Absolute Constructions: On the Distribution of Predicative Idioms

Susanne Z. Riehemann & Emily Bender Stanford University

1 Introduction

Many syntactic theories, following Chomsky (1985), strive for a kind of modularity where statements of grammar (rules, constraints, or principles) refer only to general grammatical items (e.g., features or configurations) and the constructions discussed by traditional grammarians are considered epiphenomena. This paper argues for an alternative conception of grammar in which constructions have primary ontological status on the basis of the partially restricted distribution of certain predicative idioms (see Kay & Fillmore 1999, Goldberg 1995, Zwicky 1994, and Sag 1997 for the notion of construction). Specifically, we will be looking at the with and with-less absolute constructions (Stump 1985, McCawley 1983).

In the remainder of this paper, $\S 2$ presents the data, $\S 3$ and $\S 4$ make the case for a constructional analysis of the data, $\S 5$ formalizes the analysis in HPSG and $\S 6$ applies this analysis to the more nuanced data of a range of individual grammars.

2 Data

The particular data we are concerned with in this paper is presented in (1) and (2). (1) gives examples of each of our four predicative idioms (in italics) in the *with* absolute construction. *With* absolutes are sentence modifi-

^{*}This paper is based upon work supported by the NSF under grant number IRI-9612682. This paper has benefited from conversations with and comments from Ivan Sag, Tom Wasow, Arnold Zwicky, Dan Flickinger, and Paul Kay, and especially from extensive discussions with Ann Copestake. We remain absolutely responsible for any errors in fact or analysis.

^{©1999} S. Riehemann & E. Bender. WCCFL 18 Proceedings, S. Bird, A. Carnie, J. Haugen, P. Norquest (eds.). pp. 476-489. Somerville, MA: Cascadilla Press.

ers. In this paper we are concerned with the case where they occur sentence-initially. Internally, with absolutes consist of the lexical item with followed by a small clause of the form NP + predicative XP. In the examples in (1), the predicative idioms head the predicate of the small clause.

- (1) a. With the negotiators still *poles apart* on so many issues, it's hard to see how these talks will ever end.
 - b. With expectations *flying high*, the Bulls have to win the championship this time.
 - c. With the media *all ears*, Clinton was very careful about what he said.
 - d. With peace talks *old hat*, it's hard to get a sense of hopefulness in the Middle East these days.

The examples in (2) are parallel to the examples in (1) except that they involve the *with*-less absolute construction.

- (2) a. The negotiators still *poles apart* on so many issues, it's hard to see how these talks will ever end.
 - b. Expectations *flying high*, the Bulls have to win the championship this time.
 - c.? The media all ears, Clinton was very careful about what he said.
 - d.*Peace talks *old hat*, it's hard to get a sense of hopefulness in the Middle East these days.

In contrast to the *with* absolute examples, not all of the idioms are acceptable in the *with*-less absolute construction. This pattern of judgments is summarized in Table $1.^2$

	with absolute construction	with-less absolute construction
poles apart	ok	ok
flying high	ok	ok
$all\ ears$	ok	?
$old\ hat$	ok	*

TABLE 1 Contrast patterns based on 14 speakers.

The following two sections argue that this data motivates a constructional analysis. §3 first makes the argument that the absolute construction

¹Here we take the class of predicatives to be those phrases that can follow *be*, with the understanding that this distributional definition should line up with a cross-linguistically applicable semantic/pragmatic notion.

applicable semantic/pragmatic notion.

²There is a considerable degree of variation in this domain. §6 describes how the pattern shown in Table 1 was derived from the judgments of 14 speakers.

must be analyzed as a construction. §4 provides (further) motivation for a constructional analysis of the idioms.

3 Constructions

In this section, we will focus on the with-less absolute construction, which has no overt lexical content uniquely associated with it. Nonetheless, it is a pairing of form (a small clause which is a sentence modifier) and meaning (its semantic and pragmatic properties). We see two possible ways to capture this pairing: a grammatical construction or a null element of some sort. Hantson (1992) develops a null complementizer analysis of with-less absolutes. Here, the null element \emptyset in (3) is syntactically parallel to with, which Hantson takes to be a complementizer.³

(3) There he sat, $[S' \ \emptyset \ [S]$ his back against the hot stones of the tower.]] In general, there is a certain formal equivalence between null elements and constructions. It is difficult to imagine a paradigm that could be described with one but not the other, as long as one is solely interested in generating the right strings. However, approaches based on null elements and those based on constructions do differ in the kinds of generalizations they can capture elegantly. Here we will argue that the distribution of predicative idioms across the two types of absolute constructions allows us to distinguish between the two approaches.

All of the idioms are acceptable in the *with* absolute, while only some are acceptable in the *with*-less absolute. On the null complementizer analysis, this data would have to be handled in terms of subcategorization of the null complementizer. On the constructional analysis, they can be handled in terms of subtyping of the constructions (elaborated in §5 below).

Subcategorization is implausible because no other complementizers (or elements that select for clauses) are selective about lexical material in those clauses. Here we mean lexical as opposed to grammatical—a complementizer could indeed select for clauses with a certain mood, where the mood is expressed on the verb. However, this is different from selecting for specific open class words. It seems to us that idioms are more like open class words in this respect than they are like grammatical properties such as mood.

On the other hand, Nunberg et al. (1994:516) mention several other idioms which are selective about which constructions they co-occur with. For example, idioms such as *Is the Pope Catholic?* only occur as polar questions and idioms such as *Believe you me!* only occur as imperatives. Thus while clause-selecting heads never care about the lexical content of

³Example (3) is adapted from Hantson 1992:89. This particular example is of a clause-final absolute. Hantson also discusses clause-initial absolutes and in fact draws no distinction, but does not happen to supply any examples of clause-initial with-less absolutes.

those clauses, idioms are known to be selective about their syntactic context.

Further, an analysis in terms of subtyping allows us to specify exactly where these restricted idioms appear—instead of having to say where they do not. For every place one might expect an idiom to occur and it does not, the null complementizer analysis requires finding a selecting head and fixing its subcategorization so that it excludes that idiom. On a constructional analysis, the subtyping mechanism only requires dealing with the contexts in which the idiom does occur.⁴

However, if it was the case that the acceptability of these idioms turned on some semantic feature, then an analysis in terms of subcategorization (of a null complementizer or of a construction) would be more appealing. To test this possibility, we conducted a second survey with a separate group of 19 native speakers. These speakers were presented with the sentences in (1) and (2) plus parallel sentences with paraphrases for the idioms. The paraphrase sentences for *poles apart*, *flying high*, and *all ears* are as in (4).

- (4) a. (With) the negotiators still far apart on so many issues, it's hard to see how these talks will ever end.
 - b. (With) the Lakers so successful, LA fans are optimistic about the playoffs.
 - c. (With) the media *intensely alert*, Clinton was very careful about what he said.

The idiom old hat was harder to find a paraphrase for, perhaps because it may be undergoing a change in meaning. (Webster's gives two meanings, 'old-fashioned' and 'lacking in freshness: trite', but a corpus⁵ search turned up many examples which were not consistent with either.) A preliminary survey indicated that a meaning like 'commonplace' was the most current. We changed the test sentences to ones in which that meaning would be plausible and asked participants at the end of the survey what they thought the best paraphrase was.⁶ The test sentences we used for old hat were:

- (5) a. (With) email and webbrowsers *old hat*, it's hard to remember what life was like before the Internet.
 - b. (With) email and webbrowsers *commonplace*, it's hard to remember what life was like before the Internet.

⁴Subtyping is not the only mechanism available in this approach. For an idiom that has the general distribution of a VP, for example, it would suffice to give it features like those of a VP.

⁵The corpus we use throughout is the North American News Text Corpus from the LDC, http://morph.ldc.upenn.edu/Catalog/LDC95T21.html

⁶The choices we offered were *boring*, *outdated*, *old fashioned*, *nothing new*, *nothing extraordinary*, *commonplace*, and *thoroughly familiar*. We also asked if there was some better paraphrase not on the list.

The results of this second survey were as follows. First, commonplace was judged to be the best paraphrase of old hat in these sentences by most of the participants and at least a possible paraphrase by most of the rest. Next, we turn to the distribution of the idioms and their paraphrases. There were 12 speakers who did not systematically reject the with-less absolute. Of these, eight accepted old hat in the with absolute but preferred commonplace to old hat in the with-less absolute.

Further, there were similar results for the idiom all ears. Because this idiom was generally more acceptable in the absolutes than old hat the numbers are smaller, but 5 speakers (of the 12) preferred intensely alert to all ears in the with-less absolute.

In all, only one speaker showed a pattern which would be consistent with a semantic subcategorization explanation for old hat. That is, he accepted the with-less absolute for other idioms and accepted both old hat and commonplace in the with absolute, but rejected both old hat and commonplace in the with-less absolute. However, this speaker was one of the 5 to get a contrast between all ears and intensely alert.

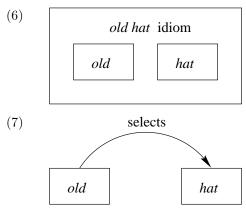
If the paraphrases we chose actually shared the relevant semantic properties of the idioms, then we can take these results to show that the restricted distribution of predicative idioms across the two types of absolutes is not a matter of the two absolute constructions having different semantic restrictions. We conclude that the null complementizer approach, which relies on subcategorization, does not provide a natural account of the pattern of grammaticality shown in Table 1. In the next section we will discuss an alternative, construction based account.

4 Idioms

In general, idioms can be analyzed as special phrases (constructional approach) or as special words with specific selectional restrictions (lexical approach). The subtyping approach that we advocate (and describe in §5) requires a constructional approach to the idioms involved. This section provides a brief discussion of the motivation for the constructional approach to idioms. For a thorough discussion, see Riehemann 1997, in preparation.

The figures in (6) and (7) schematize the two approaches to idioms. On the constructional approach (6), the *old hat* idiom is a big phrase that contains somewhere within it the words *old* and *hat*. The meaning of the idiom is associated with the construction. On the lexical approach (7), there is a special word *old* that selects for the (special) word *hat*. Here the meaning of the idiom is associated with the selecting word (old).

⁷14 of the 19 speakers chose *commonplace* as the best paraphrase (possibly tied with some others) or said it was an acceptable one if they picked something else. 10 of the 12 speakers we discuss below are in this category.



One problem with a lexical approach is that only one of the words which are part of the idiom is constrained to co-occur with the others—and there is no mechanism for controlling the distribution of the others. For example, if old hat is analyzed as a special idiomatic lexical entry for old that only modifies the idiomatic word hat, then there is no way of preventing the idiomatic hat from occurring without old. Note also that this approach requires the selecting word to be made special, even when it is actually has the literal meaning, like old in old hat or bad in bad blood.⁸

As noted above, the constructional approach to idioms allows for the possibility of constructional licensing of restricted idioms. Of interest here, then, is the distribution of these idioms in a corpus across constructional and lexically selected contexts. By constructional here we mean where the licenser appears to be a construction (such as the *with*-less absolute construction) rather than a lexical item (such as *is* in *That's old hat*).

If constructionally licensed occurrences of idioms were rare, one might be tempted to write them off as peripheral, i.e., as instances where for whatever reason of performance or telegraphic register the licensing lexical item was omitted. However, constructional occurrences of idioms are not rare—in our corpus we found that 22% of the occurrences of our four idioms do not involve a form of the verb be. Some of these involve other verbs like seem, but 9% of the overall occurrences are constructional. Further, since we are dealing with written and edited data, these instances cannot be explained away as slips of the tongue. Also, constructional occurrences are not restricted to telegraphic contexts such as headlines.

Note that not all constructional representations of idioms are complete syntactic trees. Many idioms can occur discontinuously, as in (8).

(8) My music career was not *flying* all that *high* and I was tired of being on the road.

⁸For further arguments against a lexical approach see Pulman 1993.

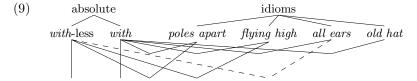
Our corpus contains 14 occurrences of non-contiguous flying high (e.g., flying so high (that)). Other idioms are even more syntactically flexible. We will describe in §5 how we handle these non-contiguous occurrences by specifying the semantic relation between idiomatic words.

In this and the preceding sections, we have argued that the distribution of predicative idioms is most satisfactorily analyzed in terms of constructions. In the following section, we give the details of our construction-based analysis and, in particular, describe the mechanism of subtyping.

5 HPSG Analysis

Before we can make this constructional approach formally precise we have to give a little bit of background about Head-Driven Phrase Structure Grammar. HPSG views grammar as a system of signs—pairings of phonological and syntactic form and semantic and pragmatic meaning. These signs are modeled with feature structures, or attribute-value matrices (AVMs). Tree structure is encoded in the AVMs by means of daughter attributes. Some of the syntactic signs (such as the head-complement construction) are very general, while others have much more idiosyncratic characteristics. The signs in HPSG are organized into a multiple inheritance hierarchy, where higher types express broad generalizations and lower types more specific details. This idea was first used to organize lexical types (Flickinger 1987) and later extended to phrasal types (Sag 1997).

An illustration of how the formal device of subtyping captures the restricted distribution of these idioms is given in the partial type hierarchy in (9). Each node in the hierarchy is a construction type. The solid lines connect actually existing types. Only the types at the bottom of the hierarchy license grammatical sentences. Dashed lines indicate marginal types. (The dashed lines are not part of the HPSG formalism but rather a placeholder for a theory of marginality judgments.)

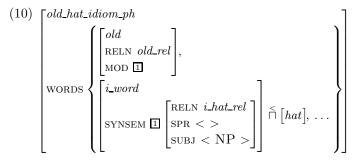


So, for example, this hierarchy expresses that the idiom *poles apart* can occur in the *with* and *with*-less absolute constructions, while the idiom *old hat* only has a mutual subtype with the *with* absolute construction—and therefore cannot occur in the *with*-less absolute. We are employing the closed-world assumption here. This means that only the types actually declared in the hierarchy exist. Since we declare no common subtype for *old hat* and the *with*-less absolute, that combination is not licensed. Note

that the hierarchy displayed here is partial. All of the idioms have other subtypes, such as the one that allows old hat to occur with forms of be. The absolute constructions each also have an unrestricted subtype. This subtype is the most common one for each construction. It combines freely with syntactically compatible phrases in the sense that those phrases can unify with one of its daughters. Old hat cannot combine with either absolute construction in this way because it has no subtype that is just old hat.

We now show how idioms can be represented constructionally. This is possible because our framework allows us to represent to varying degrees of specificity how the pieces of a phrase fit together (syntactically or semantically). That is crucial because some idioms do not have to be contiguous, as discussed above.

The feature structure in (10) gives a partial description of the construction that licenses the *old hat* idiom. This analysis of idioms is based on that in Riehemann 1997, which was developed to account for the syntactic variability of many idioms.



Legend:

 $\stackrel{\widehat{}}{\sqcap}$ skeptical default unification primary semantic contribution

i_word idiomatic word

Intuitively, when this construction is involved in the derivation of a sentence, all of the specified words must occur lower within that sentence. This is the function of the WORDS feature. However, just calling for the word hat would lead to the wrong meaning, as the usual mechanisms of grammar would include the usual meaning of hat in the semantic representation of the sentence. So we need a way to state that hat does not have its usual meaning but an idiomatic one. This is done formally by using skeptical default unification of the kind proposed in (Carpenter 1993) and extended to typed feature structures in (Lascarides & Copestake 1999). That is, the idiomatic word (i_word) on the left side of the f symbol is just like the literal word on the right, except for the properties that are explicitly

changed. In this case those are the following: the meaning (which we mark as $i(diomatic)_hat_rel$), and the syntactic property of taking an NP subject instead of a specifier. Note that this idiomatic word is a characteristic of the idiom construction and has no independent lexical entry.

The construction also specifies how the idiomatic words fit together syntactically, i.e. that *old* modifies *hat*. This can be done to a greater or lesser extent depending on the variability of the idiom.

In this particular idiom we analyzed only *hat* as semantically idiomatic and assumed that *old* has its literal meaning and is not an *i_word*. This may not be the right interpretation for all speakers, and we could easily represent both words as idiomatic. We also have a way of representing idioms whose meanings cannot be distributed to their parts (see Riehemann 1997).

The feature structure in (11) describes the with absolute construction. It is a subtype of the absolute construction and, from it, inherits semantic and pragmatic information. More precisely, it inherits all the information it shares with the with-less absolute construction. This construction specifies constituent structure by means of its DTR (daughter) features. The HEAD-DTR is the lexical item with and the COMP-DTR is a predicative small clause.

The feature structure in (12) is the subtype of the *with*-absolute construction (11) and the idiom *old hat* (10), as can be seen in (13).

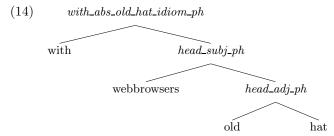
(12)
$$\begin{bmatrix} with_abs_old_hat_idiom_ph \\ \text{COMP-DTRS} \left\langle \begin{bmatrix} \text{RELN } i_hat_rel \end{bmatrix} \right\rangle \end{bmatrix}$$

$$(13) \quad with_absolute_ph \qquad old_hat_idiom_ph$$

$$with_abs_old_hat_idiom_ph$$

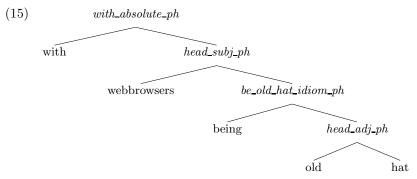
Since the type old_hat_idiom_ph phrase does not specify any constituent structure, and since all nodes in the parse tree which dominate both old and hat are consistent with the constraints on old_hat_idiom_ph, it is not im-

mediately obvious which node(s) in the tree are of this type. In fact, there is only one, as shown in (14). This is the only node which is compatible with all of the constraints on one of the subtypes of old_hat_idiom_ph.



Note that we cannot specify on with_abs_old_hat_idiom_ph that the COMPDTR is of type old_hat_idiom_ph. This is because in order to have that old_hat_idiom_ph be just old hat and not be old hat or with NP old hat, there would have to be a stand alone subtype of old_hat_idiom_ph. This would nullify all of the advantages of the subtyping analysis.

The type old_hat_idiom_ph has another subtype, be_old_hat_idiom_ph. Phrases licensed by this subtype are free to combine with the with absolute as part of its complement daughter. In this case, only the node labeled be_old_hat_idiom_ph in (15) is an instance of the type old_hat_idiom_ph.9



Given that the *old_hat* idiom is involved at the same level in the tree as the *with* absolute construction in (14), we need to specify that the words *old hat* show up as the predicate of the small clause inside this construction, and not any further down the tree. That is, we do not want the matrix *with* absolute to license *old hat* in (16).

(16)*With John thinking, "This issue old hat, I'd better move on," things are certainly going to get worse.

⁹At first glance, it looks like the node labeled *with_absolute_ph* might be compatible with the constraints on *with_abs_old_hat_idiom_ph*, but the RELN of the COMP-DTR is contributed by *being* rather than by (idiomatic) *hat*.

On the other hand, it would not do to say that the predicate of the small clause is exactly the two words *old hat*, since some modification is allowed:

(17) With these issues already old hat, we'll have to look for some more topics for position papers.

Our solution is to have the type with_old_hat_idiom_ph specify that the primary semantic contribution of the complement daughter is that of the idiom old hat. This works because the feature RELN always points to the semantic contribution of the syntactic head of a phrase. The RELN feature is shared between the head daughter and the mother in all headed constructions. In (17), old hat is the head of the head-modifier phrase already old hat, so the RELN gets passed up from old hat. In (16), the small-clause in the absolute is John thinking ... move on, which is headed by thinking, so it has a different value for RELN.

To summarize, here is how our analysis captures the important properties of our data. Both the absolute constructions and the idioms are represented phrasally. This has all the advantages discussed above, including naturally capturing the fact that idiomatic words occur only as part of the idiom and cannot have those idiomatic meanings when they occur alone.

Given a phrasal representation for both the idioms and the constructions, the distribution of the idioms can be restricted by only allowing them to occur in certain environments. This is formally expressed by cross-classifying them only with some constructions but not others.

6 Individual Systems

In the first survey, we asked 21 native speakers (8 linguists and 13 non-linguists) for their judgments on the sentences in (1) and (2) plus some others, presented in random order. 7 speakers systematically rejected the with-less absolutes. 10 speakers systematically accepted the with absolutes. 9 speakers neither systematically rejected nor systematically accepted either of these constructions. Even with this small number of idioms studied, only 2 speakers were systematic for both of these constructions—it is quite likely that they might not be as systematic with other idioms.

The data presented in §2 was based on the statistical mode of the patterns given for each idiom by those 14 informants who did not systematically reject the *with*-less absolutes. For example, the contrast for *old hat* as presented in (1d) vs. (2d) was the most frequent pattern given. Not everyone had the same contrasts, but 19 speakers (=90%) allowed some combi-

 $^{^{10}}$ For readers familiar with MRS (Copestake et al. 1998): RELN corresponds to KEY.

nations of these idioms and syntactic constructions, and not others. A constructional analysis is necessary for all these speakers, although it may not be identical to the one presented here.

Here are two examples of individual speakers' systems:

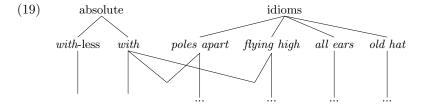
	Speaker 1		Speaker 2	
	with absolute construction	with-less absolute construction	$\begin{array}{c} with \\ \text{absolute} \\ \text{construction} \end{array}$	with-less absolute construction
$poles\ apart$	ok	*	ok	ok
flying high	ok	*	ok	ok
$all\ ears$	*	*	ok	ok
$old\ hat$	*	*	ok	*

TABLE 2 Contrast patterns for two speakers.

The first speaker does not have the contrast for *old hat* that we discussed above—he finds this idiom ungrammatical in both constructions. But an analysis like the one we presented is needed to capture the contrast he has for *poles apart* and *flying high*, since this speaker did accept the *with*-less absolute, just not with any of these idioms. It is also needed for *all ears* and *old hat* because these were accepted by this speaker in other constructions, like as a pre-nominal modifier:

(18) We'll just get more old hat conspiracy theory journalism.

The hierarchy representing this speaker's system would be:¹¹



The second speaker shows basically the pattern we discussed in §2.

7 Conclusion

We have shown that the distribution of predicative idioms motivates a formal notion of construction for its analysis in a generative grammar. This approach handles the unpredictable distribution of these idioms by specify-

 $^{^{11}\}mathrm{The}$ '...' in this figure indicate subtypes of these idioms other than those which involve the absolutes.

ing the environments in which they can occur without missing the generalization that the SAME idiom occurs in all these contexts. Because in this approach the idiomatic instances of words do not have an existence outside of the idiomatic phrases, the problem of having to restrict them to these environments does not even arise.

Note that we are not claiming that all idioms have to be treated this way. Some idioms may be totally permissive about the constructions they occur with, and for others, the restrictions might be explained semantically. However, some idioms require the kind of approach we describe. This is not surprising, since psycholinguistic evidence shows that speakers process canonical forms of idioms faster (McGlone et al. 1994). This suggests that speakers have representations for specific combinations of idioms and constructions, in addition to general knowledge of the idioms.

References

- Carpenter, B. 1993. Skeptical and Credulous Default Unification with Applications to Templates and Inheritance. Inheritance, Defaults and the Lexicon, eds. E. J. Briscoe, A. Copestake, & V. de Paiva, 13–37, Cambridge University Press.
- Chomsky, N. 1985. Knowledge of Language: Its Nature, Origin, and Use, New York: Praeger.
- Copestake, A., D. Flickinger, C. Pollard, & I. A. Sag. 1998. Minimal Recursion Semantics: An Introduction. ftp://ftp-csli.stanford.edu/linguistics/sag/mrs.ps.gz
- Flickinger, D. 1987. Lexical Rules in the Hierarchical Lexicon. PhD Thesis, Stanford University.
- Goldberg, A. E. 1995. Constructions: A Construction Grammar Approach to Argument Structure. University of Chicago Press.
- Kay, P. & C. J. Fillmore. 1999. Grammatical Constructions and Linguistic Generalizations: The 'What's X doing Y?' Construction. *Language* 75:1–33.
- Hantson, A. 1992. Case Assignment and Be-Deletion in Non-Finite Clauses: With Special Reference to (Absolute) Free Adjuncts. Belgian Journal of Linguistics 7:75–94.
- Lascarides, A. & A. Copestake 1999. Default Representation in Constraint-based Frameworks. *Computational Linquistics* 25:55–105.
- McCawley, J. D. 1983. What's with with? Language 59:271–287.
- McGlone, M. S., S. Glucksberg, & C. Cacciari. 1994. Semantic Productivity and Idiom Comprehension. *Discourse Processes* 17:167–190.
- Nunberg, G., I. A. Sag, & T. Wasow. 1994. Idioms. *Language* 70:491–538.
- Pulman, S. G. 1993. The Recognition and Interpretation of Idioms. *Idioms—Processing, Structure, and Interpretation*, eds. C. Cacciari et al., 249–270, Lawrence Erlbaum.
- Riehemann, S. 1997. Idiomatic Constructions in HPSG. Paper presented at the Cornell HPSG Workshop. http://doors.stanford.edu/~sr/idioms.ps
- Riehemann, S. in preparation. A Constructional Approach to Idioms and Derivational Morphology. PhD Thesis, Stanford University.

- Sag, I. A. 1997. English Relative Clause Constructions. $\it Journal~of~Linguistics~33:431-483.$
- Stump, G. T. 1985. The Semantic Variability of Absolute Constructions. Dordrecht: Reidel.
- Zwicky, A. M. 1994. Dealing out Meaning: Fundamentals of Syntactic Constructions. *Proceedings of BLS 20*, 611–625.