Channel and constructional meaning: A collostructional case study

Anatol Stefanowitsch and Stefan Th. Gries

Abstract

Research on grammatical structure, including construction-based research, is sometimes criticized for not paying sufficient attention to variational dimensions such as register, channel, etc. Our recent research on constructions, which is based on a set of quantitative corpus methods that we refer to as *collostructional analysis*, is theoretically subject to such criticism, as we have investigated the meaning of a range of constructions without paying attention to these variables at all. In this paper, we address this potential criticism. Using recent methodological extensions of our method, we show how variables like channel can be included in collostructional analysis. On the basis of three case studies, we show that (i) constructions may display channel-specific associations to individual lexical items, (ii) constructions differ with respect to their channel sensitivity, and (iii) the meaning of a given construction does not vary across channels. We argue that the inclusion of channel-specific information is a necessary addition to the specification of a construction's properties even though it does not interact substantially with constructional semantics.

Keywords: collostructional analysis, configural frequency analysis, quantitative corpus linguistics, construction grammar, channel, spoken vs. written language

1. Introduction

Researchers interested in language structure, especially syntacticians, are sometimes criticized by broadly usage-oriented linguists for failing to include in their analyses a range of usage-related aspects such as social meaning, interactional meaning, or – most importantly for the present paper – register (i.e. contextually induced variation) and channel (i.e. variation induced by the choice of spoken vs. written language) (see, for example, Tummers et al. 2006; Grondelaers this volume; Speelman this volume).

This criticism potentially also applies to a strand of research that we have developed in a series of recent publications (Stefanowitsch and Gries 2003, 2005; Gries and Stefanowitsch 2004a, 2004b, in press). In this work, we have outlined a usage-based procedure for investigating the semantics of constructions (in the Construction Grammar sense of the term, cf. Goldberg 1995: 4, cf. also Croft 2001: 18 ff.).

Construction grammar is a cover term for a group of more or less closely related linguistic theories most of which all share two assumptions: first, that grammatical structures are meaningful linguistic signs, and second, that these signs, referred to as *grammatical constructions*, are the basic units of grammar. Grammatical constructions can vary in complexity and schematicity, ranging over the following broad types (cf. the discussion in Croft 2001: 17 f.):

- *simple specific*, i.e. morphemes such as *give* or *-ing*;
- *simple schematic*, i.e. grammatical categories – for example, word classes (NOUN, VERB, etc.) or grammatical relations (SUBJECT, OBJECT, etc.);
- *complex specific*, i.e. multimorphemic words (like *caregiver* or *give up*) or fixed expressions (like *Don't give up the day job* or *He gives twice who gives quickly*);
- *complex schematic*, i.e. partially filled expressions (like SUBJECT be given to NP, as in *Billy is given to hasty decisions*) or fully abstract grammatical structures (like the ditransitive construction SUBJECT + VERB + OBJECT + OBJECT, as in *Billy gave Diane a diamond ring*).

Most of these theories take a broad approach to meaning, taking this notion to cover not just semantics proper (i.e. propositional semantics and frame semantics) but also 'contexts of use' (cf. Goldberg 1995: 229), which, presumably, include channel, register, etc. In any theory attributing meaning to grammatical structure, it is desirable to have empirical discovery procedures that allow the researcher to uncover this meaning.

Our procedure, which is based on a quantitative corpus-linguistic method referred to as *collostructional analysis*, essentially involves identifying the association strength between a given construction and the lexical items occurring in one (or more) particular slot(s) in that construction (i.e. the strength of the preference of lexical items to occur in particular constructional slots); the lexical items are then ranked according to their association strength and grouped into semantic classes (typically on the basis of common-sense criteria arrived at inductively, but cf. Gries and

Stefanowitsch in press for a more objective procedure based on hierarchical cluster analysis). These classes can then guide the researcher in uncovering the meaning or meanings of the construction in question as well as allowing statements about the relative importance or centrality of particular subsenses as compared to others.

On the basis of this method, we have (among other things) uncovered systematic and semantically highly coherent distributional differences between members of 'alternating' pairs of constructions in English such as active/passive (*Billy wrote this poem* vs. *This poem was written by Billy*), ditransitive/*to*-dative (*Billy sent Diane a poem* vs. *Billy sent a poem to Diane*), *s*-genitive/*of*-construction (*the poem's beauty* vs. *the beauty of the poem*), the *will*-future and the *going*-*to*-future (*Diane will marry Billy* vs. *Diane is going to marry Billy*), and the two verb-particle constructions (*Diane's father gave away the bride* vs. *Diane's father gave the bride away*).

Let us give a concrete example to illustrate this method and introduce some crucial terminology. Consider the alternation between the ditransitive and the *to*-dative. Many verbs in English can occur in both of these constructions (for example, *give*, *tell*, *bring*, *teach*, *send*, etc.), a fact which has led a number of researchers to claim that the two constructions are purely formal variants. However, the fact that a large number of verbs *can* occur in both constructions does not warrant the conclusion that all of these verbs actually *do* so randomly. Instead, some or all of these verbs may have significant preferences towards one of the two (we call words that have such a significant preference to one member of a given pair of constructions (*significant*) *distinctive collexemes* of that construction). Take the verb *give*, which is actually one of the most frequently found verbs in both constructions. More precisely, in the British Component of the *International Corpus of English* (ICE-GB), it occurs in the ditransitive 461 times and there are 574 occurrences of this construction with other verbs, and it occurs in the *to*-dative 146 times and there are 1,773 occurrences of this construction with other verbs. In order to test whether this distribution shows a significant association to one of the two constructions, these frequencies are entered into a two-by-two contingency table, which can then be submitted to a distributional statistic such as the chi-square test or the Fisher-Yates exact test (in our work, we use the latter, since, as its name suggests, it is an exact test, cf. Stefanowitsch and Gries 2003 and Gries and Stefanowitsch 2004 for details on the statistical test and extensive justification). Table 1 shows the relevant contingency table (for expository reasons, the table also shows the frequencies expected from chance alone in parentheses; these were, of course, derived by the

Table 1. The distribution of give in the ditransitive and the to-dative

	give	Other Verbs	Row Totals
Ditransitive	461 (213)	574 (822)	1,035
To-dative	146 (394)	1,773 (1,525)	1,919
Column Totals	607	2,347	2,954

standard procedure of multiplying the marginal frequencies for each cell and dividing the results by the table total).

Submitting these frequencies to the Fisher-Yates exact test yields a p-value of 1.84E-120, rounded off to 1.84E-120. This p-value shows that *give* is significantly associated with one of the two constructions; however, it does not in itself tell us with which of the two. In order to determine this, we need to compare the observed frequencies with the expected ones. This comparison shows that *give* occurs in the ditransitive more than twice as frequently as expected, but only occurs about two thirds as frequently as expected in the *to*-dative. Thus, *give* is significantly associated with – is a (*significant*) *distinctive collexeme* of – the ditransitive construction. Since the comparison is only between these two constructions, this automatically entails that *give* is repelled by the *to*-dative. One can now apply the same procedure to all verbs that occur at least once in each of the two constructions in the ICE-GB, and rank the results in descending order of the p-values.

Table 2 shows the significantly distinctive collexemes for each construction.

Table 2. Distinctive collexemes in the ditransitive and the to-dative

Ditransitive		*To*-Dative	
Collexeme	p	Collexeme	p
give (461:146)	1.84E-120	*bring* (7:82)	1.47E-09
tell (128:2)	8.77E-58	*play* (1:37)	1.46E-06
show (49:15)	8.32E-12	*take* (12:63)	2.00E-04
offer (43:15)	9.95E-10	*pass* (2:29)	2.00E-04
cost (20:1)	9.71E-09	*make* (3:23)	6.80E-03
teach (15:1)	1.49E-06	*sell* (1:14)	1.39E-02
wish (9:1)	5.00E-04	*do* (10:40)	1.51E-02
ask (12:4)	1.30E-03	*supply* (1:12)	2.91E-02
promise (7:1)	3.60E-03		
deny (8:3)	1.22E-02		
award (7:3)	2.60E-02		

Clearly, the distinctive collexemes of both constructions encode the notion 'transfer' (either literally or metaphorically). However, the collexemes of the ditransitive all encode a relatively direct transfer of an object from an agent to a recipient in a face-to-face situation; in contrast, the collexemes of the *to*-dative encode a transfer of an object over some distance to some location (note that the verb *play*, whose presence in this list may seem puzzling at first, occurs in these constructions frequently in sports commentary (e.g. *Billy plays the ball to Diane*). This semantic contrast between the two sets of distinctive collexemes reflects a corresponding semantic contrast between the two constructions that has been posited by a number of authors (for example, Goldberg 1995).

In all our previous analyses, we have taken a relatively narrow approach to constructional meaning, restricting ourselves to semantics proper (in the sense defined above) and, occasionally, some general pragmatic aspects such as illocutionary force or discourse structure. We have almost completely ignored differences concerning channel or register. We assumed that these variables would not interact significantly with constructional semantics proper; however, we did note in passing that the specific verbs instantiating the semantic classes identified by our procedure were possibly influenced by asymmetries in the distribution of a given construction across channels (Gries and Stefanowitsch 2004: 128 f., cf. also Stefanowitsch and Gries 2003: 233 f.).

In this paper, we take initial steps toward a systematic inclusion of channel variation into collostructional analysis by extending the method of distinctive collexeme analyses first introduced in Gries and Stefanowitsch (2004a), specifically, the three-dimensional version introduced in Stefanowitsch and Gries (2005). We focus on three questions, namely (i) whether there are channel-specific associations between constructions and individual lexical items (i.e., whether there are verbs that are associated with a given construction in channel X, but not in channel Y), (ii) whether constructions differ with respect to such associations (i.e. whether some constructions display more cases of such channel-specific associations to individual lexical items), and (iii) whether these associations point to channel-specific differences in the meaning of constructions (i.e. whether the individual lexical items associated with a given construction in channel X form a different semantic class than those items associated with the same construction in channel Y). With these tests, we can determine whether constructional semantics is in fact sensitive to channel differences or not.

In order to investigate these questions, we need to operationalize the notion *channel* in a manner suitable to the requirements of collostructional analysis. Collostructional analysis in any of its manifestations requires an exhaustive and thus largely manual retrieval of the construction(s) in question, and thus large amounts of syntactically annotated corpus data are needed. Such data are hard to come by: one of the few corpora annotated in sufficient detail is the British component of the *International Corpus of English* (ICE-GB) already mentioned above, which is a mere one million words in size (cf. Greenbaum 1996 for detailed descriptions of this corpus). The files in the ICE-GB are classified according to a variety of dimensions that would allow us to categorize them into different sub-channels or even registers (or register-like entities) in various ways, but even a moderately sophisticated approach to this task would result in sub-corpora too small for quantitative analysis. Thus, we are forced to settle on a very general operationalization and simply draw the broadest distinction that the ICE-GB allows us to draw, that between spoken and written language. Note that this distinction will – to some extent – correlate with classes of registers that are more likely to be associated with one or the other of these channels, but this correlation will, of course, not be a perfect one.

2. Case Studies

Aims and Methods. In order to investigate the potential influence of channel on constructional semantics, we chose three pairs of constructions from among those analyzed in our previous work that have been claimed to interact with channel or register: active vs. passive voice, the two verb-particle constructions, and two English future-tense constructions; since we have already investigated these constructions in earlier work, we will be in a position to compare the results of these case studies to the earlier analyses, where channel was not included.

With respect to passives there is by now general agreement that they are distributed asymmetrically across channel: while several early studies (Blankenship 1962; Poole and Field 1976) failed to find clear preferences of the passive for either spoken or written language, more recent work has consistently reported that passives are more frequent in writing (cf. Chafe 1982; Brown and Yule 1983; Biber et al. 1999: 938). In addition, Biber et al. (1999: 937) observe that in English "short dynamic *be*-passives are sharply differentiated by register, with conversation and academic prose at opposite poles"; specifically, they report that in conversation stative and dynamic passives are about equally frequent while in academic writing dynamic passives are seven times as frequent as stative ones.

With respect to the two verb-particle constructions, there is little previous work on their distribution across channel or register. In one recent corpus-based study, Gries (2003: 97) finds that the verb-particle construction in which the particle precedes the direct object (which we will refer to as the *particle-first variant* here) is significantly more frequent in writing while the construction in which the particle follows the direct object (which we will refer to as the (*object-first variant*), is significantly more frequent in speaking. However, Gries also argues that this correlation does not reflect a direct causal relation between channel and constructional choice, but that it is an epiphenomenon arising from the interaction of other factors.

Finally, with respect to the two future-tense constructions, there also appears to be general agreement that *will* and *going-to* are distributed asymmetrically across written and spoken channels, with the latter occurring more frequently in speech and speech-like texts (cf. e.g. Quirk et al. 1985, § 4.43, and Berglund 1997 for corpus-based verification). At the same time, most reference works agree that there is a semantic difference between the two constructions, with the *going-to* future expressing a greater degree of premeditation, certainty, and/or immediacy than *will* (e.g. Thompson and Martinet 1986: 185; Murphy 1986: 16). In addition, Quirk et al. (1985, § 4.43) claim that *going-to* is associated with more agentive events (cf. Gries and Stefanowitsch 2004: 113 ff. for corpus-based verification).

For each of the three constructions, we followed the procedure outlined in Stefanowitsch and Gries (2005) for three-dimensional collostructional analysis. This procedure is based on Configural Frequency Analysis (cf. von Eye 1990), in particular, the binomial version discussed in Krauth (1993, Sec. 1.10) (the binomial test is an exact test used to determine the probability to obtain x hits out of n trials when the probability of each hit is p). Our procedure involves constructing a three-dimensional frequency table for each potential collexeme along the dimensions COLLEXEME × CONSTRUCTION × CONTEXT. What precisely constitutes a CONTEXT variable depends on the research question. It may be an additional collexeme (as in covarying-collexeme analysis, cf. Stefanowitsch and Gries 2005), it may be regional dialect (as in Wulff, Gries, and Stefanowitsch 2005), or it may be any other variable that can be systematically assessed on the basis of corpus data.

For the present study, we chose the variables COLLEXEME (specified as lemma of the verb occurring in the finite verb slot of the construction vs.

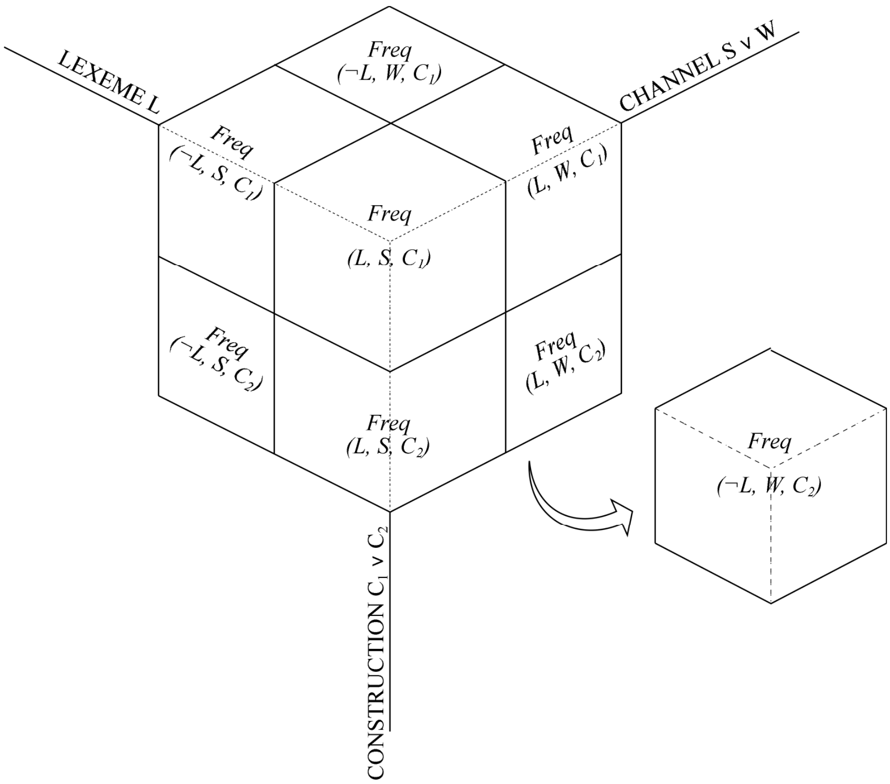

Figure 1. Three-dimensional distinctive collexeme analysis with the variables LEXEME × CONSTRUCTION × CHANNEL

all other verb lemmas), CONSTRUCTION, (specified, respectively, as Active vs. Passive, V-OBJ-Prt vs. V-Prt-OBJ, and *will*-future vs. *going-to*-future), and CHANNEL (specified as spoken vs. written language). This gives us the design COLLEXEME × CONSTRUCTION × CHANNEL. The three-dimensional frequency table that needs to be constructed for each verb in this design is represented schematically in Figure 1 (the extra cell shown at the bottom right of the table is the one opposite the top cell).

Once such tables are constructed for each verb lemma, a binomial test is performed for each cell (i.e., each combination of values of the three dimensions) to test whether its observed frequency deviates significantly from the expected one. The p-values resulting from these tests are taken to reflect the degree of association of the combination in question (cf. Stefanowitsch and Gries 2003, 2005, Gries and Stefanowitsch 2004a, b for extensive discussion and justification). For reasons that are partly mathematical and partly expository, these p-values are then transformed into their base ten logarithms and then the sign of the resulting value is set to plus for positive associations – i.e., for cases where the observed frequency is greater than expected –, and to minus for negative associations – i.e. for cases where the observed frequency is smaller than expected (these reasons cannot be discussed here; the mathematically-minded reader will find a discussion in Stefanowitsch and Gries 2005, and in Gries, Hampe, and Schönefeld 2005 n. 13).

The result of this procedure is a measure of association for each combination of values of the three dimensions. The final step of the analysis consists in choosing a relevant set of comparisons between these combinations and then drawing these comparisons by calculating the difference between the values corresponding to these combinations. In the present study, we focused on three comparisons, namely:

(1) Construction 1 vs. Construction 2 by Channel: this will show us which verbs are distinctive for each of the two constructions in the spoken channel and which verbs are distinctive for each of the two constructions in the written channel – no direct comparison between the two channels takes place;
(2) Spoken Channel vs. Written Channel by Construction: this will show us which verbs are distinctive for each of the two channels in the first construction and which verbs are distinctive for each of the two channels in the second construction – no direct comparison between the two constructions takes place;
(3) Channel S/W by Construction 1/2: this will show us which verbs display crossover effects, i.e. which verbs are associated with construction 1 in speaking but with construction 2 in writing or vice versa.

While the first two comparisons are meant to yield potential general differences between the behavior of the two constructions in the two channels, the third comparison is the one that most directly tests the hypothesis that the meaning of constructions may vary across channels.

Let us briefly illustrate our procedure by means of an example, namely the distribution of the verb *have* across the active and the passive construction and across the two channels. The three-dimensional table for this verb is shown in Table 3 (for details concerning extraction and lemmatization of the data see further below).

Table 3. Three-dimensional frequency table HAVE × CONSTRUCTION × CHANNEL

VERB	CONSTRUCTION	CHANNEL	Obs. frequency
Have	active	Spoken	2,642
Have	active	Written	1,291
Have	passive	Spoken	0
Have	passive	Written	1
other verbs	active	Spoken	29,893
other verbs	active	Written	19,335
other verbs	passive	Spoken	4,886
other verbs	passive	Written	7,046

As outlined above, we then computed for each of these configurations of VERB × CONSTRUCTION × CHANNEL (i) the expected frequency, and (ii) the probability to obtain the observed frequency by means of an exact one-tailed binomial test (whose p-value was then transformed into the base ten logarithms reflecting the strength and the direction of association). These interim results are shown in Table 4.

Table 4. Computing association strengths for HAVE × CONSTRUCTION × CHANNEL

Configuration	Observed frequency	Expected frequency	$p_{binomial}$	$\log_{10} p_{bin.}$ with +/− for attraction/ repulsion
have active spoken	2,642	1,847	5.57E-70	69.25
have active written	1,291	1,365.8	2.03E-02	−1.69
have passive spoken	0	414.6	2.35E-181	−180.63
have passive written	1	306.6	1.06E-131	−130.97
other active spoken	29,893	28,714	7.67E-21	20.12
other active written	19,335	21,234.2	6.13E-58	−57.21
other passive spoken	4,886	6,445.4	9.09E-101	−100.04
other passive written	7,046	4,766.4	1.44E-227	226.84

From the upper half of Table 4, each of the configurations with *have* can now be evaluated in isolation. For example, the configuration [*have* spoken active] is significantly more frequent than expected on the basis of complete independence of the three variables while the configuration [*have* spoken passive] is significantly less frequent than expected. However, there is more this data can offer since one can now also perform the pairwise comparisons mentioned above. Thus, while the use of loglinear analysis and similar methods was ruled out given the many tables with extremely low observed frequencies, it is still possible to measure individual preferences by simply subtracting the logarithms of the p-values of the configurations to be compared.

For example, in the spoken channel, *have* is strongly preferred in the active (as indicated by the positive value of 69.25) and strongly dispreferred in the passive (as indicated by the negative value of −180.63) so that the difference of the two (69.25− (−180.63) = 249.88) reflects that, within this channel (the variable that is held constant in the subtraction) – *have* exhibits a strong preference of the active.

We can perform similar comparisons by simply choosing the appropriate values for subtraction accordingly. For example, while the above subtraction aimed at identifying the preference of *have* in speaking, it is equally possible to determine the preference of *have* in writing: we simply subtract the value −130.97 (which indicates a very strong avoidance of have in the passive in writing) from the value −1.69 (which indicates a slight avoidance of *have* in the active in writing), yielding the value 129.28, which indicates that, just as in speaking, *have* also exhibits a strong preference for the active in writing.

Before we discuss the results of these and additional comparisons in quite some detail, note that a highly positive value such as 128.93 for *have* in writing does not necessarily mean that *have* strongly prefers actives in writing and strongly disprefers passives in writing: as is obvious from the above values, *have* in writing is dispreferred in actives (cf. −1.69) *as well as* in passives (−130.97). The dispreference in passive constructions is simply much stronger than that in active constructions and thus outweighs the much weaker one for actives, yielding the strong overall preference of *have* for actives.

2.1. Active and Passive Voice

We first extracted all main verbs from the ICE-GB and generated separate lemmatized frequency lists for the spoken and written channels (lemmatization was done manually). We then repeated the process for all main verbs in the passive voice, based on the annotation provided in the ICE-GB. By subtracting the passive frequencies from the overall frequencies for each verb, we calculated the active frequencies (active voice is not explicitly coded in the ICE-GB's annotation). We then followed the general procedure outlined above. The resulting association measures served as a basis of two comparisons: (i) a separate comparison of the active and the passive construction for each of the two channels, and (ii) a separate comparison of the spoken and the written channel for each of the two constructions.

2.1.1. Results

The results of the first comparison are shown in Table 5, where the verbs are rank-ordered in each column (i.e., for each construction); in other words, when one compares active vs. passive in speaking, the verb *have* is the verb most strongly associated with active while *concern* is most strongly associated with passive etc.

Table 5. Comparison of active and passive by channel

SPOKEN		WRITTEN	
ACTIVE	PASSIVE	ACTIVE	PASSIVE
have (249.88)	*concern* (43.74)	*have* (129.28)	*base* (55.42)
think (240.41)	*base* (29.99)	*want* (22.85)	*think* (54.36)
get (182.67)	*involve* (29.77)	*thank* (13.46)	*use* (52.95)
say (128.44)	*bear* (21.98)	*hope* (11.9)	*do* (47.13)
do (120.23)	*use* (13.7)	*include* (11.57)	*associate* (24.44)
want (91.94)	*engage* (13.35)	*try* (11.53)	*make* (18.87)
know (71.73)	*publish* (11.83)	*wish* (11.11)	*publish* (17.45)
see (62.04)	*enclose* (10.6)	*mean* (10.74)	*entitle* (16.84)
like (43.59)	*marry* (10.1)	*enclose* (9.87)	*deposit* (15.52)
mean (42.31)	*associate* (9.7)	*ensure* (9.4)	*relate* (15.32)
try (36.32)	*damage* (9.34)	*see* (8.26)	*design* (14.85)
remember (20.24)	*confine* (9.3)	*get* (8.12)	*derive* (14.7)
read (14.21)	*design* (9.2)	*provide* (8.02)	*require* (13.81)
believe (12.71)	*aim* (9.13)	*like* (7.98)	*report* (12.3)
suppose (12.69)	*distribute* (8.66)	*know* (7.7)	*store* (12)
feel (11.63)	*compare* (8.41)	*contain* (7.02)	*confine* (11.98)
take (11.07)	*release* (8.14)	*help* (6.96)	*engage* (11.34)
hope (9.29)	*injure* (7.17)	*reach* (6.68)	*link* (10.83)
hear (8.19)	*build* (6.98)	*believe* (6.22)	*record* (10.46)
find (7.58)	*entitle* (6.96)	*increase* (4.98)	*concern* (10.45)

In Gries and Stefanowitsch (2004a: 108 ff.) we showed that the active construction prefers low-dynamicity verbs whose logical objects are not easily construable as patients (i.e. verbs encoding states like *want* or *know* or time-stable processes like *see* or *remember*) while the passive construction prefers high-dynamicity verbs whose patients will be in a salient and typically permanent end state as a result of the event (i.e. verbs encoding actions involving transfer of energy like *damage* or *melt*). The results in Table 5 essentially confirm these observations: the top twenty distinctive collexemes for the active construction consist mainly of stative (often mental) verbs, with only five exceptions in the spoken data (*say*, *do*, *try*, *read*, and *take*); the trend is less clear in the written data, where only half of the top twenty collexemes are stative while the other half are relatively dynamic (*thank*, *include*, *try*, *enclose*, *ensure*, *get*, *provide*, *help*, *reach*, *increase*). For the passive construction, the distinctive collexemes are almost exclusively dynamic in both the spoken and the written data (with the exception of *think* and possibly *relate* in the latter).

Crucially for our research questions, the two channels behave essentially identically with respect to the dynamicity contrast between active and passive, even though it seems slightly more pronounced in the spoken data. The two constructions do not differ greatly across channels with respect to specific associations to individual verbs either; there is a substantial overlap of verbs (active: *have*, *get*, *want*, *know*, *see*, *like*, *mean*, *try*, *believe*; passive: *base*, *use*, *engage*, *publish*, *associate*, *design*, *entitle*). Even the collexemes that do differ in their distribution across channels do not show any obvious channel specificity, but such differences would be more likely to emerge in the second comparison, to which we will turn presently. Before we do so, note that there are three apparent cases of crossover, i.e., of verbs that are associated with one construction in one channel and the other construction in the other: *think*, *do*, and *enclose.* However, as pointed out above, it is not strictly speaking possible to contrast an individual construction across the two channels on the basis of the data in Table 5. Instead, the identification of crossover effects has to be done separately on the basis of the original cell values (see further below). Briefly, the reason is that the values in Table 5 are based on channel-internal contrasts between the two constructions, and thus the fact that, for example, *think* is listed for *active-spoken* and *passive-written* does not mean that it is generally associated with these two combinations, but only that it is vastly more strongly associated with *active-spoken* than with *active-written* and vastly more strongly with *passive-written* than with *passive-spoken* (recall the discussion of *have* in writing above). A direct comparison of channels shows, however, that *think* is so frequent with *active-written* that all other combinations occur less frequently than expected with all other combinations.

The results of the second of the three possible comparisons introduced above are shown in Table 6, where the verbs are once again rank-ordered within columns so that, for example, when one looks only at the active verbs in order to compare spoken vs. written language, *think* is the verb most strongly associated with the spoken channel whereas *enclose* is most strongly associated with the written channel.

Table 6. Comparison of channels by active/passive

ACTIVE		PASSIVE	
SPOKEN	WRITTEN	SPOKEN	WRITTEN
think (280.16)	*enclose* (23.91)	*concern* (23.73)	*use* (54.91)
do (178.84)	*provide* (23.25)	*involve* (13.63)	*have* (49.65)
get (157.08)	*include* (21.29)	*do* (11.48)	*see* (32.93)
say (122.35)	*use* (15.66)	*bear* (7.35)	*base* (25.8)
have (70.95)	*thank* (13.34)	*marry* (7.31)	*associate* (20.67)
want (57.93)	*contain* (11.84)	*arrest* (4.61)	*know* (19.35)
know (44.69)	*display* (11.64)	*crossexamine* (4.49)	*get* (17.47)
like (27.99)	*increase* (10.79)	*cube* (4.49)	*make* (17.23)
mean (23.49)	*involve* (10.02)	*readmit* (4.49)	*require* (15.62)
see (20.85)	*produce* (9.76)	*put* (4.21)	*think* (14.62)
read (19.83)	*receive* (9.61)	*knock* (3.99)	*entitle* (13.97)
try (16.07)	*concern* (9.57)	*hear* (3.6)	*deposit* (12.65)
put (15.28)	*satisfy* (8.42)	*inspire* (3.23)	*provide* (11.94)
remember (9.49)	*wish* (8.09)	*found* (3.23)	*derive* (11.94)
suppose (9.04)	*reach* (8.02)	*addict* (3.16)	*report* (11.47)
measure (8.91)	*fold* (7.88)	*adduce* (3.16)	*want* (11.15)
hear (7.34)	*bear* (7.75)	*extrude* (3.16)	*describe* (10.82)
believe (5.74)	*influence* (7.67)	*win* (3.04)	*store* (10.79)
take (5.59)	*ensure* (7.57)	*damage* (3.04)	*channel* (10.45)
make (5.34)	*require* (7.54)	*dress* (2.88)	*replace* (10.03)

This comparison confirms our observations concerning dynamicity: again, the active construction is clearly associated with stative (low-dynamicity) verbs and the passive construction with dynamic verbs. However, in this direct comparison of each construction across channels, clear formality differences emerge in the case of the active construction. In the spoken channel, it tends to occur with short (mostly monosyllabic) verbs of Germanic origin, while in the written channel it tends to occur with polysyllabic verbs of Romance/Latinate origin. This is, of course, the kind of formality difference that we would expect to characterize these two channels in general, so it is, perhaps, not altogether surprising to see it emerge in an individual construction. However, this difference does *not* emerge in the case of the passive construction, which occurs with both formal and informal vocabulary to approximately the same degree in both channels. In this case, there does seem to be an interaction of an individual construction and channel: the passive construction itself seems to be associated with formal vocabulary to a higher degree than the active construction, and it retains this association even in the more informal registers likely to be found in spoken language.

Finally, let us return to the crossover effects mentioned above. These are properly identified by identifying verbs for which the following conditions hold: (i) in each channel, they are significantly attracted to one and only one of the constructions in question, and (ii) they are attracted to different constructions in the different channels. For the active/passive comparison, only two such verbs were identified: *find* and *work* are significantly attracted to the combinations spoken/active and written/passive (the opposite case did not occur at all). However, both verbs point to an interaction between lexical semantics and channel rather than differences in constructional semantics: a closer examination reveals that we are dealing with different senses of these words in the two channels. While *work* in the written/passive combination occurred mainly with the meaning 'use' (e.g. *work one's muscles*, *work the sails*), its main use in the spoken/active combination is the phrasal verb *to work out sth*. Similarly, while *find* in the written/passive combination occurred mainly with the meaning 'exist' (e.g. *The fibres of group B are found in the autonomic nervous system* [ICE-GB W2A-026]), its main meaning in the spoken/active combination is 'realize' (e.g. *I found that I was not behaving well* [ICE-GB S1A-072]).

2.1.2. Discussion

In sum, the results of channel-sensitive collostructional analysis are essentially identical to those yielded by a 'channel-ignorant' analysis as far as constructional meaning in the narrow sense is concerned: we found no interaction between channel and semantics proper at all. This suggests that, in this respect, the method is comparable to the narrower procedure used in our previous work. With respect to channel-specific vocabulary, however, we did find a general tendency of the active construction, but not of the passive, to occur with channel-specific vocabulary: the passive construction occurs relatively frequently with formal vocabulary in both channels. This result clearly could not have been arrived at by the narrower method used in our previous work.

2.2. Verb-Particle Constructions

We extracted all verb phrases containing a phrasal adverb (the ICE-GBs label for particles) and a direct object from the ICE-GB based on the annotation provided and manually annotated the results as cases of either the object-first or the particle-first variant. We then generated separate lemmatized frequency lists for each channel and each variant; as before, lemmatization was done manually. We then followed the procedure outlined above. Again, the resulting association measures served as a basis for (i) a separate comparison of the two verb-particle constructions for each channel, and (ii) a separate comparison of spoken and written language for each construction.

2.2.1. Results

The results of the first comparison are shown in Table 7 (rank-ordered within columns as above).

Table 7. Comparison of the two verb-particle constructions by channel

SPOKEN		WRITTEN	
V-PRT-OBJ	V-OBJ-PRT	V-PRT-OBJ	V-OBJ-PRT
carry out (8.94)	*get back* (5.61)	*carry out* (6.79)	*get out* (1.95)
find out (5.83)	*play back* (4.74)	*set up* (6.14)	*get in* (1.81)
point out (5.67)	*get out* (4.51)	*point out* (5.82)	*get back* (1.56)
give up (3.72)	*turn off* (4.09)	*find out* (5.77)	*cheer up* (1.54)
work out (3.25)	*ring up* (3.64)	*take on* (5.13)	*fold across* (1.54)
set up (3.1)	*take out* (3.2)	*take up* (4.23)	*lock in* (1.54)
set out (3.03)	*get on* (3.07)	*build up* (3.61)	*psyche up* (1.54)
take on (2.85)	*get together* (3.07)	*rule out* (3.19)	*shove back* (1.54)
bring out (2.83)	*put up* (2.68)	*wipe out* (2.88)	*slow down* (1.54)
build up (2.53)	*put in* (2.58)	*pull off* (2.61)	*put back* (1.51)
take up (2.19)	*follow up* (2.45)	*write off* (2.61)	*put out* (1.51)
bring about (2.06)	*take off* (2.43)	*bring in* (2.46)	*hold together* (1.31)
poke out (1.85)	*phone up* (2.36)	*hold up* (2.29)	*have back* (1.09)
cut down (1.66)	*play forward* (2.32)	*scrape down* (2.06)	*trace back* (1.09)
hold off (1.6)	*turn round* (2.32)	*put on* (2.04)	*send out* (1.05)
cut off (1.51)	*let down* (2.19)	*leave out* (2.01)	*send back* (1.02)
radiate away (1.48)	*write down* (2.06)	*make out* (2.01)	*bring back* (0.97)
strike out (1.48)	*chip forward* (1.94)	*fill in* (1.98)	*ask home* (0.91)
type in (1.48)	*get up* (1.94)	*give up* (1.84)	*attract back* (0.91)
make out (1.37)	*have off* (1.94)	*keep up* (1.84)	*bang in* (0.91)

In Gries and Stefanowitsch (2004a), we showed that the particle-first variant is mainly associated with idiomatic verb-particle combinations, while the object-first variant is mainly associated with spatial and/or resultative readings where the particle encodes a final location or state (cf. also Gries 2003: 87 f. on data from the BNC). As before, this result is confirmed by the data in Table 7, and the two channels behave identically with respect to this semantic difference. Again, there is also a substantial overlap of individual collexemes for the particle-first variant (nine of the top twenty collexemes are shared – *carry out*, *find out*, *point out*, *give up*, *set up*, *take on*, *build up*, *take up*, and *make out*); interestingly, however there is very little overlap for the object-first variant. There seem to be a few channel-specific lexical choices, but their number is far from overwhelming – *poke out*, *phone up*, *turn 'round* and *have off* have an informal flavor, while *rule out* or *write off* could perhaps be argued to be more formal; however, both the spoken and the written data contain a substantial number of informal verb-particle combinations, especially in the case of the object-first variant. The results of the second comparison are shown in Table 8.

Table 8. Comparison of channels by the two verb-particle constructions

V-PRT-OBJ		V-OBJ-PRT	
SPOKEN	WRITTEN	SPOKEN	WRITTEN
work out (2.61)	*set up* (3.81)	*put in* (3.77)	*carry out* (4.78)
bring out (1.79)	*keep up* (3.03)	*play back* (3.65)	*give up* (2.8)
hold off (1.64)	*hold up* (2.96)	*take out* (3.52)	*lay down* (1.88)
send out (1.57)	*pull off* (2.94)	*ring up* (3.35)	*cheer up* (1.69)
poke out (1.56)	*write off* (2.94)	*turn off* (3.02)	*fold across* (1.69)
give out (1.5)	*wipe out* (2.89)	*get back* (2.61)	*lock in* (1.69)
put in (1.32)	*carry out* (2.63)	*get on* (2.48)	*psyche up* (1.69)
set out (1.27)	*rule out* (2.36)	*get together* (2.48)	*shove back* (1.69)
radiate away (1.27)	*scrape down* (2.32)	*get out* (2.32)	*slow down* (1.69)
strike out (1.27)	*lay down* (2.31)	*phone up* (2.27)	*curl up* (1.52)
type in (1.27)	*hand out* (2.14)	*write down* (2.12)	*find out* (1.37)
miss out (1.23)	*follow up* (2.09)	*tidy up* (1.96)	*bring about* (1.27)
blow out (0.95)	*leave behind* (2.05)	*play forward* (1.91)	*bring back* (1.25)
feed in (0.95)	*leave out* (2.05)	*turn round* (1.91)	*keep up* (1.14)
lay aside (0.95)	*bring together* (1.94)	*put up* (1.77)	*call in* (1.03)
line up (0.95)	*carry on* (1.76)	*bring in* (1.7)	*call up* (1.01)
mix up (0.95)	*put up* (1.74)	*follow up* (1.69)	*ask home* (0.98)
pack away (0.95)	*put on* (1.73)	*make up* (1.65)	*attract back* (0.98)
pass over (0.95)	*build up* (1.67)	*take off* (1.62)	*bang in* (0.98)
pull in (0.95)	*build in* (1.67)	*chip forward* (1.61)	*bend back* (0.98)

These results confirm the impression that idiomaticity is a determining factor for the choice between the constructions in both channels. In fact, Gries (2003: 100 f., 196) found that the effect of idiomaticity is slightly stronger in written language. These results also confirm the impression that channel does not strongly influence lexical choice in the verb-particle construction: there are no clear formality differences for either variant. There do seem to be some choices that are plausibly tied to situations that are more likely to be verbalized in one channel than the other (for example, *get together*, or *play back* and *play forward*, both of which occur exclusively in live football commentary in the ICE-GB); however, even these differences are not overwhelmingly obvious.

Finally, note that in the case of the verb-particle constructions, there was not a single crossover effect, i.e. there was not a single verb-particle combination that was significantly attracted only to object-first/spoken and particle-first/written (or vice versa).

2.2.2. Discussion

Two aspects of our results seem particularly noteworthy. First, that, again, neither of the two variants show evidence for channel-specific semantics, and that, in fact, there does not even seem to be a general tendency towards channel-specific vocabulary (instead, the vocabulary in both registers is relatively informal). Second, and more importantly for our present purposes, however, the two variants of the verb-particle construction nevertheless behave asymmetrically with respect to the variable *channel*: the particle-first variant is clearly less sensitive to channel influences than the object-first variant (as witnessed by the vastly higher lexical overlap for this variant). It seems to be the case that the strong association of the particle-first variant to idiomatic verb-particle combinations overrides channel influences in the same way as was the case for the strong association of the passive to formal verbs in the preceding section.

2.3. *Will*-future vs. *going*-*to*-future

We extracted from the ICE-GB all main verbs following the modal auxiliary *will* and all verbs following the string *going to* (either directly, or after some intermediate material). The latter were manually post-edited to remove false hits. We generated separate lemmatized frequency lists for each construction in each channel (as before, lemmatization was done manually). We then followed the procedure outlined above, with the same two comparisons as before.

2.3.1. Results

The results of the first comparison are shown in Table 9 (rank-ordered within columns).

Table 9. Comparison of the two future constructions by channel

SPOKEN		WRITTEN	
will	*going to*	*will*	*going to*
see (4.38)	*say* (16.06)	*be* (17.71)	*do* (4.76)
know (2.32)	*be* (14.7)	*give* (3.7)	*let* (1.38)
want (2.24)	*have* (9.33)	*find* (3.57)	*chubb lock* (1.27)
notice (1.86)	*do* (9.16)	*have* (3.41)	*conduct* (1.27)
read (1.55)	*go* (6.03)	*make* (2.95)	*extradite* (1.27)
find (1.42)	*happen* (5.99)	*consider* (2.83)	*spoil* (1.27)
speak (1.32)	*use* (4.87)	*increase* (2.76)	*go* (1.26)
agree (1.27)	*win* (4.22)	*receive* (2.64)	*say* (1.14)
explain (1.27)	*stay* (3.97)	*depend* (2.48)	*manage* (1.11)
recall (1.23)	*get* (3.71)	*send* (2.28)	*deliver* (1.1)
learn (1.23)	*buy* (3.41)	*add* (2.25)	*feed* (1.1)
lie (1.23)	*show* (3.39)	*include* (2.25)	*talk* (0.81)
teach (1.23)	*put* (2.94)	*write* (2.24)	*die* (0.73)
hear (1.15)	*increase* (2.79)	*become* (2.24)	*hit* (0.73)
ring (1.11)	*ask* (2.68)	*provide* (2.21)	*mention* (0.7)
take (1.11)	*suggest* (2.63)	*pay* (2.19)	*win* (0.61)
give (1.07)	*talk* (2.52)	*occur* (2.08)	*develop* (0.61)
work (1.07)	*pass* (2.24)	*reach* (1.97)	*move* (0.56)
mention (1.07)	*pay* (2.22)	*cause* (1.94)	*spend* (0.53)
answer (1.06)	*cause* (2)	*wait* (1.81)	*introduce* (0.48)

In Gries and Stefanowitsch (2004a), we were able to confirm earlier claims that the *going-to*-future has a much stronger preference for verbs encoding dynamic actions than the *will*-future. These results are further confirmed by the data in Table 9. Among the top twenty distinctive collexemes of the *going-to*-future there are only three low-dynamicity verbs in the spoken data (*be*, *have*, *stay*) and one in the written data (*die*); in contrast, among the top twenty distinctive collexemes of the *will*-future there are nine low-dynamicity verbs in the spoken data (*see*, *know*, *want*, *notice*, *find*, *agree*, *recall*, *learn*, and *hear*), and the same number in the written data (*be*, *find*, *have*, *consider*, *receive*, *depend*, *become*, *occur*, and *reach*). As before, then, we find the same semantic contrast in both channels. Also as before, we do not find strong channel influences on lexical choice; but recall that these are really only expected to emerge in the direct comparison of channels, to which we now turn. The results of the second comparison are shown in Table 10.

Table 10. Comparison of channels by the two future constructions

will		*going to*	
SPOKEN	WRITTEN	SPOKEN	WRITTEN
do (4.89)	*be* (12.57)	*be* (19.83)	*chubb lock* (1.21)
see (4.8)	*increase* (5.01)	*say* (16.22)	*conduct* (1.21)
go (4.51)	*pay* (3.5)	*have* (11.88)	*extradite* (1.21)
come (3.61)	*consider* (3.41)	*do* (9.29)	*spoil* (1.21)
want (2.39)	*receive* (3)	*go* (9.28)	*deliver* (0.98)
tell (2.15)	*add* (3)	*happen* (5.01)	*feed* (0.98)
know (2.07)	*cause* (2.95)	*use* (4.63)	*bear* (0.85)
talk (1.95)	*write* (2.84)	*get* (4.05)	*tour* (0.83)
notice (1.92)	*occur* (2.75)	*show* (3.96)	*serve* (0.81)
mention (1.87)	*depend* (2.66)	*talk* (3.66)	*close* (0.51)
hear (1.8)	*include* (2.64)	*win* (3.54)	*move* (0.51)
introduce (1.75)	*assist* (2.48)	*put* (3.41)	*provide* (0.47)
develop (1.65)	*base* (2.48)	*come* (3.26)	*depend* (0.42)
read (1.55)	*last* (2.48)	*stay* (3.19)	*let* (0.41)
speak (1.53)	*qualify* (2.48)	*ask* (3.15)	*receive* (0.39)
learn (1.53)	*understand* (2.48)	*make* (3.01)	*consider* (0.36)
lie (1.53)	*apply* (2.3)	*buy* (2.73)	*remain* (0.36)
teach (1.53)	*entitle* (2.3)	*tell* (2.21)	*finish* (0.33)
spend (1.42)	*reach* (2.26)	*suggest* (1.86)	*include* (0.3)
try (1.39)		*play* (1.71)	*hold* (0.3)

The dynamicity differences show up less clearly than before in this comparison, but instead, clear formality differences emerge between the channels for both constructions: the top collexemes in the spoken data consist mainly (though not exclusively) of short words of Germanic origin while in writing the collexemes tend to be longer Romance or Latinate words.

Finally, there was one case of crossover: the verb *be* is significantly associated with *going-to*/spoken and with *will*/written. Given the wide range of functions served by this verb, it is impossible to come up with an explanation for this without a more detailed investigation

2.3.2. Discussion

The results of this case study do not differ fundamentally from those of the previous two, and thus there is little to say about them beyond what was already said above. One noteworthy fact is that in this case both constructions showed a relatively high sensitivity to channel with respect to lexical choice in that their top collexemes in the written data were much more formal than those in the spoken data. Thus, none of the two constructions seems to have its own formality preferences (unlike in the previous two case studies).

3. General Discussion

Despite possible overgeneralizations in the way in which 'channel' was operationalized in the present study, three important conclusions follow from the results presented above.

The first conclusion is that there is no evidence so far to suggest that constructional semantics in the narrow sense interacts with channel in such a way that there are differences in a construction's meaning across channels. The semantic characterizations of the constructions in question arrived at in earlier work were found to hold in both spoken and written language. This was shown in each case by the two major comparisons we drew (Construction 1 vs. Construction 2 by Channel and Spoken Channel vs. Written Channel by Construction) as well as by our investigation of crossover effects. To be fair, two caveats are in order, one concerning the major comparisons, and one concerning the investigation of crossover effects. With respect to the comparisons, the method employed here (like corpus-based methods in general) is necessarily superficial in two important respects: first, it only captures those semantic differences that have consequences for lexical choice, and second, it ignores the polysemy of lexical items and constructions. More fine-grained annotation of large corpora may allow future research to overcome this problem, but for now, corpus linguists will have to live with it. With respect to the crossover effects, it has to be kept in mind that the criterion we apply here is extremely strict since we only accept cases where a word is significantly associated to a single construction in each channel. It is conceivable to take a less categorical approach and to accept cases whose association strengths exhibit crossover effects.

Even given these caveats, however, the results seem plausible and are perhaps not entirely unexpected. Linguistic signs may differ in their connotative meaning and thus have different likelihoods of occurrence in different channels. The (invariant) meaning of a given sign may certainly be exploited to yield different communicative effects in different contexts, and future construction-based research would certainly profit from taking this into account (cf. Stefanowitsch and Gries 2005 for a brief discussion of one such case). However, if meaning itself were to differ according to context, this would seriously threaten the integrity of the linguistic system and hence its usefulness for communication. This is true of lexical items, which are relatively accessible to conscious introspection and hence to meta-linguistic discourse; how much more true should we expect it to be of grammatical constructions, which are inaccessible to conscious introspection.

The second conclusion is that there can be channel-specific associations between constructions and individual lexical items. Again, this is perhaps not entirely surprising, since it is a well-known fact that channel may influence lexical selection (it is, in fact, reflected in most modern dictionaries by labels such as *formal*, *informal*, *colloquial*, etc.) Since spoken and written channels differ in vocabulary, they inevitably also differ in terms of specific collostructional relationships between words and constructions. Such specific collostructional relationships are doubtless of theoretical and of practical interest and thus could fruitfully be integrated into any collostructional analysis and into construction-based research in general.

The final, and perhaps most important conclusion from a construction-grammar viewpoint is that different constructions differ with respect to the degree to which they exhibit channel-specific collostructional relationships: while some constructions (the active construction, the particle-first variant of the verb-particle construction, and the two future-tense constructions) are relatively sensitive to the formality differences associated with spoken vs. written channels, other constructions are rather insensitive, and seem to have their own, construction-specific preferences which they retain regardless of the channel they are used in (the passive construction always has a relative preference for formal lexical items, the particle-first variant of the verb-particle construction always has a relative preference for informal, idiomatic lexical items. Such differences in the degree to which lexical choices within a given construction reflect either general properties of a particular channel or specific properties of the construction itself constitute an important fact about it that has to be recorded as part of the construction's specification if construction grammar takes serious its commitment to a broad understanding of meaning.

References

Berglund, Ylva
1997 Future in Present-day English: Corpus-based evidence on the rivalry of expressions. *ICAME Journal* 21: 7–19.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan
1999 *Longman Grammar of Spoken and Written English*. London: Longman.

Blankenship, Jane
1962 A linguistic analysis of oral and written style. *Quarterly Journal of Speech* 48: 419–422.

Brown, Gillian and George Yule
1983 *Discourse Analysis*. Cambridge: Cambridge University Press.

Chafe, Wallace L.
1982 Integration and involvement in speaking, writing, and oral literature. In: Deborah Tannen (ed.), *Spoken and Written Language: Exploring Orality and Literacy,* 35–54. Norwood, NJ: Ablex.

Croft, William
2001 *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Eye, Alexander von
1990 *Introduction to Configural Frequency Analysis: The Search for Types and Antitypes in Cross-classifications*. Cambridge: Cambridge University Press.

Goldberg, Adele E.
1995 *Constructions. A Construction-Grammar Approach to Argument-Structure Constructions*. Chicago/London: The University of Chicago Press.

Greenbaum, Sidney (ed.)
1996 *Comparing English Worldwide: The International Corpus of English*. Oxford: Clarendon Press.

Gries, Stefan Th.
2003 *Multifactorial Analysis in Corpus Linguistics: A Study of Particle Placement*. London/New York: Continuum.

Gries, Stefan Th., Beate Hampe and Doris Schönefeld
2005 Converging evidence: Bringing together experimental data on the association between verbs and constructions. *Cognitive Linguistics*.

Gries, Stefan Th. and Anatol Stefanowitsch
2004a Extending collostructional analysis: A corpus-based perspectives on 'alternations'. *International Journal of Corpus Linguistics* 9.1: 97–129.
2004b Co-varying collexemes in the *into*-causative. In: Michel Achard and Suzanne Kemmer (eds.), *Language, Culture, and Mind*, 225–236. Stanford, CA: CSLI
in press Cluster analysis and the identification of collexeme classes. In: John Newman and Sally Rice (eds.), *Experimental and Empirical Methods in the Study of Conceptual Structure, Discourse, and Language*. Stanford, CA: CSLI Publications.

Grondelaers, Stefan
this volume National variation in the use of *er* "there". Regional and diachronic constraints on cognitive explanations Regional constraints on cognitive explanations.

Krauth, Joachim
1993 *Einführung in die Konfigurationsfrequenzanalyse (KFA)*. Weinheim: Beltz.

Murphy, Raymond
1986 *English Grammar in Use*. 4th rev. ed. Cambridge: Cambridge University Press.

Poole, Millicent E. and T.W. Field
1976 A comparison of oral and written code elaboration. *Language and Speech* 19: 305–311.

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

Siegel, Sidney.
1956 *Nonparametric Statistics for the Behavioral Sciences*. Tokyo: McGraw-Hill & Kogakusha Ltd.

Speelman, Dirk
this volume Variation in the choice of adjectives in the two main national varieties of Dutch.

Stefanowitsch, Anatol and Stefan Th. Gries
2003 Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8.2: 209–243.
2005 Co-varying collexemes. *Corpus Linguistics and Linguistic Theory* 1.1: 1–46.

Thompson, Audrey J. and Agnes V. Martinet
1986 *A Practical English Grammar*. 4th ed. Oxford: Oxford University Press.

Tummers, José, Kris Heylen, and Dirk Geeraerts
2006 Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1.2: 225–261

Wulff, Stefanie, Stefan Th. Gries, and Anatol Stefanowitsch
2005 Brutal Brits and argumentative Americans: What collostructional analysis can tell us about lectal variation. Paper presented at the International Cognitive Linguistics Conference 2005, Seoul, South Korea.