Converging evidence from corpus and experimental data to capture idiomaticity*

STEFANIE WULFF

Abstract

It is a by now established fact that idiomaticity cannot be equated with non-compositionality alone, but is a complex concept that is also associated with various aspects of formal flexibility. This raises the question to what extent speakers call up these different factors when judging the overall idiomaticity of a phrase.

In the present paper, experimental and corpus-linguistic methodology are combined to address this question. For a total of 39 V NP-idioms of the kind make a point *or* take the plunge*, comprising more than 13,000 tokens obtained from the British National Corpus, their compositionality, syntactic, lexico-syntactic, and morphological flexibility were assessed corpus-linguistically. The corpus-based results thereby obtained were then correlated with native speakers' overall idiomaticity judgments in a multiple regression analysis to determine each factor's impact on the overall judgments. The results indicate that speakers indeed rely on multiple factors simultaneously, with lexico-syntactic and morphological factors being even more important than compositionality, and verb-related being more important than NP-related information. Overall, the results back up the theoretical concept of a collocation-idiom continuum, and demonstrate how various, and sometimes competing, motivations determine a phrase's position on this continuum.*

Keywords: Compositionality; construction grammar; corpus linguistics; idiomaticity; V NP-construction.

1. Introduction

Over the past three decades, various psycholinguistic, discourse-functional, and phraseological studies have suggested that idiomaticity is best conceived

Corpus Linguistics and Linguistic Theory 5–1 (2009), 131–159 1613-7027/09/0005-131
DOI 10.1515/CLLT.2009.006

of as a complex phenomenon that comprises a variety of factors, the most recurrent of which are the following:

– compositionality (Comp): to what extent do the component words contribute to the meaning of an idiomatic construction?
– tree-syntactic flexibility (SF): what kinds of syntactic variation does the idiomatic construction license?
– lexico-syntactic flexibility (LF): what kind of material is found inserted into the idiomatic construction?
– morphological flexibility (MF): what kinds of morphological variation do component words/does the idiomatic construction occur in?

The present study is the first to define all the above parameters on the basis of corpus data (rather than merely applying given definitions to or checking them against corpus data). However, it is an established fact that idiomaticity is a psychological, intuition-based concept that does not manifest itself in corpus data in a unique way, and consequently, it cannot be tracked down directly from the corpus data. All you can do is retrieve all instances of a construction and count the number of times it occurs in different syntactic configurations, morphological variants, or with or without inserted lexical material. So why should one want to turn to corpus data at all? I would like to argue that, from a usage-based perspective, this is exactly what native speakers also do: all linguistic behavior is determined by the speaker's linguistic environment, and large-scale, balanced corpora like the British National Corpus used here can be regarded as the best approximation to a speaker's linguistic environment that is available today. With respect to idiomaticity judgments, it is assumed that speakers monitor the distributional characteristics of a construction at different levels, and by weighting this distributional information in a particular way, they arrive at a judgment of overall idiomaticity. While the weighting or idiomaticity formula itself is not retrievable from corpus data, the distributional properties entering into that formula are. Accordingly, the term *idiomaticity* is used here to refer to the quality that speakers construct on the basis of different *idiomatic variation parameters*. Bearing this crucial distinction in mind, there are two central questions addressed in this study: first, what does a usage-based model of idiomatic variation look like? And second, which aspects of the linguistic environment figure in speakers' assessments of idiomaticity, and how important are they relative to each other?

2. A constructionist perspective

This study adopts a constructionist approach to language (Goldberg 1995, 2006). In Construction Grammar, constructions are defined as form-meaning pairings that range from simplex morphemes to complex patterns which are schematized (i.e., lexically unspecified) to different extents; all constructions are assumed to be stored in the constructicon (i.e., the extended mental lexicon) as represented in Figure 1.

Morpheme	*pre–*, *–ing*
Word	*avocado*, *anaconda*, *and*
Complex word	*Dare-devil*, *shoo-in*
Complex word (partially filled)	[N-s] (for regular plurals)
Idiom (filled)	*going great guns*, *give the Devil his due*
Idiom (partially filled)	*jog* <someone's> *memory*, *send* <someone> *to the cleaners*
Covariational Conditional	The Xer the Yer (e.g. *the more you think about it, the less you understand*)
Ditransitive (double object)	Subj Obj1 Obj2 (e.g. *he gave her a fish taco*; *he baked her a muffin*)
Passive	Subj aux VPpp (PP_{by}) (e.g. *the armadillo was hit by a car*)

Figure 1. *A schematic representation of the constructicon (adapted from Goldberg 2006: 5)*

While the integration of non-compositional expressions has been a major impetus for the development of Construction Grammar, constructions are not restricted to non-compositional phrases; even highly transparent expressions that are used sufficiently often to become entrenched in the speaker's mental lexicon qualify as constructions (Goldberg 2006: 64). Indeed, both lexico-syntactic variability and compositionality prevail at all levels of the constructicon (Croft and Cruse 2004), albeit in different shades of prominence and relative importance. Accordingly, the fact that only some of the constructions in Figure 1 are referred to as idioms is slightly misleading since it is not only these constructions that can be labeled as (more or less) idiomatic. However, it is plausible to argue that due to their low degree of schematization and moderate complexity, idiomaticity effects are most obvious in these constructions. For this reason, the present study focuses on one such type: V NP-constructions, i.e. monotransitive verbs followed by a direct object noun phrase.

3. Data

All instantiations (including all variant forms) of the 39 V NP-constructions shown in (1) (13,141 tokens total) were retrieved from the British National Corpus (BNC); their corpus frequencies are given in parentheses.[1]

(1) *bear* DET[2] *fruit* (90), *beg* DET *question* (163), *break* DET *ground* (133), *break* DET *heart* (185), *call* DET *police* (325), *carry* DET *weight* (157), *catch* DET *eye* (491), *change* DET *hand* (212), *close* DET *door* (827), *cross* DET *finger* (150), *cross* DET *mind* (140), *deliver* DET *good* (145), *do* DET *trick* (155), *draw* DET *line* (310),

fight DET *battle* (192), *fit* DET *bill* (116), *follow* DET *suit* (135), *foot* DET *bill* (109), *get* DET *act together*[3] (142), *grit* DET *tooth* (164), *have* DET *clue* (232), *have* DET *laugh* (98), *hold* DET *breath* (292), *leave* DET *mark* (145), *make* DET *headway* (136), *make* DET *mark* (213), *make* DET *point* (1,005), *make* DET *face* (371), *meet* DET *eye* (365), *pave* DET *way* (269), *play* DET *game* (290), *scratch* DET *head* (100), *see* DET *point* (278), *take* DET *course* (294), *take* DET *piss* (121), *take* DET *plunge* (115), *take* DET *root* (113), *tell* DET *story* (1,942), *write* DET *letter* (1,370)

The collected corpus data were complemented with overall idiomaticity judgments for the same 39 V NP-constructions. 39 subjects from the University of Sheffield[4] participated in a questionnaire experiment; all were first year students of English. Each subject was presented a different construction as a reference construction and was asked to assign any number to this reference construction that they felt should represent its idiomaticity. The syntactic structures in which the V NP-constructions were presented were selected from the corpus data such that the structures are typical of the V NP-construction in question.[5] Subjects were then asked to judge the remaining 38 constructions relative to that reference construction. This method is called magnitude estimation (Bard et al. 1996), and it outperforms other judgment scaling techniques in various ways: since there are no restrictions on the number of values used to measure the property of interest, both the range of responses as well as the distribution of individual responses within that range are informative; moreover, differences in the values of these ratio-scaled judgments directly reflect the subjects' perceived differences.

Subjects were given the following instructions:

Dear participant,
The present questionnaire is concerned with so-called idiomatic sentences. Idiomatic sentences are the kind of sentences you typically find in dictionaries or phrase books. Some examples are the following:
The government got its fingers burnt.
Vincent has spilled his guts.
The knives are out for me at the moment.
Your task in the questionnaire attached to this instruction sheet is to judge how idiomatic each sentence is. In other words, you are asked to decide to what degree the sentence is different from 'normal' sentences, and how reasonable you think it is that this phrase is included in dictionaries or phrase books.
[. . .]

Note that none of the potentially contributing parameters, such as compositionality or syntactic flexibility, was introduced in further detail so as not to bias the participants' ratings towards any parameter. Given these deliberately vague instructions, the question arises how confident we can be that the subjects actually judged the constructions according to idiomaticity and not something else. There are several pieces of evidence that support the validity of the questionnaire design. First, an inspection of the resulting ranking of constructions according to a normed idiomaticity index from 0 (non-idiomatic) to 1 (idiomatic) showed that the ranking largely reproduced established findings from the literature, with metaphorical-opaque idioms ranking highest (e.g. *make* DET *headway*, *take* DET *plunge*, and *foot* DET *bill*), metaphorical expressions obtaining the middle ranks (e.g. *see* DET *point*, *draw* DET *line*, and *fight* DET *battle*), and transparent constructions ranking lowest (e.g. *write* DET *letter*, *tell* DET *story*, and *call* DET *police*). Second, inter-subject consistency was extremely high (Cronbach's alpha=.923; cf. Cronbach 1951), so subjects agreed on which constructions rank high or low in overall idiomaticity. Third, while it is theoretically possible that the data reflect some concept other than idiomaticity very consistently, the judgments obviously do not (exclusively) reflect subjects' familiarity with the V NP-constructions: the correlation between the mean normed values and the corpus frequency of the constructions is only moderately high ($r_{Pearson} = -.635$).

4. Corpus-linguistic definitions of idiomatic variation parameters

4.1. A collocation-based compositionality measure

Nunberg et al. (1994: 498) departed from the earlier understanding of compositionality as a binary phenomenon such that constructions were either idioms or non-idioms, and instead defined it as "the degree to which the phrasal meaning, once known, can be analyzed in terms of the contributions of the idiom parts". Various studies have provided empirical support in favor of this 'compositional view' of language (Cacciari and Tabossi 1988, Peterson and Burgess 1993, Titone and Connine 1994, to name but a few), showing that even in highly non-compositional constructions, the semantics of the component words are activated at least to some degree (cf. e.g. Glucksberg 1993), and that consequently, the literal meanings of words facilitate the comprehension of idioms to the extent that they semantically overlap with the idiomatic meaning (cf. e.g. Gibbs and Nayak 1989, Gibbs et al. 1989).

Starting out from the assumption that compositionality is a function of the semantic similarity of the constituent words and the phrasal expression, a number of corpus-linguistic compositionality measures have been proposed lately. Some measure compositionality via the ability to replace component words without losing the idiomatic interpretation of the construction (Lin 1999, McCarthy et al. 2003); others measure it via the semantic similarity of the contexts of the constructions compared with those of the component words (Schone and Jurafsky 2001, Bannard et al. 2003, Bannard 2005). The measure presented here also adopts the latter approach; more specifically, it elaborates on a study on verb particle constructions (VPCs) by Berry-Rogghe (1974), who defined the compositionality of a VPC as the overlap between the sets of collocates associated with the particle and the VPC respectively (Berry-Rogghe 1974: 21–22). She computed a compositionality index value R as shown in (2) by dividing the overlap by the total number of collocates of the VPC. R can range between 0 when there is no overlap at all, so the VPC is perfectly non-compositional, and 1 when the collocate sets of the particle and the VPC match perfectly, i.e. the VPC is fully compositional.

(2) $$R = \frac{\text{no. of collocates of VPC} \cap \text{P}}{\text{no. of collocates of VPC}}$$

For the present study, the original R was improved in several ways. First of all, the analysis was based on the 100 million word British National Corpus rather than the 202,000 word corpus (of texts by D. Lessing, D.H. Lawrence, and H. Fielding) used by Berry-Rogghe, which provides a more comprehensive semantic profile of the component words and the construction. Secondly, instead of the z-score, the Fisher Yates Exact (FYE) test was used (for a detailed account of why this can be argued to be the preferable choice, cf. Stefanowitsch and Gries 2003: 217–218). Last but not least, a constructionist approach as adopted here postulates that every component word makes some contribution to the higher order construction it occurs in, so for the V NP-constructions, the R-values for both the verb and the noun phrase were considered. In order to weight the contributions made by the component words relative to each other, another piece of information was added to the formula: while the original R-value reflects how much of the construction's semantics is accounted for by the component word, the extended R-value to be presented here also takes into account how much of itself every component word contributes. Let me illustrate the relevance of this information with the example of *take the plunge*. *Take* is a high frequency verb, and consequently, it attracts a high number of significant collocates. *Plunge*, on the other hand, is much less frequent and accordingly attracts fewer significant collocates. Given this frequency bias, *take* stands a much higher chance to share a substantial number of its collocates with any other collocate set than *plunge* does, yet this does not necessarily mean that *take* is indeed semantically similar to the phrase in question. This potential misrepresentation can be avoided by also considering how many of its collocates any verb contributes to the phrasal collocate set; for *take the plunge*, we see that *take* actually shares only a fraction of its collocates with *take the plunge*, whereas *plunge* shares nearly all of its collocates with that one construction.

Accordingly, these two pieces of information were combined as follows. For each component word W, the original *R*-value was calculated and then relativized against what I refer to as the share of the component word, which is the ratio of collocates shared between the component word and the construction C divided by the total number of the component word's collocates; cf. the formula in (3). The overall compositionality value for a construction is the sum of all component words' relativized *R*-values. Table 1 provides an overview of the results, sorted according to the extended *R*-values.[6]

(3) $$\text{contribution W} = \overbrace{\frac{\text{n colls C in colls W}}{\text{n colls C}}}^{R_{\text{W}}} \times \overbrace{\frac{\text{n colls C in n colls W}}{\text{n colls W}}}^{\text{share of } R_{\text{W}}}$$

Table 1. *Extended R-values for V NP-constructions*

Construction	*R*	Construction	*R*
make DET *headway*	.003	*carry* DET *weight*	.137
take DET *plunge*	.004	*follow* DET *suit*	.147
take DET *piss*	.008	*beg* DET *question*	.150
make DET *face*	.021	*bear* DET *fruit*	.160
get DET *act together*	.026	*deliver* DET *good*	.161
pave DET *way*	.033	*cross* DET *finger*	.171
change DET *hand*	.051	*draw* DET *line*	.174
take DET *course*	.058	*take* DET *root*	.185
foot DET *bill*	.058	*cross* DET *mind*	.225
see DET *point*	.062	*hold* DET *breath*	.232
leave DET *mark*	.074	*break* DET *heart*	.238
grit DET *tooth*	.079	*fight* DET *battle*	.288
break DET *ground*	.079	*do* DET *trick*	.340
meet DET *eye*	.101	*make* DET *point*	.359
make DET *mark*	.106	*scratch* DET *head*	.368
have DET *laugh*	.106	*close* DET *door*	.421
fill DET *bill*	.108	*catch* DET *eye*	.432
have DET *clue*	.117	*tell* DET *story*	.730
call DET *police*	.117	*write* DET *letter*	.844
play DET *game*	.132		

A compositionality continuum emerges from Table 1, with phrases like *make* DET *headway* and *take* DET *plunge* ranking lowest, metaphors such as *break* DET *ground*, *bear* DET *fruit*, and *draw* DET *line* obtaining middle ranks, and phrases such as *close* DET *door*, *tell* DET *story*, and *write* DET *letter* ranking highest. By and large, this ranking not only accords with established phraseological models (cf. Fernando 1996); the measure also reproduces the established fact that compositionality and token frequency correlate highly

($r_{Pearson}$=.802; cf. Barkema 1994a). Moreover, the compositionality measure derives a lot of plausibility from its compatibility with theoretical premises: it is the first measure which considers all component words, reflecting the constructionist view that a complex phrase is a manifestation of several smaller constructions (Goldberg 2006: 10). Secondly, the measure weights the contributions made by the component words relative to each other on an item-specific basis (that is, it licenses the possibility that the contribution made by *point* in *make* DET *point* may be higher or lower than in *see* DET *point*), implementing the assumption that constructions are differently entrenched in the mental lexicon (Langacker 1987: 59). Thirdly, by combining the original *R*-value and the share value, the measure leaves room for a potential backward influence of the constructional semantics on the component word's semantics (Langacker 1987).

4.2. Measuring formal flexibility: An extension of Barkema (1994b)

The formal behavior of the V NP-constructions was measured corpus-linguistically elaborating on Barkema's (1994b) formula for describing the syntactic and morphological flexibility of English noun phrases such as *cold war*. In a first step, the so-called 'received form profile', i.e. an inventory of *cold war*'s variant forms in the corpus, is determined. The following profile emerges from the 20-million word *Birmingham Corpus* (Barkema 1994b: 43):

(4) *renewed Cold War; the melting Cold War; the world Cold War; continuing, ever-present Cold War; the Cold War won by Europeans who 'destalinized' Eastern Europe; the cold war which threatened to divide the world into two ideological armed camps; a not-so-cold war against Kaddafi; the awkward cold war thought up by the American paranoids, who should be back in the law offices of middlewestern towns; a period of hot and cold civil war which ended with Hitler's invasion of Austria; a kind of cold civil war; the cold war that existed between the two giants, the United States and* [. . .]; *the Cold War in Washington; the cold war between the Nature Conservancy Council and the farmers*

Cold war occurs 124 times in the corpus, among which 111 times out of those in its base form. (4) lists the remaining 13 variant forms.

The standard of comparison for the flexibility of *cold war* is a corresponding inventory of all variant forms of the syntactic construction underlying *cold war*, which is an adjective premodifying a singular common head noun (subsequently referred to as adjective-noun construction). This construction occurs 3,171 times altogether in the *Birmingham Corpus*, and 1,257 times in its base form. The logic behind comparing the flexibility of *cold war* with that of the adjective-noun construction is that if *cold war* behaves like a free expression, then for any particular form (base or variant), the ratio of the frequency of occurrence of this form to the total number of occurrences of *cold war* should be about the same as the ratio of the frequency with which the adjective-noun construction takes this form to the total number of occurrences of the adjective-noun construction; cf. (5).

(5) $$\frac{\text{n form A } \mathit{coldwar}}{\text{n base + variant forms } \mathit{coldwar}} = \frac{\text{n form A in adj} - \text{noun construction}}{\text{n base + variant forms in adj} - \text{noun construction}}$$

The expected frequency of any variant form A can accordingly be computed on the basis of the formula in (6).

(6) $$\text{n exp form A } \mathit{coldwar} = \text{n base + variant forms } \mathit{coldwar} \times \frac{\text{n form A in adj} - \text{noun construction}}{\text{n base + variant forms in adj} - \text{noun construction}}$$

According to (6), we would expect *cold war* to occur 49 out of 124 times in its base form; cf. (7).

(7) $$\text{n exp base form } \mathit{coldwar} = 124 \times \frac{1{,}257}{3{,}171} = 49.15(\sim 49)$$

However, the actually observed frequency of *cold war* in its base form is 111/124, which already hints at the limited flexibility of that construction. Table 2 provides a detailed flexibility profile for *cold war*. The leftmost column specifies the form, followed by its expected frequency and observed frequencies (all accompanied by their values in percent of all 124 occurrences of *cold war*). The rightmost column provides the difference between observed and expected frequency; positive numbers mean that *cold war* occurs more often in that particular form than the adjective-noun construction in general, negative numbers mean that *cold war* occurs less frequently.

For the purpose of the present study, Barkema's method was slightly modified in several respects. Firstly, it has to be noted that while Barkema collapses morphological, lexico-syntactic and tree-syntactic aspects of flexibility, these were treated as separate independent variables here. Table 6 in the appendix provides an overview of all formal flexibility factors considered; the leftmost column provides the abbreviated parameter labels subsequently used, followed by an example and, in the right-hand column, the information level at which the parameter was coded.

Secondly, the flexibility profile of the V NP-construction, henceforth referred to as the baseline, was established slightly differently. Since large-scale, syntactically and/or morphologically annotated corpora are not available yet (at least not to the general public), Barkema suggested using a large corpus to extract the lexically specified constructions (like *make* DET *headway* or *make* DET *point*) in order to retrieve a sufficient number of tokens, and to obtain the corresponding baseline data from a small, fully annotated corpus in order

Table 2. *Flexibility profile for* cold war *(adapted from Barkema 1994b: 50)*

Form	$n_{exp.}$ (%)	$n_{obs.}$ (%)	Diff. %
base form	49.15 (39.64)	111 (89.52)	+49.88
premodifying adjective	4.18 (3.37)	3 (3.2)	−.17
postmodifying prepositional phrase	7.59 (6.12)	2 (1.6)	−4.52
premod. adjective + postmodifying prepositional phrase	19.24 (15.52)	2 (1.6)	−13.92
premod. adjective + postmodifying clause	1.17 (.94)	1 (.008)	−.93
postmod. past participial clause	.98 (.79)	1 (.008)	−.78
premodifying adverb + postmodifying prepositional phrase	1.96 (1.58)	1 (.008)	−1.57
premodifying adjective (in expression)	.08 (.0006)	1 (.008)	+.0074
premodifying noun	.04 (.0003)	1 (.008)	+.0077
coordinating conjunction + premodifying adjective + postmodifying clause	.04 (.0003)	1 (.008)	+.0077
noun is in plural	24.64 (19.87)	0 (-)	−19.87
noun is in plural + postmodifying prepositional phrase	2.85 (2.3)	0 (-)	−2.3
premodifying intensifying adverb	2.35 (1.9)	0 (-)	−1.9
postmodifying past participial clause	.98 (.79)	0 (-)	−.79
noun is in plural + premodifying intensifying adverb	.86 (.69)	0 (-)	−.69
noun is in plural + postmodifying finite clause	.66 (.53)	0 (-)	−.53
superlative premodifying adjective + postmodifying prepositional phrase	.55 (.44)	0 (-)	−.44
superlative premodifying adjective	.5 (.4)	0 (-)	−.4
noun is in plural + postmodifying past participial clause	.46 (.37)	0 (-)	−.37
Totals	124 (100)	124 (100)	–

to retrieve frequency information about the different variation criteria with reasonable speed and accuracy. Accordingly, the baseline data for the present study were taken from the British component of the International Corpus of English (ICE-GB), a small but fully syntactically annotated corpus of spoken and written British English. However, comparing ICE-GB and the BNC poses a potential problem since the register distributions of these two corpora differ drastically, which may have a huge impact on the (frequencies of the) variations found. In ICE-GB, 60% are spoken and 40% are written data; the BNC, on the contrary, comprises only 10% spoken language. In order to enhance the compatibility of the two corpora, it is therefore reasonable to compile a (sufficiently large) sample from the smaller corpus (here: ICE-GB) that reflects the register distribution of the larger corpus (here: the BNC) as accurately as possible. To that end, ICE-GB was searched for V NP-sequences (with optional adjectival or adverbial modifiers intervening; allowing the presence of any kind of determiner/no determiner) in all their syntactic forms. 2,295 (1,046 written and 1,249 spoken) sentences were retrieved. The 1,046 written sentences were taken over into the sample to represent 90% of the total baseline sample, and another 105 sentences were randomly selected from the spoken sentences to represent the remaining 10%, so the overall baseline sample comprised 1,151 V NP-constructions.

A third modification of Barkema's original measure was motivated by the desire to have an overall flexibility index value for each variation parameter that is easier to interpret than the rather complex flexibility profiles as shown in Table 3, and that makes it easier to compare different V NP-constructions. The basic idea for collapsing the flexibility profile information into one single index value was that small deviations from the baseline have only little effect on the overall flexibility value, whereas large deviations have a considerable impact. Consider the V NP-construction *foot* DET *bill* and its morphological flexibility in terms of the idiomatic variation parameter Tense, which comprises four levels: 'past', 'present', 'future', and 'nonfinite'. The 109 occurrences of *foot* DET *bill* are distributed across the parameter levels of Tense as follows: 10 times as 'past', 45 times as 'present', 9 times as 'future', and 45 times as 'nonfinite'. However, if *foot* DET *bill* actually behaved like the baseline with regard to Tense, we would expect 28 occurrences in the past tense, 68 in the present tense, 2 in the future tense, and 11 occurrences of the verb as a nonfinite form. Accordingly, the differences between observed and expected frequencies amount to −16.80% for the parameter level 'past', −20.66 for 'present', 6.69 for 'future', and 30.77 for 'nonfinite'. In order to combine these values and also weight them adequately, they are squared and then added: small deviations will contribute only little to the overall value, while big deviations will contribute much more. The overall sum of squared deviations (SSD) for Tense for *foot* DET *bill* is 1700.952, with the parameter level 'nonfinite' contributing 946.904 to this overall value. 'Future', on the contrary, which deviates only 6.69% from the baseline, contributes only 44.8 to the overall value. Note also that the weighting of the parameter levels for each parameter is always determined item-specifically: for another construction like *do* DET *trick*, the parameter level 'nonfinite' is apparently much less important, because the deviation only amounts to -1.48, and correspondingly, it contributes only little to *do* DET *trick*'s overall Tense flexibility value. Since the resulting overall SSD values can take on any value higher than 0, they may be difficult to interpret if one wants to compare values of different parameters or parameter levels for one construction, or compare different constructions with respect to one variation parameter or parameter levels. Therefore, a normalized version of these values (NSSD) is reported here. That is, for every variation parameter, an index is created by distributing the values on a scale between 0 (representing small deviations from the baseline) and 1 (representing large deviations form the baseline).[7] While space does not merit a detailed discussion of the results (cf. Wulff 2008), consider Table 7 in the appendix for an overview.

5. Results I: How idiomatic variation parameters cluster

In order to address the question how much and what kind of internal structure can be detected in the corpus data at large, the 20 idiomatic variation parameters, i.e. the 18 formal flexibility parameters, compositionality (Comp), and the corpus frequency (CorpFreq) of the constructions, were subjected to a Principal Components Analysis (PCA). A PCA is geared towards detecting structure within a set of variables and groups them into over-arching principal components (cf. Bortz 2005: 511–564).

The PCA grouped the 20 idiomatic variation parameters into 8 principal components; cf. Table 3. The leftmost column provides the eigenvalues[8] for each component; the third column shows how much of the total variance of the data the components account for. The respective columns to the right of these provide cumulative numbers.

Table 3. *Principal components identified by the PCA*

Eigenvalue	Cum. eigenvalue	% variance	Cum. %
4.224	4.224	21.121	21.121
2.882	7.106	14.41	35.53
2.12	9.226	10.599	46.129
1.901	11.127	9.505	55.635
1.357	12.484	6.786	62.421
1.241	13.726	6.207	68.628
1.168	14.894	5.842	74.47
1.009	15.903	5.047	79.517

Summing up, the first four components are by far the most important ones: they have an information value that is equivalent to more than half of the original variables (their cumulative *eigenvalue* amounts to 11.127), and they account for more than half of the total variance in the data (55.635%). Components 5, 6, 7, and 8 also yield eigenvalues higher than 1, yet they add much less explanatory power than the first four – they only increase the cumulative eigenvalue by about the size of one original variable. Overall, about forty percent of the number of original variables account for 74.47% of the total variance, a solid result that testifies to a considerable amount of distributional cohesiveness.

A look at the component loadings of each variation parameter on each principal component in Table 4 reveals which of the variables are grouped together (please cf. Table 6 for an explanation of the variable names). Component loadings higher than .7 (absolute values) are highlighted in bold print, since these actually constitute the component. For those parameters which did not yield values higher than .7, their highest component loadings are italicized to indicate on which component they load most highly, even if their contribution is not significant.

Table 4. *Component loadings of idiomatic variation parameters according to the PCA*

Variable	1	2	3	4	5	6	7	8
SF	**.901**	−.087	.069	.2	.098	−.009	−.106	.002
MF_Person	*.434*	.326	−.394	.15	−.022	.328	.277	.241
MF_NumV	−.082	**.943**	.077	.081	.006	.107	.019	−.069
MF_Tense	.173	.388	−.063	−.136	.007	*.47*	.46	−.268
MF_Aspect	.31	.158	.259	.128	.142	−.29	*−.543*	.391
MF_Mood	.04	**.95**	.098	.008	.08	.032	.029	−.026
MF_Voice	**.894**	.03	.074	.218	.121	−.092	−.064	.079
MF_Neg	.064	−.104	−.032	.016	−.055	−.035	−.001	**.939**
MF_Det	*.625*	.046	.157	−.187	−.196	.304	−.219	−.055
MF_NumNP	−.129	.431	−.224	.398	−.103	*.486*	.073	.003
MF_Gerund	−.012	.057	−.012	−.008	**.907**	.084	.005	−.071
LF_Addition	.2	.009	**.739**	.09	−.367	.175	−.04	−.01
LF_AttrAdj	.137	−.078	−.176	.118	−.083	.201	**−.795**	−.09
LF_AttrNP	*.611*	−.071	.146	.377	−.274	.08	.116	.11
LF_PP	.152	.147	.172	.049	.143	**.836**	−.152	−.042
LF_RelCl	−.145	−.152	.064	*−.583*	−.01	.525	.026	.009
LF_NoAdv	.103	.151	**.926**	.059	.12	−.01	.062	−.043
LF_KindAdv	.067	−.005	*.648*	.195	.548	.036	.136	.14
Comp	−.128	−.026	−.082	**−.925**	−.121	.013	.081	−.054
CorpFreq	−.29	−.026	−.182	**−.801**	.074	−.119	.136	.003

According to Table 4, the most important principal component 1 (important here in the sense that it accounts for more variance than any other component) comprises the idiomatic variation parameters tree-syntactic flexibility (SF) and the morphological flexibility parameter Voice (Mf_Voice) – in other words, if one wants to describe the overall distribution of the V NP-constructions, these two parameters which provide most information for this description. This result should actually be interpreted with caution since passives are obligatorily marked in English both at the syntactic level in terms of constituent order and at the morphological level in terms of the voice of the verb involved, so this result does not come as a surprise. The morphological flexibility parameters MF_Det and MF_Person, as well as the lexico-syntactic flexibility parameter LF_AttrNP, have their highest loadings on this principal component.

The second most important component according to the PCA comprises the morphological flexibility parameters MF_NumV and MF_Mood; the third component comprises two lexico-syntactic flexibility parameters, namely LF_Addition and LF_NoAdv. Note how the factor loading for LF_KindAdv of .648 only marginally misses the threshold value of .7, so one could consider it part of this principal component as well. The fourth component, which is the last of the top most important components, comprises compositionality (Comp) and corpus frequency (CorpFreq) – so the strong correlation between these two parameters is actually higher than any correlation of either parameter with any other one among the whole set of 20 variation parameters. Components 5, 6, 7, and 8 each comprise only one single variation parameter, in order of decreasing importance: presence or absence of a gerund (MF_Gerund), presence of modifying prepositional phrases (LF_PP; the tense of the verb (MF_Tense) and the number of the noun phrase (MF_NumNP) load highest on this component, too), the number of modifying attributive adjectives (LF_AttrAdj; the aspect of the verb (MF_Aspect) has its highest loading on this component, too), and the absence or presence of negation (MF_Neg).

In sum, the PCA by and large stands in accord with previous monofactorial studies. For one, the relevance of tree-syntactic flexibility is supported (cf. e.g. Gibbs and Gonzales 1985, Abeillé 1995); the same holds for the acknowledged connection between the ability of adverbial modification and idiomatic variation (cf. Gibbs et al. 1989, Nicolas 1995). Although compositionality and corpus frequency do not turn out to be the most decisive variation parameters, the centrality assigned to both parameters throughout the literature is also captured by the multifactorial analysis. The prominence of the morphological flexibility parameters NumV and Mood, however, has not been acknowledged in previous studies.

Next to distinguishing relevant from irrelevant parameters, the PCA goes beyond previous studies by quantifying the (relative) relevance of each parameter. From a multifactorial perspective, the relevance of tree-syntactic flexibility can hardly be underrated. Moreover, the results stand at odds with the widely-held claim that compositionality is the most important variation parameter – while other studies have shown that speakers can be motivated to zoom in on this parameter, the present multifactorial results indicate only a limited explanatory power with respect to the idiomatic variation continuum. However, we have to interpret this result with caution since it cannot be ruled out that the corpus-linguistic definition of compositionality is just not optimal yet.

Interestingly enough, the fact that corpus frequency loads high on principal component 4 speaks to a substantial influence of frequency, yet by far not an exclusive one – once more fine-grained classification schemes as provided by the variation parameters are considered, a much more complex picture emerges.

So which pieces of this rather complex picture do speakers take into consideration when they judge the overall idiomaticity of a V NP-construction?

6. Results II: Bringing together corpus and experimental data

In order to address this question, a multiple regression analysis was computed: the independent variables were the values of the idiomatic variation parameters for each V NP-construction, and the dependent variable was the corresponding average normed idiomaticity judgment. Taking all parameters into account, nearly 80% of the variance in the average idiomaticity judgments is accounted for, a highly significant result that testifies to a solid relationship between the parameters and the judgments (adjusted R^2=.565, p=.005**).[9] More specifically, the regression analysis provides so-called beta weights[10] for all parameters; the closer an idiomatic variation parameter is to 1, the more important (in the sense of covering variance) it is. Parameters with beta weights +.22 can be considered relevant because they account for 5% of the variance. Consider Table 5 for an overview.

As Table 5 shows, the most important variation parameters are the morphological flexibility parameters MF_NumV and MF_Mood, followed by two lexico-syntactic flexibility parameters, LF_KindAdv and LF_NoAdv. Next in line are compositionality and tree-syntactic flexibility. The morphological flexibility parameters MF_Voice and MF_Neg also yield sufficiently high beta weights to be considered relevant. The last variation parameter with a value higher than +.22 is the lexico-syntactic flexibility parameter LF_Addition (.265). Corpus frequency (CorpFreq) obtains a beta weight of .209.

Comparing these results with those of the PCA, the fit is striking: if we look at the top nine parameters speakers rely on according to the multiple regression and check where they occur in the PCA, we find that all of them also form the most important principal components: MF_NumV and MF_Mood formed one principal component; the lexico-syntactic flexibility parameters formed another, tree-syntactic-flexibility and MF_Voice formed yet another one. According to the regression analysis, speakers furthermore rely on the morphological flexibility parameter MF_Neg – which is one of the parameters which constitute a principal component of their own. And corpus frequency, which the PCA identifies as a component alongside compositionality, just about misses the threshold level of +.22.

Not only do the parameters correlating most strongly with the overall idiomaticity judgments coincide with the principal components, what is more, they single out those parameters which form the <u>most important</u> components

Table 5. *Beta weights for variation parameters as determined by a multiple regression of corpus and judgment data*

Variation parameter	Absolute beta weight
MF_NumV	.757
MF_Mood	.695
LF_KindAdv	.651
LF_NoAdv	.632
Comp	.578
SF	.573
MF_Voice	.351
MF_Neg	.275
LF_Addition	.265
CorpFreq	.209
MF_Person	.197
MF_Gerund	.16
MF_Tense	.125
MF_NumNP	.109
LF_AttrNP	.083
MF_Det	.055
MF_Aspect	.046
LF_PP	.043
LF_RelCl	.038
LF_AttrAdj	.032

(the only exception here is MF_Neg, which, according to the PCA, is only the 8^{th} most important component): MF_NumV and MF_Mood form the second component, the adverbial flexibility parameters the third, and tree-syntactic flexibility and Voice are the parameters comprising component 1. At the same time, none of the parameters which the PCA did not identify as belonging to either component is present among those parameters that the regression considers relevant.

Last but not least, looking at the ordering of the parameters in the regression analysis, we find that those parameters are closest to each other which also form components according to the PCA – alternatively, it would have been possible that, say, both MF_NumV and MF_Mood obtain high beta weights, but that they are not back to back in the ranking.

In sum, the extraordinary overlap between the results of the PCA and the multiple regression analysis suggests that the parameters most relevant for the judgments are those which have the highest information value in terms of the amount of variance they cover. This is most evident in the morphological flexibility parameters MF_NumV and MF_Mood obtaining the top ranks – a result that, while unexpected from previous research, fits this kind of interpretation well. At the same time, the results stand in opposition to the widely held assumption that compositionality is the most decisive parameter contributing to idiomaticity. However, the results are compatible with the hypothesis that the frequency with which speakers are exposed to a given parameter (level) plays a major role: most of the parameters and parameter levels which neither the PCA nor the regression consider relevant have low average values across the V NP-constructions and low baseline values, e.g. LF_AttrNP (average V NP-construction value=.032; baseline value=.147), LF_RelCl (.033/.149), PP (.111/.13), and LF_AttrAdj (.208/.227; cf. Table 7). While tree-syntactic flexibility, MF_Voice, and MF_Neg, on the other hand, also yield only low average V NP-construction values, their baseline values are among the highest – which could explain how this parameter gains salience for speakers' judgments.

More generally, it appears that parameters relating to the verb-slot on the whole rank higher than those related to the noun phrase-slot of the V NP-constructions; actually, none of the noun-related parameters yields a beta weight higher than +.22. This suggests that speakers strongly focus on the verb-slot when judging the idiomaticity of the V NP-construction.

7. Conclusions: Towards a new model of idiomaticity

Methodology-wise, the present study has gone beyond previous studies in

- presenting a multifactorial approach that weighs the relative influence of each variation parameter, thereby doing justice to the assumption that cognitive processes are complex in nature;
- presenting maximally objective and bottom-up parameter definitions which do not fall back on any given classification scheme which is not itself empirically founded, such as different classes of compositionality;
- integrating theoretical assumptions of usage-based Construction Grammar, such as the existence of scalar categories, differences in the cognitive entrenchment of variation parameters and their parameter levels, and an active interplay between lexical and constructional semantics.

With respect to the theoretical implications of the present study, the results tie in well with many widely-established claims about idiomaticity, with tree-syntactic flexibility being a key characteristic. Likewise, the central role of compositionality is reproduced by the Principal Components Analysis. However, the multifactorial perspective reveals that, when being considered *in toto*, other variation parameters turn out to be even more important than syntactic flexibility and compositionality, foremost aspects of morphological and lexico-syntactic flexibility.

With regard to the morphological parameters, one could argue that this result is a statistical artefact because morphological marking is obligatorily present. However, this does not necessarily entail that they are not of any psychological relevance with regard to perceived idiomaticity. Neither would this objection explain why only certain morphological parameters, and specifically those of the verb, are so much more important than others. Thirdly, according to Newman and Rice's (2005) *Inflectional Island Hypothesis*, verbs can be strongly associated with particular inflected forms, just like lemmas can exhibit strong preferences for a restricted set of argument structures and lexical meanings. Since verbs with extremely biased inflectional profiles are particularly susceptible to grammaticalization, this may explain speakers' (subconscious) sensitivity towards morphological parameters.

Similarly, the results suggest that speakers are paying attention to the lexico-syntactic flexibility of V NP-constructions, in particular that of the verb slot. While previous studies have pointed towards the association between lexico-syntactic flexibility and idiomaticity, the high ranking of these parameters found here ascribes them a much more prominent role. As Elizabeth Traugott pointed out (p.c.), speakers' focus on aspects concerning the adverbial modification potential is easily motivated from the perspective of language change: one of the first properties constructions tend to lose during grammaticalization processes is their adverbial flexibility.

At a more general level, the fact that speakers are obviously sensitive to correlational clusters of variation parameters as modeled by the PCA particularly supports a usage-based approach that assumes only little grammatical hardware in the sense that speakers apply given grammatical concepts or categories to the data in order to make sense of them; rather, they build their categories in a bottom-up fashion, the primary task being to cover as much variance in their input as possible.

Given the findings of the present study, a major issue is how to implement this probabilistic and complex information into existing schematic models of the mental lexicon. As outlined above, the constructicon is mainly specified with regard to what I would like to call the vertical axis; the primary process creating diversification along this axis beyond the level of one-word constructions is schematization. To a certain extent, idiomatization can be conceptualized as being diametrically opposed to delexicalization: the more idiomatic a construction is, that is, the more formally fixed and non-compositional, the less likely it is that this construction will delexicalize. In other words, on a continuum of idiomatic phrases ranging from collocations to idioms, the more idiomatic the phrase, the less delexicalization potential it has. These probabilistic tendencies aside, however, it is desirable to find a form of representation that also captures constructions which are relatively idiomatic and, at the same time, delexicalized, such as intensity expressions like *burning ambition/burningly ambitious/to burn with ambition* (cf. Zeschel 2007).

Accordingly, I propose to extend the constructicon by adding another dimension, a horizontal axis as shown in Figure 2. This horizontal axis cuts across the range of the vertical axis where fully lexically specified complex constructions are located. More precisely, one can think of the constructicon as bifurcating at the level of complex constructions, opening a quadrant space in which constructions can be positioned according to their degree of schematization and idiomaticity. The closer a phrasal construction is located on the horizontal axis to the vertical axis, the more compositional and formally less constrained it is (e.g. *write a letter*); the more formally frozen and semantically opaque a construction is (such as *take the plunge*), the further away from the vertical axis that construction is positioned on the horizontal axis.

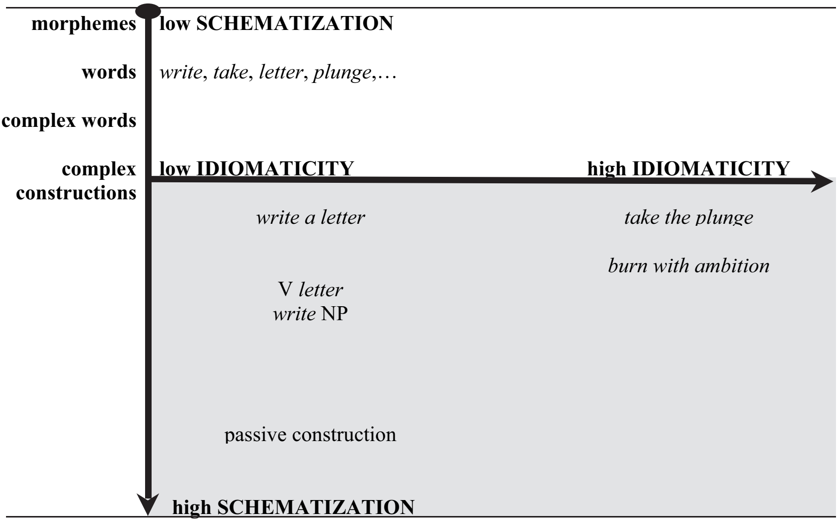

Figure 2. *Extended schematic representation of the constructicon*

The horizontal axis itself has multiple layers (which are not depicted in Figure 2), one representing each variation parameter. These layers form clusters that, at a level of coarse granularity, can be likened to principal components, while at a finer level of granularity, they can be likened to individual parameters. For instance, one layer represents information about aspects of morphological flexibility in terms of the number of the verb, another that of the mood of the verb – since they correlate highly, they form a cluster. Each V NP-construction is represented once on each of these (clusters of) layers, i.e. it is assigned a value on the morphological flexibility layers, the lexico-syntactic layers, the compositionality layer, etc.

The higher the overall idiomaticity of a construction, the less its representation is connected to its constituting lexical constructions. For instance, *take the plunge* is both semantically highly non-compositional and formally restricted, so its overall idiomaticity is high; accordingly, *take the plunge* is only weakly connected with the lexical representation of *take* and *plunge*. *Write a letter*, on the other hand, ranges relatively low on all idiomatic variation parameters, and so it is more strongly connected with the lexical representations of *write* and *letter* further up the vertical axis – and consequently, it is also more strongly connected with other lexical constructions that are associated with *write* and *letter*, such as *type*/*compose* or *email*/*paper*, which in turn makes *write a letter* a likely candidate for subsequent schematizations.

To conclude, by adding the horizontal axis to the constructicon, the results of the present study can be represented, and the representation also stands in line with established findings from previous studies:

- the horizontal axis basically represents the idiom-collocation continuum, specifying the (slightly misleading) term *idiom* on the vertical axis;
- conceptualizing the vertical axis as multilayered, the fact that idiomaticity is a multifactorial and scalar concept is represented;
- by assigning each construction a value on each of these layers, item-specific differences in the weightings of the different variation parameters are represented;
- the distributional similarity of the different variation parameters can be represented via the distance between the different layers;
- the construction-specific representation accommodates the fact that speakers are able to judge constructions according to individual variation parameters;
- the item-specific, multilayered, and therefore storage-redundant representation accommodates Langacker's (1987) *Rule-List Fallacy* and therefore accords with other contemporary models of linguistic representation (e.g. Pierrehumbert's (2003) *exemplar theory*).

Despite these achievements, this study is certainly not the final answer to an adequate characterization of idiomaticity. A number of limitations need to be pointed out, and several caveats remain. Strictly speaking, the fact that the performance-based variation parameters correlate highly with the idiomaticity judgments surely points towards their relevance – yet neither does it follow from the present study that idiomaticity is <u>exclusively</u> based upon performance data, nor can we conclude that the parameters subjects rely on are necessarily also <u>grammatically represented</u>. Alternatively, at least some of the parameters speakers rely on could be *ad hoc* created categories (Barsalou 1992) that are stipulated by the performance data because they cover so much of the variance that speakers attempt to systematize; these categories do not need to have any grammatical status (although it is a plausible working assumption). In order to shed more light on the question which of the idiomatic variation parameters are actually grammatically represented, a greater variety of constructions has to be considered. Moreover, further experimental studies should focus on individual variation parameters; for instance, in the questionnaire used in the present study, all V NP-constructions were presented in their most typical morphological forms, so it would be interesting to see if similarly high correlations with morphological flexibility can be observed if the morphological contexts (or that of any other variation parameter, for that matter) are controlled for and/or varied systematically.

All in all, however, this study strongly suggests that a complex and intuitive phenomenon like idiomaticity can be modeled on the basis of performance data, thereby providing further evidence for a mutual interplay between grammar and usage and the relevance of studies of authentic language data. In doing so, it forms a queue with the increasing number of recent studies (cf., among others, Gries et al. 2005; Kepser and Reis 2005; Arppe and Järvikivi 2007) that demonstrate the vast potential that resides in combining quantitative corpus-linguistic and experimental methods for the study of language.

Appendices

Table 6. *An overview of the formal flexibility parameters and their parameter levels with corresponding examples*

Parameter	Parameter levels	Example	Level
SF	declarative active	She *told the story*.	nominal
	declarative passive	The story *has been told* countless times.	nominal
	relative cl. active	Mary told *the story* she *had written*.	nominal
	particle cl. active	There are *many* more *stories to tell*.	nominal
	relative cl. passive	They agreed that this is a story that needs to be told.	nominal
	particle cl. passive	Inevitably there is a little *story to be told* about this.	nominal
	interrogative active	I mean, *didn't* you *have a good story to tell*?	nominal
	imperative active	*Don't tell me that story*!	nominal
	interrogative passive	*Can a* new *story be told* here?	nominal
	imperative passive	Ruth heard Adam *teeth grit* at the memory.	nominal
LF_ADD	any absent/present	It is a story never to be told *in full*.	nominal
LF_ATTRADJ	attr. adjective for NP	He would tell the *rudest* stories out loud	interval
LF_ATTRNP	attr. noun for NP	I will tell you a *horror* story.	interval
LF_PP	post-mod PP for NP	The runner has gone into it and told the story *of the battle*.	interval

Parameter	Parameter levels	Example	Level
LF_RelCl	post-mod. relative cl.	A story was told *whose smear value demands immediate publication*.	interval
LF_NoAdv	adverb(ial)s	Now it is important for me to tell the story *correctly*.	interval
LF_KindAdv	no adverb(ial)	I will tell you a story.	nominal
	adverb modifying N	I could tell *quite* a story.	nominal
	space adverbial	People tell me stories *on the doorsteps*.	nominal
	time adverbial	The story has *often* been told.	nominal
	process adverbial	The story is told *in seven episodes each covering a day*.	nominal
	respect adverbial	He told them a story *about his own children when they were very young*.	nominal
	contingency adverbial	I tell this story *for two reasons*.	nominal
	modality adverbial	She *really* is telling this story.	nominal
	degree adverbial	. . . tell the story of the Poison Feast from the Drachenfels novel *in full* . . .	nominal
MF_Person	infinitive	In the city nobody was allowed *to tell* stories.	nominal
	first	*I* will *tell* you a funny story about that.	nominal
	second	Well, *you* sometimes *tell* stories for fun, don't you?	nominal
	third	She may not be brilliant but *she tells* a straightforward story.	nominal
	vocative	Come on, Dandelion, *tell* us a story.	nominal
	other	The act of *telling* one's life story is an encounter with reality.	nominal
MF_NumV	singular	But even *he* may not *tell* the whole story.	nominal
	plural	Cumulatively, these studies *are telling* a similar story.	nominal
	nonfinite	No, I am not here *to tell* you a story.	nominal
MF_Tense	past	She *told* us a terrible story.	nominal
	present	*Tell* your story based on the poem.	nominal
	future	I am *going to tell* you a story.	nominal
	nonfinite	He loved *telling* the story.	nominal
MF_Aspect	simple	Even the abrupt and anticlimatic conclusion *tells* a story.	nominal
	progressive	He gave me an odd look as if I *was telling* strange stories.	nominal
	perfective	I *have* also been *told* the same story time and time again.	nominal
MF_Mood	indicative	I *tell* this story for two reasons.	nominal
	subjunctive	We *ought to tell* the whole story to Zacco.	nominal
	nonfinite	This is the most difficult part of my story for me *to tell*.	nominal

Parameter	Parameter levels	Example	Level
MF_VOICE	active	I *told* him your cover story and he swallowed it quite happily.	nominal
	passive	Penny's story *was* bravely *told*.	nominal
MF_NEG	absent/present	But even he may *not* tell the whole story.	nominal
MF_DET	*the/my/no/what*	She tells *their* story well enough.	nominal
	some/any/enough	He told *some* funny stories.	nominal
	this/that	I never told my wife *this* story.	nominal
	these/those	Barbara must not be permitted to tell *these* stories.	nominal
	a[*n*]/*each*/*every*/[*n*]*either*	Daddy is going to tell you *a* bedtime story.	nominal
MF_NUMNP	singular	Then you can tell this *story* until you die, brother.	nominal
	plural	Then he told them terrible *stories* of his wild and criminal life at sea.	nominal
MF_GERUND	absent/present	I shall summarize … *by telling* a particular story about beavers.	nominal

Table 8. *Mean (normalized) sums of squared deviations from baseline ([N]SSD) for formal flexibility parameters*

	SF	LF_ADD	LF_ATTRADJ	LF_ATTRNP	LF_PP	LF_RELCL	LF_NOADV	LF_KINDADV	MF_PERSON	MF_NUMV	MF_TENSE	MF_ASPECT	MF_MOOD	MF_VOICE	MF_NEG	MF_DET	MF_NUMNP	MF_GERUND
bear DET *fruit*	.976	.272	.033	1.000	.470	.169	.316	.187	.256	.273	.321	.573	.324	1.000	.014	.719	.130	.001
beg DET *question*	.928	.004	.003	1.000	.227	.159	.245	.223	.258	.026	.122	.601	.002	.900	.035	.017	.003	.014
break DET *ground*	.909	.067	1.000	1.000	.895	.125	.038	.077	.157	.034	.074	.572	.006	.802	.033	.920	.153	.035
break DET *heart*	.945	.391	.073	1.000	.523	.169	.300	.163	.120	.031	.167	.270	.025	.912	.021	.071	.030	.015
call DET *police*	.216	.284	.068	.429	.909	.169	.162	.088	.068	.353	.424	.003	.030	.002	.021	.067	.983	.000
carry DET *weight*	.847	.000	.126	1.000	.253	.169	.066	.109	.220	.061	.046	.537	.042	.897	.015	.388	.153	.000
catch DET *eye*	.827	.336	.082	.955	.128	.157	.372	.180	.192	.014	.549	.491	.009	.718	.035	.990	.086	.025
change DET *hand*	1.000	.046	.115	1.000	.974	.169	.015	.175	.214	.082	.633	.177	.048	1.000	.022	1.000	1.000	.565
close DET *door*	.725	.044	.061	.009	.645	.154	.008	.041	.102	.071	.554	.286	.029	.738	.034	.140	.079	.855
cross DET *finger*	.649	.168	.115	1.000	.974	.169	.280	.150	1.000	1.000	.914	.593	1.000	.757	.038	.070	1.000	.169
cross DET *mind*	.958	.185	.095	1.000	.692	.127	.091	.156	.434	.086	1.000	.009	.019	1.000	.087	.119	.085	.023
deliver DET *good*	.966	.142	.090	.715	.902	.063	.032	.052	.175	.245	.246	.395	.191	1.000	.000	.143	1.000	.014
do DET *trick*	.923	.626	.115	1.000	.778	.169	.480	.237	.290	.006	.159	.324	.119	.930	.017	.990	.153	.023
draw DET *line*	.379	.001	.012	.030	.718	.082	.132	.295	.137	.170	.252	.047	.183	.085	.025	.009	.036	.127
fight DET *battle*	.906	.004	.297	.284	.668	.109	.022	.062	.152	.218	.344	.565	.177	1.000	.037	.259	.000	.003
fit DET *bill*	.934	.346	.107	1.000	.798	.169	.158	.147	.245	.007	.006	.691	.059	.906	.011	.104	.153	.011
follow DET *suit*	1.000	1.000	.108	1.000	.974	.125	1.000	.444	.264	.403	.421	.761	.300	1.000	.033	.987	.146	.003
foot DET *bill*	.940	.137	.034	.347	.909	.169	.994	.438	.249	.391	.234	.510	.261	1.000	.040	.907	.125	.044
get DET *act together*	1.000	.514	.095	1.000	.761	.169	.376	.204	.089	.139	.009	.330	.020	.814	.001	.157	.146	.007
grit DET *tooth*	.904	.501	.115	1.000	.974	.169	.270	.147	.147	.102	.571	.455	.107	1.000	.046	.148	1.000	.002
have DET *clue*	.835	.281	.064	1.000	.798	.097	.176	.230	.435	.021	.054	.739	.019	1.000	1.000	.525	.071	.016

Table 8. (cont.)

	SF	LF_ADD	LF_ATTRADJ	LF_ATTRNP	LF_PP	LF_RELCL	LF_NOADV	LF_KINDADV	MF_PERSON	MF_NUMV	MF_TENSE	MF_ASPECT	MF_MOOD	MF_VOICE	MF_NEG	MF_DET	MF_NUMNP	MF_GERUND
have DET *laugh*	.695	.139	.009	1.000	.586	.169	.275	.166	.218	.367	.149	.332	.159	1.000	.034	.413	.112	.289
hold DET *breath*	.943	.787	.100	.962	.956	.159	.696	.329	.075	.004	.476	.492	.016	.871	.008	.137	.146	.122
leave DET *mark*	.826	.002	.003	.851	.832	.169	.125	.251	.298	.082	.497	.004	.002	.925	.027	.031	.099	.001
make DET *face*	.886	.220	.005	1.000	.862	.169	.531	.250	.156	.068	.636	.617	.027	.985	.031	.617	.045	.133
make DET *headway*	.988	.000	.008	1.000	.974	.126	.129	.080	.161	.178	.422	.143	.169	.806	.008	.645	.153	.009
make DET *mark*	.463	.014	.013	1.000	.974	.141	.013	.053	.241	.394	.600	.188	.369	1.000	.028	.735	.148	.142
make DET *point*	1.000	.165	.037	1.000	.235	.079	.255	.128	.133	.046	.160	.178	.027	.484	.038	.010	.038	.004
meet DET *eye*	.941	.356	.017	1.000	.728	.048	.426	.209	.159	.121	.413	.555	.123	1.000	.012	.071	.467	.624
pave DET *way*	.558	.072	.111	1.000	.974	.169	.813	1.000	.169	.212	.357	.290	.296	.899	.046	.149	.153	1.000
play DET *game*	.978	.086	.006	.093	.399	.110	.227	.154	.114	.227	.186	.612	.192	.717	.000	.083	.021	.704
scratch DET *head*	.935	.611	.105	1.000	.974	.169	.454	.236	.242	.158	.604	.591	.178	1.000	.046	.985	.020	.275
see DET *point*	.607	.023	.115	1.000	.059	.011	.063	.284	.508	.005	.002	.819	.007	.980	.353	.064	.146	.017
take DET *course*	.897	.081	.013	1.000	.010	.015	.340	.179	.192	.218	.415	.101	.256	.642	.034	.012	.129	.045
take DET *piss*	.912	.016	.115	1.000	.974	.169	.070	.213	.121	.082	.104	1.000	.064	.910	.009	.152	.153	.682
take DET *plunge*	.978	.341	.077	1.000	.883	.169	.183	.100	.151	.282	.363	.286	.207	.952	.036	.955	.153	.150
take DET *root*	.520	.138	.115	1.000	.974	.169	.021	.094	.198	.062	.470	.060	.028	1.000	.004	1.000	.153	.002
tell DET *story*	.436	.003	.003	.335	.043	.015	.130	.082	.097	.063	.109	.216	.066	.479	.029	.009	.009	.072
write DET *letter*	.835	.000	.002	.453	1.000	1.000	.106	.075	.080	.035	.484	.047	.030	.259	.037	.243	.015	.078

Bionote

Stefanie Wulff received her Ph.D. in 2007 from the University of Bremen, Germany, with a corpus-linguistic dissertation on idiomaticity in English. A recent postdoctoral fellow at the English Language Institute at the University of Michigan and currently a lecturer in the Linguistics Department at the University of California at Santa Barbara, Stefanie's current research interests include variation at the syntax-lexis interface, second language acquisition, and academic discourse. E-mail: swulff@linguistics.ucsb.edu

Notes

* I would like to express my gratitude to (in alphabetical order) Nick C. Ellis, Stefan Th. Gries, and Ute Roemer for their comments on an earlier version of this paper. The usual disclaimers apply.

1. The construction types were primarily selected on the basis of the *Collins Cobuild Dictionary of Idioms* (2002). 262 V NP-constructions are listed in the dictionary, 33 of which occur more than 90 times in the BNC (this frequency threshold had to be met to license statistical evaluation). While the definition of idiom in this dictionary already captures a substantial part of the idiomaticity continuum, in order not to bias the sample towards one end of the continuum, this set was extended by another six constructions randomly selected from a concordance of all verb-noun phrase sequences in the BNC that occur more than 100 times.
2. DET stands for any kind of determiner, including a zero determiner.
3. While this is not a V NP-construction, it was included in several pre-tests and is reported alongside the V NP-constructions.
4. I am indebted to Ewa Dąbrowska and Philip Shaw, who made this part of the study possible by distributing the questionnaires for me.
5. The most typical contexts were identified as follows. For every V NP-pattern, an Excel sheet was created in which each attestation is represented in one row, and all variation parameters included in the present study (tree-syntactic flexibility and the different kinds of lexico-syntactic flexibility and morphological flexibility) are represented in one column each. The different parameter levels of each variation parameter are coded with numbers. For instance, the morphological flexibility factor Tense exhibits four parameter levels: "0" for past tense, "1" for present tense, "2" for future tense, and "3" for cases with nonfinite verb forms. Every attestation of a V NP-pattern was assigned one number for each variation parameter. If, for example, a V NP-construction preferably occurs in the present tense, this is reflected by the predominance of the number "1" in the corresponding column. So with respect to Tense, the most typical context for that V NP-pattern is one in present tense. After having identified the most frequent parameter levels for all variation parameters (that is, in all columns), that attestation of a V NP-pattern which unites variable-specific preferences to a maximum extent was taken to represent the V NP-pattern in its most typical context. One might object that maximal control over context effects is only achieved if the contexts are identical for all V NP-constructions, such that, say, the subject slot is always lexicalized as a personal pronoun, and the verb is always in the simple past tense. However, such a procedure does not do justice to the contextual preferences (or even restrictions in some cases) that the different V NP-constructions exhibit.
6. Theoretically, it would be reasonable to divide these sums by the number of component words (here: two) to obtain an index value between 0 and 1; however, for the results presented here, this was not done since the values were extremely small already. This is not a problem as long as only the results of one analysis are compared (since the ranking of the constructions remains the same); however, once results from several analyses are compared, it may be reasonable to do the division in order to stay within the index from 0 to 1.

7. There is one piece of information that is neither accessible on the basis of Barkema's original method nor the extension presented here: the directionality of the deviation. Is, say, *foot* DET *bill* generally more or less flexible with respect to Tense variation compared to the baseline? Here, we can only consider the directionality of the deviation for the different variation parameter levels, but at the level of the more general variation parameters, this information is lost due to the squaring process. Therefore, Wulff (2008) also presents another flexibility measure to complement the Barkema-extension which provides exactly this missing piece of information: relative entropy.
8. The eigenvalue of a component predicates how many of the original variables are represented by the component, so a component with an eigenvalue of 1 captures as much variance as one original variable.
9. While the adjusted R^2-value only amounts to .565, it has to be borne in mind that this value is lowered by the overall number of variables entering into the computation: the more variables are required to account for all the variance in the data, the lower the adjusted R^2 will be.
10. Beta weights quantify the contribution of each individual independent variable to the overall correlation observed.

References

Abeillé, Anne
1995 The flexibility of French idioms: A representation with lexicalised tree adjoining grammar. In Everaert, Markus, Erik-Jan van der Linden, André Schenk, and Rob Schreuder (eds.), *Idioms: Structural and Psychological Perspectives*. Hillsdale: Erlbaum, 15–42.

Arppe, Antti and Juhani Järvikivi
2007 Every method counts – combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory* 3 (2), 131–160.

Bannard, Colin
2005 Learning about the meaning of verb particle constructions from corpora. *Journal of Computer Speech and Language* 19 (4), 467–478.

Bannard, Colin, Timothy Baldwin, and Alex Lascarides
2003 A statistical approach to the semantics of verb-particles. *Proceedings of the ACL-Workshop on Multiword Expressions: Analysis, Acquisition, and Treatment*, 65–72.

Bard, Ellen G., Dan Robertson, and Antonella Sorace
1996 Magnitude estimation of linguistic acceptability. *Language* 72 (1), 32–68.

Barkema, Henk
1994a The idiomatic, syntactic and collocational characteristics of received NPs: Some basic statistics. *Hermes Journal of Linguistics* 13, 19–40.

Barkema, Henk
1994b Determining the syntactic flexibility of idioms. In Fries, Udo, Gunnel Tottie, and Peter Schneider (eds.), *Creating and Using English Language Corpora.* Amsterdam: Rodopi, 39–52.

Barsalou, Lawrence R.
1992 *Cognitive Psychology: An Overview for Cognitive Scientists*. Hillsdale, NJ: Erlbaum.

Berry-Rogghe, Godelieve L. M.
1974 Automatic identification of phrasal verbs. In Mitchell, John Lawrence (ed.), *Computers in the Humanities*. Edinburgh: Edinburgh University Press, 16–26.

Bortz, Jürgen
2005 *Statistik für Human-und Sozialwissenschaftler*, 6th ed. Heidelberg: Springer.
Cacciari, Cristina and Patricia Tabossi
1988 The comprehension of idioms. *Journal of Memory and Language* 27, 668–683.
Collins Cobuild Dictionary of Idioms
2002 2nd ed. London: Harper Collins.
Croft, William and D. Alan Cruse
2004 *Cognitive Linguistics*. Cambridge: Cambridge University Press.
Cronbach, Lee J.
1951 Coefficient alpha and the internal structure of tests. *Psychometrika* 16 (3), 297–334.
Fernando, Chitra
1996 *Idioms and Idiomaticity*. Oxford: Oxford University Press.
Gibbs, Raymond W. and Gayle P. Gonzales
1985 Syntactic frozenness in processing and remembering idioms. *Cognition* 20 (3), 243–259.
Gibbs, Raymond W. and Nandini Nayak
1989 Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology* 21 (1), 100–138.
Gibbs, Raymond W., Nandini Nayak, John Bolton, and Melissa Keppel
1989 Speakers' assumptions about the lexical flexibility of idioms. *Memory and Cognition* 17 (1), 58–68.
Glucksberg, Sam
1993 Idiom meanings and allusional content. In Cacciari, Cristina and Patrizia Tabossi (eds.), *Idioms: Processing, Structure, and Interpretation*. Hillsdale, NJ: Erlbaum, 3–26.
Goldberg, Adele E.
1995 *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.
Goldberg, Adele E.
2006 *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
Gries, Stefan Th., Beate Hampe, and Doris Schönefeld
2005 Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics* 16 (4), 635–676.
Kepser, Stephan and Marga Reis (eds.)
2005 *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives*. Berlin/New York: Mouton de Gruyter.
Langacker, Ronald W.
1987 *Foundations of Cognitive Grammar. Volume I: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.
Lin, Dekang
1999 Automatic identification of noncompositional phrases. *Proceedings of the 37th Annual Meeting of the ACL*, 317–324.
McCarthy, Diana, Bill Keller, and John Carroll
2003 Detecting a continuum of compositionality in phrasal verbs. *Proceedings of the ACL-Workshop on Multiword Expressions: Analysis, Acquisition, and Treatment*, 73–80.
Newman, John and Sally Rice
2005 Inflectional islands. Paper presented at the International Cognitive Linguistics Conference 9, July 2005, Seoul.

Nicolas, Tim
1995 Semantics of idiom modification. In Everaert, Markus, Erik-Jan van der Linden, André Schenk, and Rob Schreuder (eds.), *Idioms: Structural and Psychological Perspectives*. Hillsdale, NJ: Lawrence Erlbaum, 233–252.

Nunberg, Geoffrey, Ivan A. Sag, and Thomas Wasow
1994 Idioms. *Language* 70 (3), 491–538.

Peterson, Robert R. and Curt Burgess
1993 Syntactic and semantic processing during idiom comprehension: Neurolinguistic and psycholinguistic dissociations. In Cacciari, Cristina and Patrizia Tabossi (eds.), *Idioms: Processing, Structure, and Interpretation*. Hilldale, NJ: Erlbaum, 201–225.

Pierrehumbert, Janet B.
2003 Probabilistic phonology: Discrimination and robustness. In Bod, Rens, Jennifer Hay, and Stefanie Jannedy (eds.), *Probabilistic Linguistics*. Cambridge, MA: MIT Press, 177–228.

Schone, Patrick and Daniel Jurafsky
2001 Is knowledge-free induction of multiword unit dictionary headwords a solved problem? *Proceedings of the 6th Conference on Empirical Methods in Natural Language Processing*, 100–108.

Stefanowitsch, Anatol and Stefan Th. Gries
2003 Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8 (2), 209–243.

Titone, Debra A. and Cynthia M. Connine
1994 The comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20 (5), 1126–1138.

Wulff, Stefanie
2008 *Rethinking Idiomaticity: A Usage-Based Approach*. London/New York: Continuum.

Zeschel, Arne
2007 Delexicalization patterns: A corpus-based approach to incipient productivity in fixed expressions. Unpublished doctoral dissertation, University of Bremen.