

Part 7

GRAMMAR

LOOKING AT *LOOKING*

Functions and contexts of progressives in spoken English and ‘school’ English

Ute Römer

Source: A. Renouf and A. Kehoe (eds.), *The Changing Face of Corpus Linguistics*, Papers from the 24th International Conference on English Language Research on Computerized Corpora (ICAME), Amsterdam: Rodopi, 2006, pp. 231–242.

Abstract

The paper gives a preliminary account of a more extensive study (Römer, 2005) of the use of progressive forms (e.g. *'re looking*, in “*let’s say we’re looking at carbonates*”) in huge collections of spoken British English and in a small corpus of ‘spoken-type’ texts from German EFL textbooks. The starting point for this data-driven approach is an examination of 100 high-frequency present participle forms. As a concrete example, the results of a detailed analysis of one of these participles, the lexical item *looking*, its different functions and semantic and syntactic contexts, will be presented. The investigation focuses on the differences observed between English as it is used in natural communicative situations and the type of English pupils are confronted with in a foreign language teaching context. It is argued that if linguists, teachers, and textbook writers aim at achieving a greater degree of naturalness or authenticity in English language teaching, corpus evidence must be taken more seriously and further comparative analyses like the one reported on in this article have to be carried out.

1 Introduction

The English progressive is known to present a significant problem for language learners and counts among the trickier features of language (cf. e.g. Hahn *et al.*, 2000; Johansson and Stavestrand, 1987; Nehls, 1988; Williams, 2002). As there is no comparable aspectual category in their native language, German-speaking learners of English tend to be particularly confused and error-prone when it comes to deciding on a simple or a progressive tense

form, and often use progressives in inappropriate contexts. However, this missing counterpart in their mother tongue may not be the only reason why German learners find it difficult to handle the progressive. One might suspect that another problem lies in inadequate or even faulty descriptions of linguistic phenomena in teaching materials. Instead of presenting authentic samples of English and trying to capture what is really there in the language, EFL (English as a Foreign Language) textbooks often contain invented texts that have been constructed with a special teaching purpose in mind or around a certain topic and/or grammatical feature.¹ An initial step towards tackling this problem must hence be a stocktake. Thus the question is: How are progressives used in ‘real’ English and in textbook English?

The present paper contributes to answering this question. It gives a first account of a larger empirical research project (Römer, 2005) on the use of progressive verb forms in spoken British English and in EFL textbooks used in German secondary schools. The main focus of the present analysis is a pedagogic one. A central aim is to find out whether, with respect to a certain lexical-grammatical phenomenon (here English progressives), the language that students are presented with in a teaching context differs significantly from native-speaker language. If there are significant differences between the two types of English (here referred to as ‘real English’ and ‘school English’), the follow-up questions are “How can this problem be handled?” and “In what ways can corpus evidence help to improve language teaching?” The study thus presents a typical example of the indirect use of corpora in language pedagogy.

While the larger study starts from an examination of the 100 most frequent present participle forms in spoken English, the focus in the present article is only on one progressive construction: a form of TO BE + *looking*. For this selected item, a detailed contextual and functional analysis has been carried out, some results of which will be discussed in sections 3 and 4 of this paper. In order to provide a more comprehensive picture, the reader will first be briefly introduced to the data and the analytical method the larger study is based on.

2 Data and method

The data used in the study come from the two largest currently available collections of spoken British English, the 10 million word spoken part of the British National Corpus (henceforth BNC_spoken) and the 20 million word spoken British subsection of the Bank of English (henceforth BoE_brspok),² and from a small corpus of spoken-type texts (e.g. dialogues, interviews) taken from 12 volumes of two best-selling German EFL textbook series geared to secondary school level. This mini-corpus of a little over 100,000 words, the German EFL Textbook Corpus (GEFL TC), was especially compiled for this study. GEFL TC consists of two subcorpora which are

directly related to the two book series *Green Line New* (henceforth GLN) and *English G 2000 A* (henceforth EG2000), and thus allows inter-textbook comparisons.³

Following the work of researchers in the Birmingham corpus linguistic tradition (see for instance Hunston/Francis, 2000, Sinclair, 1991, and particularly Tognini-Bonelli, 1996 and 2001), the analytic approach is corpus-driven, meaning that it works bottom-up, starting from the data, and tries not to let existing theories or traditional categories influence the findings. It is crucial for such an analysis that corpus evidence comes first and that highest possible adequacy is reached at the observational and descriptive levels before the findings are interpreted and theoretical statements are made.⁴ For the present study, this meant that a close analysis of progressive forms in BNC_spoken and BoE_brspok had to be the initial step. Corpus queries were carried out on the so-called ‘-ing form’ of 100 high-frequency English verbs (e.g. DO, HELP, LOOK, WRITE). The resulting concordances of randomly selected corpus examples were then manually filtered in order to exclude from further analyses all instances of the search items which were not part of a progressive construction, e.g. nouns (as in “she’ll play about with some *helping* to get some answers to these fractions”) or adjectives (as in “students need a *helping* hand”).⁵ Included in the analysis were forms that are usually referred to as ‘present progressive’, ‘past progressive’, ‘present perfect progressive’, and ‘past perfect progressive’. Other tense form combinations that are quite rare in spoken English, such as the ‘progressive infinitive’ (e.g. in “just *be looking* at that”) or ‘modal progressives’ (“somebody *might be looking* at this in a hundred years time” [apostrophe missing in original]), were regarded marginal, especially in a pedagogical context, and were therefore excluded. The same analytical steps were applied to the two GEFL TC subcorpora GLN and EG2000. What remained after the concordancing and filtering processes were sets of ‘progressive’ concordance lines from four corpora. These concordance lines, 10,171 altogether, with a contextual span of 200 characters per line, were then imported to an Access database for further semantic and syntactic annotation. Of particular interest in connection with the database annotation (i.e. the interpretation of more than 10,000 progressives in context) were contextual features (e.g. adverbials that usually go with the progressive) and functions (What do speakers typically express with a progressive?). The following sections will centre on the different functions and semantic and syntactic contexts of *looking*.

3 Looking at *looking*: contexts

Table 1 lists the contextual features that were included in the examination of the concordance lines around *looking* (and the remaining 99 progressive forms) and gives examples of possible values each feature can take.

Table 1 Contextual features under analysis.

<i>Feature</i>	<i>Possible feature values</i>
tense form distribution	present progressive (e.g. <i>is looking</i>) past progressive (e.g. <i>were looking</i>) present perfect progressive (e.g. <i>'ve been looking</i>) past perfect progressive (<i>hadl'd been looking</i>)
TO BE contraction	short form (e.g. <i>'m, 've been</i>) long form (e.g. <i>am, have been, was</i>)
subject	e.g. <i>I, she, they, the customer</i>
preposition ⁶	e.g. <i>at, for, up</i>
object	e.g. <i>it, a job, the woman on the bridge</i>
time adverbial	e.g. <i>now, at the moment, last year</i>
place adverbial	e.g. <i>here, there, in the dining room</i>
other adverbial	e.g. <i>actually, desperately, quite</i>
negation	negated (+) / not negated (-)
question	+/-
if-clause	+/-
relative clause	+/-

This list of features emerged from an initial analysis of several hundred examples of progressive forms in context from BNC_spoken. The selection is supposed to capture the most important typicalities in the use of forms like *looking* and to account for frequent co-occurrences, collocations, clusters, or syntactic peculiarities in the concordance lines. An important aim of the analysis is to identify salient patterns around progressives. A comparison of the results of the contextual analysis from the four corpora (BNC_spoken, BoE_brspok, GLN, EG2000) can then show us whether ‘school’ English presents progressives with the same collocates as are found in ‘real’ spoken English. Sections 3.1 and 3.2 will look at two selected features, one in the left-hand and one in the right-hand context of *looking*: TO BE contractions and prepositions.

3.1 Looking and TO BE contractions

Progressive constructions can be realised by a non-contracted ‘long’ form of TO BE plus present participle as in example (1) or by a contracted ‘short’ form of TO BE plus present participle as in (2).

- (1) What we *are looking* to do here is to water down what we have at the moment.
- (2) We *'re looking* for two couples to come to Australia with us and to act as roving reporters.

As short forms are in general “strongly associated with the spoken language” (Biber *et al.* 1999: 1129), high percentages of short forms in the data from

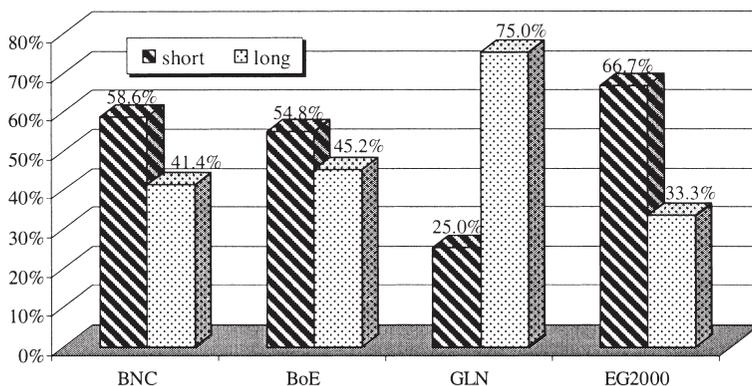


Figure 1 Looking and to be contractions: short forms (e.g. 's looking) vs. long forms (e.g. is looking) in BNC_spoken, BoE_brspok, GLN, and EG2000.

BNC_spoken and BoE_brspok, but also from the two 'spoken-type' based textbook corpora (GLN and EG2000), were expected. Figure 1 shows the proportions of long and short forms of TO BE in the four corpora analysed.

As can be seen in figure 1, the distribution of contracted and non-contracted forms in the two 'real' English corpora (BNC_spoken and BoE_brspok) is quite similar. With 58.6% vs. 41.4% (BNC_spoken) and 54.8% vs. 45.2% (BoE_brspok), both text collections show a dominance of short forms over long forms. This short-form dominance can also be found in the EG2000 data, where the percentages differ even more (66.7 as opposed to 33.3 per cent). The striking thing is that this distribution is reversed in the second GEFL TC subcorpus: in GLN, only 25% of all *looking* progressives are formed with a contracted form of TO BE, while the rest of the progressive constructions contain a long form.

I find it difficult to give plausible reasons for these discrepancies between GLN and the other three corpora. It might be that the textbook authors have chosen this under-representation of short forms for a particular purpose but, at present, it is unclear to me what that purpose may have been. As mentioned above, the 'spoken-type' texts in both GEFL TC subcorpora do not contain "used language" in Brazil's (1995) terms, but constructed examples and invented dialogues. Different distributions of contractions hence may not be due to decisions during transcription (the textbooks are *not* based on any authentic recordings of spoken English), but rather to idiosyncrasies of the authors of GLN.

3.2 Looking and prepositions

The second selected contextual feature relates to the right-hand context of *looking*. Personal experience in assessing university students' writing shows

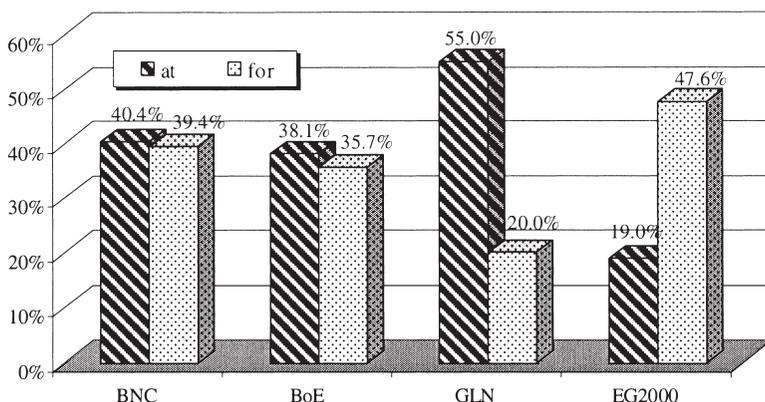


Figure 2 Looking and prepositions: looking at vs. looking for in BNC_spoken, BoE_brspok, GLN, and EG2000.

that learners, even at an advanced level, often find it difficult to decide which preposition goes with which verb. It may therefore be worth taking a closer look at the prepositions that follow the verb form *looking*.

In all corpora, *looking* is most frequently followed by *at* and *for*. Other prepositions like *after*, *back*, or *over* also occur but are very rare. Figure 2 presents the distribution of *looking at* and *looking for* in our four corpora.

The percentage figures mirror the results of *looking* and TO BE contractions presented in section 3.1. Of all *looking* concordance lines from the two 'real' speech corpora, almost equal shares go to *looking at* and *looking for*, with slightly lower numbers for BoE_brspok. But not only do we get comparable figures for BNC_spoken and BoE_brspok, we also again find significant differences between 'real' English and 'school' English, and observe wild discrepancies among the textbook English corpora.

While there are many more occurrences of *looking at* than *looking for* in the texts taken from the GLN series (see example (3)), EG2000 authors apparently favour constructions like the one in (4) with the preposition *for* following *looking*.

- (3) You saw the way they were *looking at* each other! (GLN)
 (4) Yes . . . my head is missing and I'm still *looking for* it. Oh please help me! (EG2000)

These results demonstrate that invented EFL textbook texts do not necessarily present items in their most likely natural lexical contexts. Again, it might be argued that a certain purpose lies behind the textbooks' preference for particular collocations or contexts over others but, again, I find it hard to see what this may be. Besides, if some contexts of occurrence are favoured deliberately and if certain pedagogical principles are at work, how

can it be that the textbook series (GLN and EG2000) differ so much from each other?

4 Looking at *looking*: functions

Let us now turn from context to function. Linguists and grammarians have different opinions about which meanings a progressive construction can carry and either put forward one central and some subsidiary/additional functions (cf. e.g. Biber *et al.*, 1999: 470ff.; Comrie, 1976: 32ff. or Quirk *et al.*, 1985: 197ff.) or suggest a list of several different uses. According to Mindt, for instance, there are nine meanings of the progressive (cf. 1997: 230 and 2000: 248). Frequently-used meaning labels are ‘limited duration’, ‘on-going process’, ‘incompleteness’, and ‘near future’. As stated above, the present analysis is not based on any existing set of progressive functions but starts from the corpora and tries to find the most suitable set of meaning features to account for what is actually there in the data. Table 2 provides an overview of the system of functional features that emerged from a close analysis of several thousand progressive concordance lines. A more detailed account of these features, their discovery procedure, and possible (and probable) feature combinations will be given in the larger study mentioned above (Römer, 2005).

Two of the features listed in table 2 were selected to provide a first account of the functions of progressives with *looking* in ‘school’ English and ‘real’ spoken English. Sections 4.1 and 4.2 deal with the functional features ‘repeatedness’ and ‘general validity’. All central and non-central (or additional) functional features or progressives found in the corpus data are exemplified in the appendix.

4.1 Looking and the central functional feature ‘repeatedness’

Depending on whether they express a single event or describe an action which occurs (or occurred) more than once, each of the 10,171 concordance lines in the database was assigned the feature value ‘non-repeated’ or

Table 2 Functional features under analysis.

<i>Central functional features</i>	<i>Values</i>
time reference	present, past, future, indeterminate
repeatedness	repeated, non-repeated
continuousness	continuous, non-continuous
<i>Non-central functional features</i>	
general validity, politeness/softening, emphasis/attitude, shock/disbelief, gradual change and development, old and new habits, framing	

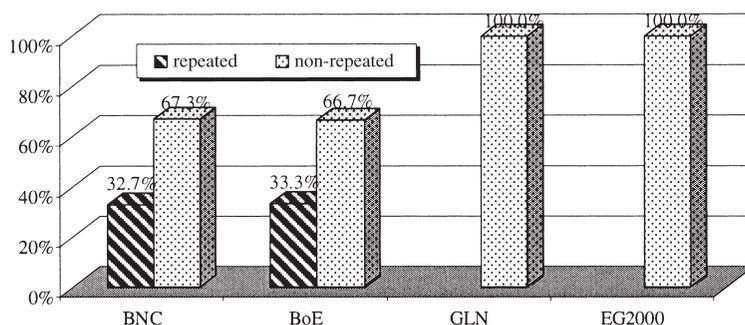


Figure 3 Looking and ‘repeatedness’: repeated vs. non-repeated in BNC_spoken, BoE_brspok, GLN, and EG2000.

‘repeated’ respectively. An example of an utterance referring to a non-repeated action would be

- (5) Well we’re really *looking* for a vegetarian one aren’t we now (BoE_brspok)

whereas repeatedness is expressed in

- (6) Yes. I remember that from when we *were looking* at houses # down there. (BoE_brspok)

Figure 3 graphically displays the proportions of corpus examples with repeated and non-repeated content. Again, the percentages for BNC_spoken and BoE_brspok are very similar and only differ by 0.6%, which hints at a high degree of representativeness of the data samples. Approximately two thirds of the ‘real’ English examples refer to non-repeated single actions, while in one third of the concordance lines, TO BE + *looking* is used to describe a repeated action.

The results from the ‘school’ English data sets are rather striking. Despite the fact that about 33% of the BNC_spoken and BoE_brspok examples express repeatedness, there is not a single instance of this feature in any of the GEFL TC texts, either in GLN or in EG2000.

In the examined EFL textbooks, progressives with *looking* always refer to a single action or event. Here are two typical examples:

- (7) Soldier 1: And the King sent us down to the cellars. What *are* we *looking* for? Monteaige: I don’t know. Gunpowder, maybe? (GLN)
 (8) Hi. I’m *looking* for a Dancing Lesson. I’m trying to work out this number, you see. (EG2000)

Learners may thus be led to think that progressive forms and repeatedness do not go together.

4.2 *Looking and the non-central functional feature 'general validity'*

Similar observations to those for repeatedness could be made for the non-central functional feature 'general validity'. It was found, although this meaning is usually not mentioned in semantic accounts of the progressive, that a number of authentic instances of TO BE + *looking* relate to generally valid facts or to observations that are true in general; 17.3 per cent of the BNC_spoken and 14.3 per cent of the BoE_bspok concordance lines express this feature (see examples (9) and (10) from the BNC).

- (9) If you *'re looking* at triangles you *'re* first thing you *'re* looking at is has it got a right angle?
- (10) I think a golden rule must be, when you *'re looking* at financial er matters is er make sure you know who you *'re* dealing with.

General validity is never expressed in the examples from the GEFL TC subcorpora, which means that another functional feature of progressive forms is withheld from the learners. It might be argued that this is done for didactic reasons and that it may be advisable to focus on one central function of a grammatical construction and leave out less frequent functions in order not to confuse learners. However, as Barlow notes with reference to reflexives, a knowledge of "less frequent patterns is important in moving the language learner from intermediate to more advanced levels of proficiency." (Barlow, 1996: 11). In my opinion, this is also true for the less frequent (non-central) functions of progressives. I do not see any good reason why, at a later stage in their course (maybe in year 5 or 6), learners should not be presented with examples of progressives expressing the 'general validity' function.

5 Conclusion

As the present corpus-driven analysis of progressive forms has demonstrated, there are significant discrepancies between the use of *looking* in 'real' spoken British English and so-called 'school' English as far as the contexts and functions of the verb forms are concerned. In their use of progressives, the analysed EFL textbooks differ significantly from natural language use, and from each other. This latter observation is particularly surprising because, in general, material writers are supposed to (and claim to) follow the same pedagogical principles. Hence, one might expect EFL textbooks to be more comparable with respect to the language they use than they actually are.

What ought to be stressed is that, for the time being, the above findings have to be regarded as preliminary and must be treated with some caution.

We should be careful with overhasty generalisations about functions and contexts of the progressive on the basis of what was observed with respect to *looking* only. Further steps in the more comprehensive study will show whether the analysis of a larger set of data can confirm these initial results. It will then be possible to discuss what kind of corpus-inspired changes in EFL teaching materials may lead to a reduction in the learning problem associated with the ‘progressive’ and to an improvement of English language teaching. In any case, what is worth stressing is the importance of further comparative approaches, especially in the field of applied corpus linguistics. I am convinced that if we want to improve language teaching (materials), we first need to compare the language that is used with the language we teach, in order to discover which language phenomena are still inadequately described and cause problems for learners.

I agree with Bald, who states that “in the field of grammar numerous problems for teachers arise from a *deficient grammatical analysis* on the part of the grammarians” (1988: 3, my emphasis) and would suggest adding *deficient grammatical descriptions* as a major problem source. To arrive at less deficient descriptions, grammarians, linguists, and authors of EFL teaching materials must take authentic linguistic evidence more seriously and should try to present learners with language as it really is – not as they imagine it to be. Of course, textbook authors usually aim at highlighting some important aspects of an item or construction at the expense of others, but they ought to be aware of the risks entailed in the exclusion of other equally important aspects and the “drifting away” from typical “used” language patterns.

Acknowledgements

I am grateful to Wolf-Dietrich Bald for helpful comments on an earlier version of this paper.

Notes

- 1 The issue of invented vs. authentic examples is discussed in Römer (2004). On the issue of faulty descriptions in teaching materials see also Sinclair (2004).
- 2 Instead of only basing the analysis on BNC_spoken data, a second large spoken English corpus was included to see whether the randomly collected data samples from the BNC could be called representative of spoken British English in its entirety.
- 3 More detailed information on the design of the corpus and on its potential for linguistic research can be found in Römer (2004).
- 4 The importance of adequate observations that have to precede any kind of explaining or theorising steps was stressed by Tognini-Bonelli in a panel discussion on “Corpus linguistics, grammar and theory” at the 24th ICAME conference (this volume).
- 5 Unless otherwise specified, all cited examples come from BNC_spoken concordances.

6 Readers might argue that “preposition” is not the correct category for *at*, *for*, *up* etc., or at least not the only one. In a multi-word verbal construction like LOOK *at*, *at* is often labelled “particle” (cf. e.g. Quirk *et al.*, 1985: 1150). As in the present analysis, whether the *looking + at/for/up/etc.* combination behaved as a single unit or not was not taken into account, and in order to avoid further complexity in description, I have decided to stick to the single label “preposition”.

Appendix

Examples from BNC_spoken expressing different progressive functions or functional features (cf. Table 2)

Continuousness

the strange black things that some of you *are holding* in your hands are called riders, and these are end leaves for the storage binder

Non-continuousness

I say real progress has been made but today I *am asking* you to think about the next step

Repeatedness

she does n't eat that much, but what she *is eating* i everything 's sweet. Mm. Chocolates and Is she eating them? Pardon?

Non-repeatedness

Oh I see. Pork's very nice. What's Geoff *eating*? Sausage roll. Oh Geoff, you 've only just had your tea!

General validity

Ca n't help it I I 'm such a give away when things like that *are happening*.

Politeness or softening

I 'm sorry I 'm not clear as to whether you *are suggesting* that there should be policy upper case criteria and some non policy lower case criteria.

Emphasis or attitude

You never drank this coffee! You're always *buying* things and then wanting Oh! to buy something better!

Shock or disbelief

You're not *suggesting* pregnancy's a disease there are you?

Gradual change and development

Er, it is a very difficult climate, it's *becoming* increasingly difficult, and indeed, it's affecting the work that we do

Old and new habits

How long *were* you *seeing* your boyfriend then before you married him? Er, about six years.

Framing

So anyway yesterday afternoon I *was checking* through it when the phone went again.

References

- Bald, W.-D. (ed.) (1988), *Kernprobleme der englischen Grammatik. Sprachliche Fakten und ihre Vermittlung*. Berlin: Langenscheidt-Longman.
- Barlow, M. (1996), 'Corpora for theory and practice', *IJCL* 1,1: 1–37.
- Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan. (1999), *Longman Grammar of Spoken and Written English*. London: Longman.
- Brazil, D. (1995), *A grammar of speech*. Oxford: Oxford University Press.
- Comrie, B. (1976), *Aspect*. Cambridge: Cambridge University Press.
- Hahn, A., S. Reich and J. Schmied. (2000), 'Aspect in the Chemnitz Internet Grammar', in: C. Mair and M. Hundt (eds.), *Corpus Linguistics and Linguistic Theory: Papers from the 20th International Conference on English Language Research on Computerized Corpora (ICAME 20)*. Amsterdam: Rodopi. 131–139.
- Hunston, S. and G. Francis. (2000), *Pattern grammar. A corpus-driven approach to the lexical grammar of English*. Amsterdam: Benjamins.
- Johansson, S. and H. Stavestrand. (1987), 'Problems in learning – and teaching – the progressive form', in: I. Lindblad and M. Ljung (eds.), *Proceedings from the third Nordic conference for English studies*. Stockholm: Almqvist & Wiksell. 139–148.
- Mindt, D. (1997), 'Complementary distribution, gradience and overlap in corpora and in ELT: Analysing and teaching the progressive', in: U. Fries, V. Müller and P. Schneider (eds.), *From Ælfric to the New York Times. Studies in English corpus linguistics*. Amsterdam: Rodopi. 227–237.
- Mindt, D. (2000), *An empirical grammar of the English verb system*. Berlin: Cornelsen.
- Nehls, D. (1988), 'Der englische Verbalaspekt', in: W.-D. Bald (ed.), *Kernprobleme der englischen Grammatik. Sprachliche Fakten und ihre Vermittlung*. Berlin: Langenscheidt-Longman. 192–208.
- Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik. (1985), *A comprehensive grammar of the English language*. London: Longman.
- Römer, U. (2004), 'Comparing real and ideal language learner input: the use of an EFL textbook corpus in corpus linguistics and language teaching', in: G. Aston, S. Bernardini and D. Stewart (eds.), *Corpora and language learners*. Amsterdam: Benjamins. 151–168.
- Römer, U. (2005), *Progressives, patterns, pedagogy: A corpus-driven approach to progressive forms, functions, contexts and didactics*. Amsterdam: Benjamins.
- Sinclair, J. (1991), *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2004), 'New evidence, new priorities, new attitudes', in: J. Sinclair (ed.), *How to use corpora in language teaching*. Amsterdam: Benjamins. 271–299.
- Tognini-Bonelli, E. (1996), *Corpus theory and practice*. Birmingham: TWC Monographs.
- Tognini-Bonelli, E. (2001), *Corpus linguistics at work*. Amsterdam: Benjamins.
- Williams, C. (2002), *Non-progressive and progressive aspect in English*. Fasano: Schena Editore.