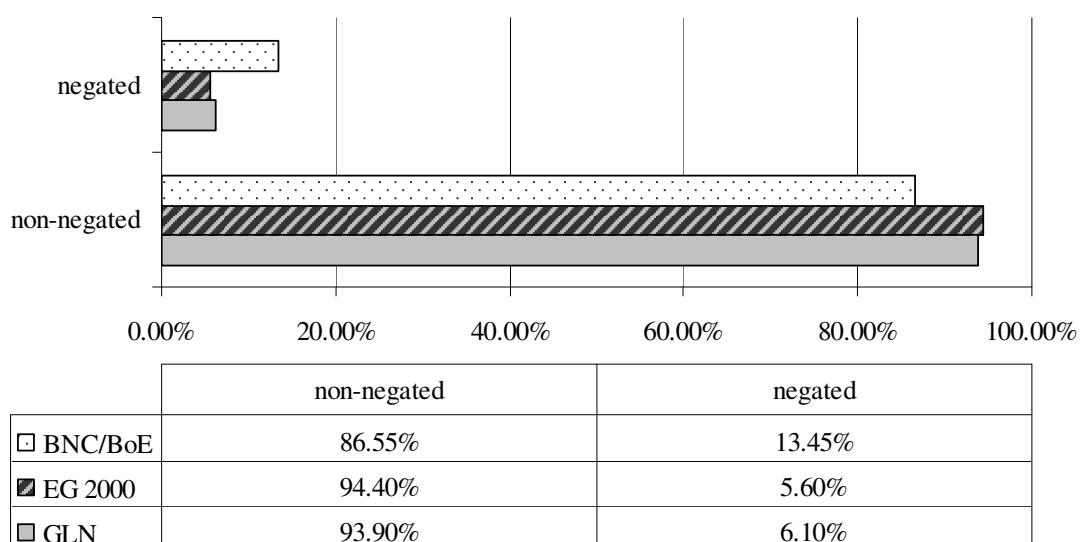
Figure 2: Shares of negated and non-negated progressives with future time reference in different corpora

As figure 2 shows, we get a comparable distribution, but altogether higher percentages, when we just focus on future time progressives. Of all progressive concordance examples from BNC_spoken and BoE_brspok which refer to the future, 13.45% are negated. One of these examples is displayed in (7) below. With only 6.10 and 5.60 per cent, the shares of negation in the GLN and EG 2000 datasets are considerably lower (cf. (8) for one of the very few instances).
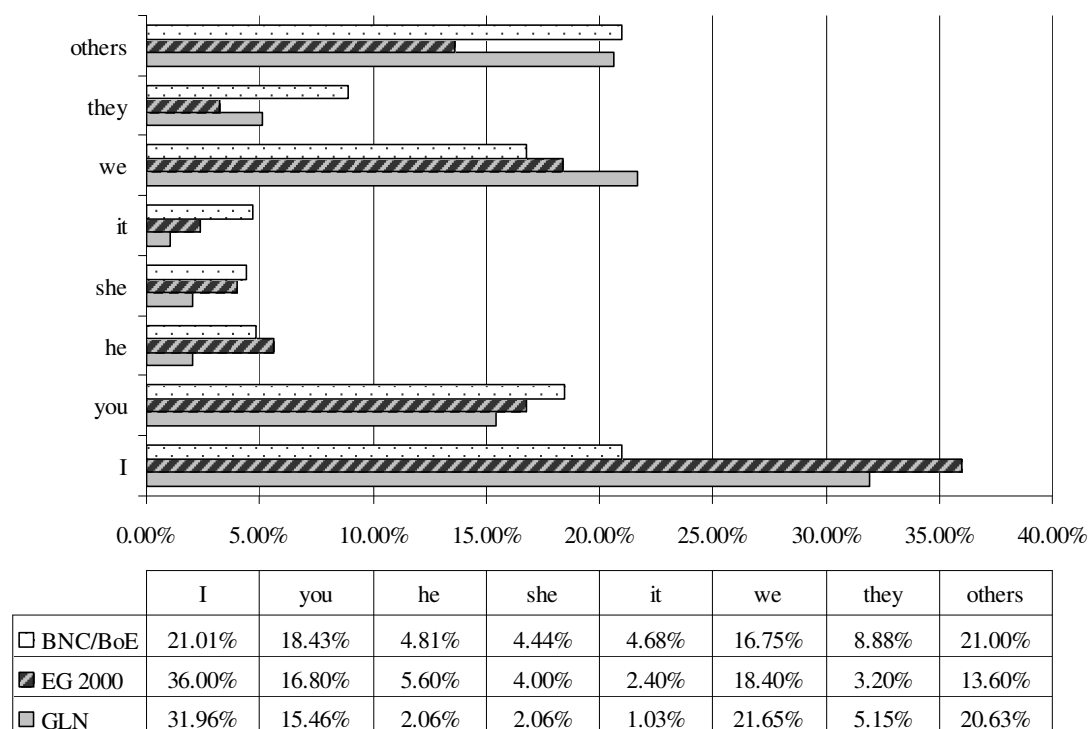
(7)   I was just coming up to Scotland for a holiday. I'*m not coming* here any more. What with the weather and this (BNC_spoken)

(8)   'Kay, I'm going and I'*m not coming* back.' 'Well, leave then! You are boring. I only sit here because our mothers are friends.' (GLN)

We can thus say that the 'future time progressive + negation' pattern is apparently much stronger in natural spoken English than in the language of EFL coursebooks. The findings in this brief section have shown that, although they generally overuse progressives with future time reference, EFL textbooks underrepresent such instances in which the verbform is negated and hence do not give enough prominence to a typical co-selection phenomenon attested in spoken British English.

3.3 Future time reference and different types of subjects

Another context feature under examination in the larger study was the co-selection of progressive verbforms and their subjects. I found that, with almost 80 per cent, personal pronouns form by far the most common group of subjects in BNC/BoE progressives (cf. Römer Forthcoming).

This is also true for the sets of progressives with future time reference. In 79% of the future time progressives from BNC_spoken or BoE_brspok, the subject is either *I*, *you*, *he*, *she*, *it*, *we*, or *they*. The frequency distribution within this group of subjects is illustrated by means of a bar chart in figure 3. We can observe that three of the co-occurrence patterns 'personal pronoun subject + progressive with future time reference' are particularly common in spoken British English: '*I* + progressive' (21.01%), '*you* + progressive' (18.43%), and '*we* + progressive' (16.75%).

	I	you	he	she	it	we	they	others
☐ BNC/BoE	21.01%	18.43%	4.81%	4.44%	4.68%	16.75%	8.88%	21.00%
▨ EG 2000	36.00%	16.80%	5.60%	4.00%	2.40%	18.40%	3.20%	13.60%
☐ GLN	31.96%	15.46%	2.06%	2.06%	1.03%	21.65%	5.15%	20.63%

Figure 3: Shares of subjects of progressives with future time reference in different corpora

A comparison of the BNC/BoE distribution of subjects with the same distribution in the two EFL coursebook datasets (GLN and EG 2000), also displayed in figure 3, highlights some significant frequency differences. With 31.96% and 36.00% the subject *I* is much more often used in the textbook progressives than in natural speech data (see (9) and (10) for examples). Also overrepresented in the teaching materials is the personal pronoun subject *we* (cf. example (11)). On the other hand, compared to real-life spoken data, we find an underuse of *you*, *they*, *it*, and *she* in subject position of future time progressives in GLN and EG 2000. Instances like the ones given in (12) and (13) with combinations of *you* or *we* and a progressive with future time orientation should perhaps replace some of the above-mentioned '*I*-progressives'. It is worth noting that, with respect to some subject types, GLN and EG 2000 deviate quite a bit from each other. The general tendencies, however, are the same.

(9)   "Well?" Lara asked. "I can't. *I'm playing* in a football match in the morning. The team needs me. It's an important game," Nathan said. (GLN)

(10)   I've never been to New York City but I'm going next week. *I'm staying* with my pen-friend in Manhattan. (EG 2000)

(11)   'My parents say you can eat with us if you want. *We're having* a Puerto Rican meal and you'd be welcome.' (EG 2000)

(12)   Erm what time *are you coming* out to play football? Oh Ooh. about half past three. (BNC_spoken)

(13)   Yeah, *they're moving* out at the end of this month. (BNC_spoken)

The observed discrepancies between natural spoken English and 'school' English could be reduced to a considerable extent, just by making different and more corpus-driven choices in the selection of examples for inclusion in the coursebooks. Crucial in this context is the fact that the examined textbooks do not contain 'used' language examples but mainly constructed sentences, invented by the materials writers for the purpose of illustrating a particular language phenomenon. Like Sinclair (2004: 41) I do not have a lot of "confidence in the ability of a human being to invent sentences which display the same patterns of meaning that are to be found in naturally occurring sentences", and would therefore like to make a call for more authentic teaching materials which incorporate language instances that have actually been used for real communicative purposes.[5]

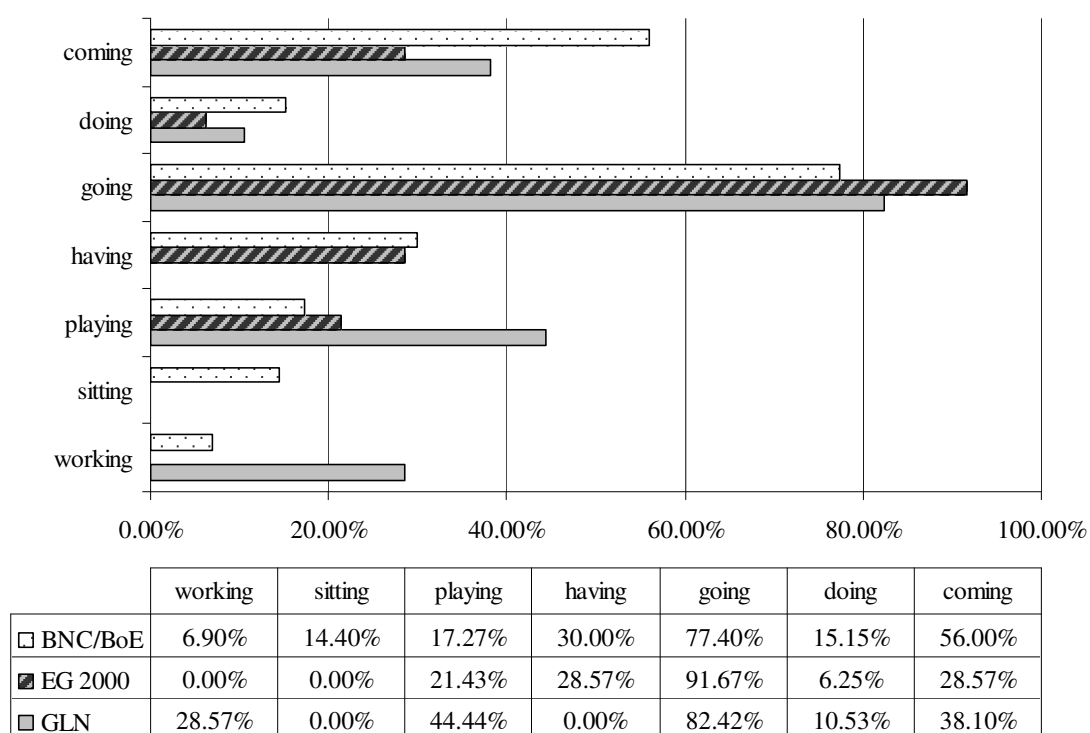### 3.4 Verb-specific distributions

In the analysis of progressives with future time reference I so far only dealt with the distribution of the progressive in general and treated the collected tokens of different progressive verb types as a uniform group. What I presented were average values and distributions of co-selection phenomena that were assumed to be generally valid for progressives in spoken British English and in so-called 'school' English.

In the introduction to this paper, however, I indicated that we cannot really be sure how many verbs actually show a prototypical behaviour, and that individual verbs may exhibit certain co-selectional preferences which deviate from the average progressive values. I formulated the question 'Is there a purely grammatical or rather a lexical-grammatical progressive?' To tackle this question, I will now shift my

---

[5]   See also my discussions of the controversial issue of "authenticity in language teaching" in Römer 2004 and Römer forthcoming.

research perspective from general to individual. Instead of further examining progressive forms of the group of 100 verbs, I shall look at progressive forms of particular verbs and determine some verb-specific shares of future time reference. This analysis was carried out for each of the 100 types in the database. In order to keep things to a manageable size, though, I will now only comment on a small number of verbs. To still make sure that we are dealing with a sufficiently large number of tokens, I selected seven particularly frequent items – frequent not only in spoken English, but also in the textbook corpora.

	working	sitting	playing	having	going	doing	coming
□ BNC/BoE	6.90%	14.40%	17.27%	30.00%	77.40%	15.15%	56.00%
▨ EG 2000	0.00%	0.00%	21.43%	28.57%	91.67%	6.25%	28.57%
▢ GLN	28.57%	0.00%	44.44%	0.00%	82.42%	10.53%	38.10%

Figure 4: Verb-individual shares of progressives with future time reference in different corpora

Figure 4 visualises the shares of future time reference for the verbforms *coming*, *doing*, *going*, *having*, *playing*, *sitting*, and *working*. Three different bar types represent the datasets from three different sources: BNC/BoE (i.e. BNC_spoken and BoE_brspok data taken together), GLN, and EG 2000. We can see that, for most of the verbs, the displayed bar charts deviate considerably in length. *Coming*, for instance, shows a high share of 56.00% of future time reference progressives in

spoken English (an example is given in (14)) but of only 38.10% in the EG 2000 coursebook volumes. The percentage for GLN (28.57%) is even lower. Here we only find very few examples of the type given in (15). With *playing* we observe an opposite tendency. While only 17.27% of all BNC/BoE progressives express futurity (like the one in example (16)), the shares are considerably higher in the coursebooks (21.43% and 44.44%), especially in GLN (see (17) for a typical example). It is also rather striking that in GLN and/or EG 2000 some of the progressives (e.g. of *having* and *sitting*) are never used to refer to a future event or action, although they frequently do so in the natural speech corpora.

(14)   I'll tell her and she can give you a buzz and tell you tell you whether she*'s coming* up on Saturday or Sunday can't she (BoE_brspok)

(15)   Alex: Great that you*'re coming* to visit us for the summer, Daniel! I'm sure you'll like it here in Issaquah. (GLN)

(16)   the football match is next week when England *are playing* (BoE_brspok)

(17)   SAMUEL: Niels is cooking Sunday lunch this weekend. Nadja *is playing* volleyball, and I'm going to the cinema (EG 2000)

Distributional differences, however, are not only observable across corpora but also, and even more significantly, across verbs, i.e. from lexical item to lexical item. Concerning the seven selected forms given in figure 4, the shares of future time reference in spoken English range from 6.90% (with *working*) to 77.40% (with *going*). This finding might cast doubt on the usefulness of average values like those given in the previous sections of this paper. Hence, generalisations about the use of progressives and their typical co-occurrence patterns may be misleading to a certain extent, and predictions on verb-individual behaviour based on average values are rather difficult to make. Coming back to the question about the lexical-grammaticalness of progressives, we can say that our corpus-driven findings suggest that we move away from the purely grammatical progressive and away from a grammar of 'empty' constructions to a grammar of lexical items.

4. Conclusion: Towards more adequate lexical-grammatical descriptions of progressives

The purpose of this article has been twofold: One aim was to argue for a broad empirical basis in the analysis of progressives and their use in spoken English and to make a claim for more adequate linguistic descriptions of these forms. The second aim was to find out how appropriate existing pedagogical descriptions of the progressive actually are and whether or not we may say that inadequate descriptions are one of the origins of the 'learning problem progressive'. With the help of a corpus-driven case study of progressives with future time reference in spoken English and in 'school' English it was possible to pursue both aims.

On the pedagogical side, I found some deviations of the distributions of coursebook progressives from the distributions in natural speech data. In the examined EFL textbook data, I also observed a couple of misrepresentations of typical attested co-selection patterns of progressives. For instance, the percentages of negated future time progressives are significantly lower in GLN and EG 2000 than in BNC/BoE, whereas GLN and EG 2000 clearly overuse future time progressives with the personal pronoun *I* in subject position. A look at the shares of future time reference of individual progressive verbforms brought to light further mismatches between spoken English and 'school' English (cf. section 3.4). We can thus say that, when put to the test against real-life language data, existing pedagogical descriptions of the progressive are not fully adequate and that there is certainly still some room for an improvement of EFL teaching materials. Also, the 'learning problem progressive' may be reduced if we manage to bring the English that is taught more closely into accordance with actual English language use. The results of my analysis suggest, for example, that more attention be paid to typical co-selection patterns of progressives and their preferred context features, and that more emphasis be put on lexical diversity within a grammatical construction.

As I hope to have shown, however, the findings of such a corpus-driven approach to progressive forms do not only have important implications for pedagogical descriptions, but for grammatical descriptions in general. In 3.1 I

pointed out that, given their relatively high frequencies of occurrence in spoken English, the marginalisation of progressives with future time reference in most grammatical accounts of the progressive does not appear justified. Evidently, new empirical findings based on new types (and larger amounts) of data necessitate certain adjustments of linguistic descriptions.

A case in point is the still prevalent focus in the grammar books on '**the** progressive' as a purely grammatical construction. While most grammarians comment on the issue which verbs can or cannot be used in the progressive, they do not comment on the fact that progressive forms of different verbs behave extremely differently concerning their preferred selection of contexts and functions. Whoever consults such a grammar may hence get the impression that all verbs which can take a progressive form are used in similar ways, express roughly the same meaning, and appear in comparable lexical contexts – which, in fact, is not true. I found that, within the progressive as a grammatical structure, individual lexical items show preferences for different formal and functional realisations of this structure. This, I argue, should be accounted for in linguistic descriptions.

What I would therefore propose is a general shift in focus from grammar to lexis. I believe that, if we are aiming at more adequate descriptions of language phenomena, we need a truly lexical grammar, not a grammar in the sense of an empty slot-and-filler model. The corpus-driven study reported on in this paper has demonstrated that 'the progressive' is not just a construction with empty slots that can be filled equally well with whatever item we like. My empirical findings confirm "that any grammatical structure restricts the lexis that occurs in it" (Tognini-Bonelli 2001: 33), and also that the selection of a particular lexical item restricts the structure in which it occurs and the function it may express. In any case, the choices a speaker (or writer) makes matter a lot, and this is probably worth reflecting in linguistic and pedagogical descriptions of the language.

References

Biber, Douglas/Stig Johansson/Geoffrey Leech/Susan Conrad/Edward Finegan (1999). *Longman Grammar of Spoken and Written English*. London: Longman.
Brazil, David. (1995). *A Grammar of Speech*. Oxford: Oxford University Press.

Comrie, Bernard (1976). *Aspect. An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.

Declerck, Renaat (1991). *Tense in English. Its Structure and Use in Discourse*. London: Routledge.

Greenbaum, Sidney/Randolph Quirk (1990). *A Student's Grammar of the English Language*. London: Longman.

Greenbaum, Sidney (1996). *The Oxford English Grammar*. Oxford: Oxford University Press.

Hahn, Angela/Sabine Reich/Josef Schmied (2000). "Aspect in the Chemnitz Internet Grammar". In: Christian Mair/Marianne Hundt (eds.). *Corpus Linguistics and Linguistic Theory: Papers from the Twentieth International Conference on English Language Research on Computerized Corpora (ICAME 20), Freiburg im Breisgau 1999*. Amsterdam: Rodopi. 131-139.

Huddleston, Rodney/Geoffrey K. Pullum (eds.). (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hunston, Susan/Gill Francis (2000). *Pattern Grammar. A Corpus-driven Approach to the Lexical Grammar of English.* Amsterdam: John Benjamins.

Johansson, Stig/Helga Stavestrand (1987). "Problems in Learning – and Teaching – the Progressive Form". In: Ishrat Lindblad/Magnus Ljung (eds.). *Proceedings from the Third Nordic Conference for English Studies, Hässelby, Sept 25-27, 1986, vol. 1*. Stockholm: Almquist & Wiksell. 139-148.

Landman, Fred (1992). "The Progressive". *Natural Language Semantics* 1. 1-32.

Legenhausen, Lienhard (1985). "Die SF/EF-Distribution in englischen Rundfunkreportagen". *Anglistik und Englischunterricht* 27. 117-135.

Levickij, Victor V./T. A. Romanova (1997). "Use of Tenses of Verbs and Adverbs in the English Language: A Statistical Study". *Journal of Quantitative Linguistics* 4/1-3. 135-138.

Ljung, Magnus (1980). *Reflections on the English Progressive*. Gothenburg: Gotab.

Mindt, Dieter (1997). "Complementary Distribution, Gradience and Overlap in Corpora and in ELT: Analysing and Teaching the Progressive". In: Udo Fries/Viviane Müller/Peter Schneider (eds.). *From Ælfric to the New York Times. Studies in English Corpus Linguistics*. Amsterdam: Rodopi. 227-237.

Mindt, Dieter (2000). *An Empirical Grammar of the English Verb System*. Berlin: Cornelsen.

Quirk, Randolph/Sidney Greenbaum/Geoffrey Leech/Jan Svartvik (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

Recktenwald, Robert P. (1975). *The English Progressive: Semantics and History*. Providence, RI: Brown University.

Römer, Ute (2004). "Comparing Real and Ideal Language Learner Input. The Use of an EFL Textbook Corpus in Corpus Linguistics and Language Teaching". In: Guy Aston/Silvia Bernardini/Dominic Stewart (eds.). *Corpora and Language Learners*. Amsterdam: John Benjamins. 151-168.

Römer, Ute (Forthcoming). *Progressives, Patterns, Pedagogy. A Corpus-driven Approach to Progressive Forms, Functions, Contexts and Didactics*. Amsterdam: John Benjamins.

Scheffer, Johannes (1975). *The Progressive in English*. Amsterdam: North-Holland Publishing Company.

Sinclair, John McH. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, John McH. (2004). "Intuition and Annotation – the Discussion Continues". In: Karin Aijmer/Bengt Altenberg (eds.). *Advances in Corpus Linguistics. Papers from the 23rd International Conference on English Language Research on Computerized Corpora (ICAME 23).* Amsterdam: Rodopi. 39-59.

Scott, Michael (1999). *WordSmith Tools* (Version 3.0). Oxford: Oxford University Press.

Smith, Nicholas (2002). "Ever Moving on? The Progressive in Recent British English". In: Pam Peters/Peter Collins/Adam Smith (eds.). *New Frontiers in Corpus Research.* Amsterdam: Rodopi. 317-330.

Tognini-Bonelli, Elena (1996). *Corpus Theory and Practice*. Birmingham: TWC Monographs.

Tognini-Bonelli, Elena (2001). *Corpus Linguistics at Work*. Amsterdam: John Benjamins.

Virtanen, Tuija (1997). "The Progressive in NS and NNS Student Compositions: Evidence from the International Corpus of Learner English". In: Magnus Ljung (ed.). *Corpus-based Studies in English: Papers from the Seventeenth International Conference on English Language Research on Computerized Corpora (ICAME 17).* Amsterdam: Rodopi. 299-309.

Williams, Christopher (2002). *Non-progressive and Progressive Aspect in English*. Fasano: Schena Editore.

Zydatiß, Wolfgang (1976). "Learning Problem Expanded Form: A Performance Analysis". *IRAL* 14. 351-371.

Ute Römer
English Department
University of Hanover