A unified lexicon and grammar? Compositional and non-compositional phrases in the lexicon

Neal Snider and Inbal Arnon

Abstract

In this chapter, we address the debate between single-system and dual-system models of language by looking at the processing of multi-word phrases. We present findings that challenge the distinction between 'stored' and 'computed' linguistic forms via two experiments. The first demonstrates parallels in the processing of words and phrases: frequent four-word phrases are processed more quickly than less frequent ones, without any evidence of a frequency threshold. The second experiment shows that idiomatic phrases prime their construction just as well as non-idiomatic phrases, suggesting that they are not stored as unanalyzed wholes, but instead have internal structure. Taken together, the findings undermine the empirical criteria traditionally used to distinguish between 'stored' and 'computed' forms: compositional phrases showed frequency effects, even though such effects are often thought to be a marker of lexical storage, while non-compositional forms (idioms) showed evidence of internal structure, unexpected if they are stored as unanalyzed wholes. The findings show that linguistic structures are processed in qualitatively the same way regardless of where they fall on the frequency and compositionality continua, and highlight the utility of models that deal with all linguistic experience in a qualitatively similar fashion, and allow for experience to influence the learning, representation and processing of all linguistic patterns.

1. Introduction

There has been long-standing tension in the study of language between approaches that assume a clear distinction between the mental lexicon and grammar (dual-system theories, Chomsky 1965, 1995; Fodor 1983; Pinker 1991, 1999; Pinker and Prince 1988; Ullman 2001, 2004) and ones that do not (single-system theories, Bates and MacWhinney 1989; Elman 1991; MacDonald, Pearlmutter and Seidenberg 1994; Rummelhart and

McClelland 1986; Seidenberg 1994). Dual-system models distinguish between the mental lexicon – an inventory of memorized forms, and the mental grammar – the rules or constraints used to combine the memorized elements. This distinction echoes the one made in many generative models of language (Chomsky 1981; Jackendoff 2002; Kaplan and Bresnan 1982; Pollard and Sag 1994). The mental lexicon is thought to contain the linguistic units that cannot be derived: simple words (e.g. *cat*), morphemes, irregular nouns and verbs, and longer non-compositional phrases like idioms. The most clearly articulated model of this kind is that of Pinker and his colleagues (Pinker 1991; Pinker and Prince 1988; Pinker and Ullman 2002). They propose that the two components of language (lexicon and grammar) are learned differently, involve different cognitive abilities and are governed by different neural substrates (Ullman et al. 2005; Ullman 2001).

In dual-system models, forms created by grammar are distinct from those originating in the lexicon. No such distinction is posited by single-system theories. Instead, all aspects of language depend on one computational system. The same cognitive mechanism processes all linguistic experience, whether a non-compositional lexical item like '*cat*' or a compositional phrase like '*I don't know*'. Word-object mappings and grammatical rules are learned in a qualitatively similar fashion – by abstracting and generalizing from linguistic experience.

A growing number of models implicitly or explicitly take a single-system stand. In connectionist models, the unity of lexicon and grammar is made explicit by using one single network to capture all linguistic experience (e.g., Rumelhart and McClelland 1986; Seidenberg 1994). Exemplar models of language also dispense with the distinction but in a different way, by having linguistic units and categories correspond to clusters of memory traces (Bod 1998, 2006; Goldinger 1996; Johnson 1997; Pierrehumbert 2001). Connectionist and exemplar models differ in several important respects, including the use of symbolic or non-symbolic representations and the implementation of higher level categories like nouns or verbs (see Bybee and McClelland 2005). But both dispense with any clear distinction between 'stored' and 'computed' forms and instead assume that all linguistic experience is learned, processed and used in a similar fashion.

These models are closely related to what are often labeled usage-based approaches to language where grammatical knowledge emerges from linguistic experience (Bybee 1998, 2006; Goldberg 2006; Barlow and Kemmer 2000; Langacker 1986, 1987; Tomasello 2003). The lexicon is not 'reserved' for atomic elements. There is no a priori limit on the size of the units that

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 128)

are stored; as long as they can be attended to and remembered, they can be of varying length (word, two-word, multi-word phrase) and levels of abstraction: from single words, through partially realized constructions to fully abstracted ones (*give*, *give me a break*, *give NP a break*, *give NP NP*, *V NP NP*).

1.1. The representational status of multi-word phrases

The contrast between dual-system and single-system approaches has been studied primarily in the domain of morphological representation and processing (Rumelhart and McClelland 1986; Pinker and Prince 1988). But given their diverging assumptions about language, these approaches make different predictions about many aspects of language use. In this chapter, we contrast the two approaches by looking at a relatively less-studied domain: the processing of larger units of language – multi-word phrases. We use this expression to refer to multi-word sequences that are syntactic constituents (e.g. *don't have to worry*, but not *in the middle of the*).

The ways the two approaches handle words are clearly articulated in existing models (e.g., Rumelhart and McClelland 1986; Pinker 1991). Their predictions about larger units are not clearly stated in any existing model but can be extrapolated from their general assumptions about language. Just as they differentiate between regular and irregular morphological forms, dual-system models maintain a distinction between compositional phrases (like *don't have to worry*) and non-compositional ones (as in idioms like *gave the surfer the creeps*). Compositional phrases are generated by the grammar while non-compositional ones originate in the lexicon and are stored together with their idiosyncratic syntactic and semantic features. Idioms should have the characteristics of stored forms while compositional phrases should not. Because compositional phrases can be derived in a predictable way, there is no need to store them in the lexicon. In fact, given the goal of minimizing storage (e.g., Pinker, 1991), compositional multi-word phrases would seem unlikely candidates for storage in the lexicon.

In contrast, single-system models do not posit such a distinction. Multi-word phrases, whether compositional or not, should be like any other linguistic pattern. Every encounter with a phrase is predicted to add to its representation and influence future processing. Compositional and non-compositional phrases should be impacted by the same factors (e.g., frequency) that impact the processing of both bare and regularly inflected words. The two kinds of phrases should also be processed in a qualitatively similar fashion: to the extent that compositional and non-compositional phrases share structural and lexical features, they should be processed in the same way. For example, hearing a compositional dative phrase like *give the man a hammer* makes one more likely to re-use the double object construction in future dative uses (e.g. Bock 1986; we expand on syntactic priming in section 3). If non-compositional phrases involve similar syntactic processes, then hearing a non-compositional phrase like *give the man a lift*, which has the same head verb and can also alternate, should also increase the likelihood of using the double object construction.

The extent to which compositional multi-word phrases are part of the mental lexicon, and the extent to which their status can be distinguished from that of non-compositional phrases, has an important role in evaluating models of language. In this chapter, we show that is it hard to differentiate compositional and non-compositional phrases empirically, and in doing this; we argue against the distinction (posited in dual-system models) between 'stored' and 'computed' forms more generally.

We do this in two ways: first, we demonstrate that compositional phrases (like *don't have to worry*) exhibit phrase-frequency effects similar to those found for words. Such a finding shows that speakers are sensitive to the frequency of a range of units (including ones that are 'computed' under dual-system models). It also undermines the empirical distinction between stored and computed forms: generated forms display frequency effects thought to be a mark of lexical storage under dual-system models (e.g. Ullman and Wellensky 2005). Second, we show that idiomatic and non-idiomatic datives prime their syntactic construction to a similar degree. Such a finding again blurs the distinction between compositional and non-compositional phrases: forms that are considered to be 'stored' (e.g. idioms) maintain internal structure and activate their constructions just like 'computed' forms. Together, these findings reveal similarities between 'stored' and 'computed' forms and undermine the possibility of coming up with empirical criteria to distinguish the two.

In Section 2 we report on studies showing that speakers are sensitive to the frequency of four-word compositional phrases. In Section 3 we present a novel study showing that idiomatic and non-idiomatic datives prime to a similar degree. In Section 4 we discuss these results in light of the contrast between dual- and single-system models of language.

2. Comprehenders are sensitive to the frequency of compositional phrases

In this section, we report a series of experiments published in Arnon and Snider (2010) showing that people process more frequent 4-word phrases faster than less frequent ones. This effect occurs across the frequency continuum, with no evidence of a threshold or cutoff. But first we take up the role frequency plays in single-system and dual-system models.

2.1. Frequency effects and mental representation

Frequency plays a very different role in single and dual-system models. In single-system models, frequency – as an approximation of experience – plays a central role in the emergence and entrenchment of linguistic units. The more often a pattern is experienced, the easier it becomes to access and use (Bybee 2006; Bod et al. 2003; Bybee and Hopper 2001). Single-system models differ in the specific mechanisms they use to explain the processing advantage of more frequent forms (by impacting the weights in a connectionist network; by lowering the threshold of activation in spreading activation networks; or by enhancing the activation of a memory trace in exemplar models). But they share a common belief that frequency effects inform us about the units that speakers attend to, and predict that frequency effects should be found for all linguistic units: simple and complex.

Frequency effects are viewed differently in dual-system models. The role of frequency in language representation and use is rarely discussed explicitly in these models (e.g. Pinker 1999). This absence echoes the traditional view in generative linguistics that frequency effects are irrelevant to the study of language because they reflect real-life probabilities or performance issues that are separate from, and immaterial to, linguistic knowledge (Chomsky 1957; recently re-argued for by Newmeyer 2003). In some models frequency effects are relegated to the mental lexicon (Ullman and Wellensky 2005). This allows them to account for the widespread frequency effects found in word production and comprehension (see Monsell 1991) while maintaining that 'stored' elements should exhibit frequency effects but 'computed' elements should not.

2.2. Lessons from morphology

Frequency effects have been used to contrast single-system and dual-system models of regular and irregular inflected forms (e.g. *walked* vs. *felt*). Dual-system models predict that irregular forms will be stored in the mental lexicon while regular forms will be generated by the grammar (Marcus et al. 1992; Pinker 1991, 1999; Pinker and Prince 1991; Pinker and Ullman 2002; Ullman et al. 1997). Single-system models predict that all forms will be represented by the same associative memory mechanism (Rumelhart and McClelland 1986; Plunkett and Marchman 1991, 1993; Marchman 1993).

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 131)

If regularly inflected forms cannot be accessed as whole words, then the base form (e.g., *walk*) should be activated every time an inflected form is encountered. Access speed should reflect the frequency of the base in all its various inflections (e.g. walks, walking, etc.). If a whole-word representation is available, then the frequency of the inflected form should also affect access speed. Finding that the frequency of the inflected form is predictive of processing time suggests a whole-form representation is available, as argued by single-system, but not dual-system, models. Indeed, the frequency of the inflected form itself (*walked*) predicts processing latencies when the frequency of the base form (*walk*) and the inflectional morphemes (*-ed*) is controlled for (e.g., Alegre and Gordon 1999; Baayen et al. 1997; Taft 1979).

A similar whole-form frequency manipulation has been extended to the study of phrases in child language (Bannard and Matthews 2008). Two and three-year-olds are faster and more accurate at repeating higher frequency phrases compared to lower frequency ones when part frequency is controlled for (e.g. *a drink of tea* vs. *a drink of milk*). Children are sensitive to phrase-frequency. This in turn suggests that they represent whole phrases at some level, just as in whole-word representation of regularly inflected words.

2.3. Phrase-frequency effects

In a series of studies we used a manipulation similar to that used by Bannard and Matthews (2008) to look at the processing of compositional phrases in adults (Arnon and Snider 2010). We wanted to see (a) whether adults are sensitive to phrase-frequency, and (b) whether this holds not only for very frequent phrases, but whenever a higher-frequency phrase is compared to a lower-frequency one. Language-users should be sensitive to phrase-frequency according to single-system, but not dual-system, models. We undertook the latter analysis to test the predictions of a slightly modified dual-system model that allowed very frequent phrases to be stored in the lexicon. Very frequent forms have privileged status also in specific usage-based models (e.g., Goldberg 2006). Therefore, asking whether there is a threshold for phrase-frequency effects has implications for those models as well (see Arnon and Snider 2010 for a further discussion).

2.3.1. Previous research

Many studies have shown that two-word (bigram) frequency affects processing: words are faster to process (McDonald and Shilcock 2005; Reali and

Christiansen 2007) and shorter to produce (Bell et al. 2003, 2009; Gregory et al. 2004; Jurafsky et al. 2001) when they appear as part of a more frequent bigram. People keep track of co-occurrence patterns for single words, but capturing such relations doesn't require any representation beyond the single word. Few studies have looked beyond the bigram, and most of those have focused on the processing of highly frequent phrases. For instance, Bybee and Scheibman (1999) found that *don't* was phonetically reduced in the frequently recurring phrase *I don't know*. Bell et al. (2003) likewise found that the ten most frequent words in English are phonetically reduced when they are more predictable given the previous and following word. Bannard and Matthews (2008) showed that children are sensitive to phrase-frequency but their frequent items were also taken from the top third of the frequency range.

A few other studies have looked at frequency beyond the bigram for a broader frequency range. Levy and Jaeger (2007) found an effect of predictability, given the previous two words, on relativizer omission in English relative clauses. Speakers were more likely to omit the relativizer when it was more predictable given the last one, two, and three words of the pre-relative clause utterance, but because they do not report the independent effect of each string size (this was not the goal of their paper), we cannot know whether their results show an effect of three-word frequency when bigram and unigram frequency are controlled for. Underwood, Schmitt, and Galpin (2004) used eye-tracking to look at participants' eye-movements while reading formulaic sequences of up to six words (e.g., *as a matter of fact*). They found fewer fixations when words appeared in formulaic sequences, which they interpreted as evidence that people represent the sequences as a whole. But since they did not control for the frequency of the substrings either, or for the plausibility of each phrase (plausibility isn't controlled also in Bannard & Matthews, 2008), it is hard to know how to interpret their results.

These effects provide limited evidence that adults are sensitive to the frequency of compositional phrases. We need more evidence from adults, with part frequency and plausibility controlled for, and from phrases across the frequency continuum.

2.3.2. Our findings

We conducted two reaction times studies where we compared processing latencies for pairs of compositional four-word phrases that differed in phrase frequency (the frequency of the four-word phrase) but were matched

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 133)

for part frequency (unigram, bigram, and trigram frequency), and for plausibility relative to the event they describe (e.g. *don't have to worry* vs. *don't have to wait*). We measured processing latencies using a phrasal decision task. People saw four-word phrases and had to judge whether they were possible in English. We used this task for two reasons. First, lexical decision tasks are often used in the study of morphologically complex words (e.g., Baayen et al. 1997). Since we are using a similar frequency manipulation (varying the frequency of the whole form vs. the parts), we wanted to use a similar task. Second, the task allows for the presentation of the phrase as a whole and encourages participants to attend to each phrase as a unit. We controlled for the frequency of the sub-strings by comparing phrases that differed only on the final word, and by controlling for the final word, the bigram, and the trigram, both in the item selection and in the statistical analysis of the results. We also controlled for the plausibility of the events depicted by the phrases using a norming study.

The two experiments together looked at phrases in three frequency bins, in order to test the effect of frequency across the spectrum. The High frequency bin compared phrases that occurred above ten times per million in the corpus, with those that occurred below ten per million. The Mid frequency bin compared phrases between five and ten per million with those below 5 per million. The Low frequency bin compared phrases between one and five per million with those below 1 per million. The items were constructed using a 20-million word corpus that consisted of the Switchboard (Godfrey, Holliman, and McDaniel 1992) and Fisher (11,699 recorded telephone conversations in American English, 18 million words; Cieri, Miller, and Walker 2004) corpora. In each bin, the high and the low variant differed in phrase-frequency but were matched on all other measures, including plausibility. Table 1 gives example items from the different bins, together with their phrase-frequency.

49 Stanford students were paid to complete the two experiments. All were native English speakers. Each participant saw one four-word phrase on the screen at a time and had to decide (as quickly as possible) whether they were possible sequences in English while their response time was measured. The experiments had an equal number of possible and impossible sequences (fillers). During a practice phase, *I saw the man* was given as an example of a possible sequence, and *I saw man the* and *jump during the pool* as impossible sequences.

We analyzed the data using mixed-model linear regression. As predicted, higher-frequency phrases were decided on faster than lower frequency phrases in all three bins. We then took the responses from all three bins

Table 1. Mean frequency (per million words) and example items in the three bins (N = number of items).

High bin ($N = 16$) (High: 19.48, Low: 3.61)		Mid bin ($N = 12$) (High: 9.75, Low: 0.75)		Low bin ($N = 17$) (High: 3.5, Low: 0.2)	
Don't have to worry	15.3	It takes a lot	7.35	Don't have any money	2.35
Don't have to wait	1.5	It takes a little	1.25	Don't have any place	0.2
I don't know why	35.5	all over the country	9.55	I want to sit	3.6
I don't know who	7.0	all over the house	0.85	I want to say	0.2

and conducted a meta-analysis of the reaction times that compared how well a continuous measure of frequency fit the data compared with a categorical one (high vs. low, calculated from the best-fitting breakpoint of frequency). We found a continuous effect of frequency on reaction times across the continuum, and this was a better fit than the categorical measure. Figure 1 shows the model fit with average log reaction times in 6 frequency bins. The fit line shows that the more frequent the phrase, the faster participants respond to it. The fit is derived from a regression

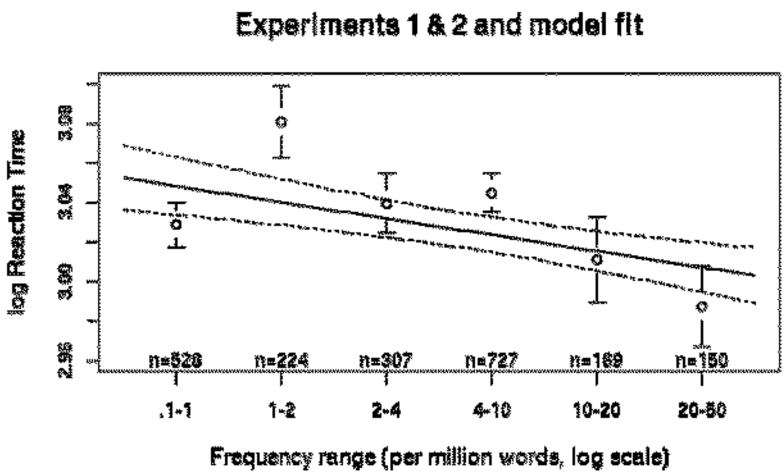

Figure 1. Model fit for reaction times to all phrases. Log reaction time by sequence frequency bin (log scale). Circles represent the means for each bin, with 95% confidence intervals. The fit line is derived from a regression model with a continuous measure of frequency and all control covariates and also includes 95% confidence intervals

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 135)

model with a continuous measure of frequency and all control covariates, so it reflects the effect of 4-gram frequency beyond the frequencies of the subparts of the phrases.

2.4. Discussion

Our findings show that higher frequency phrases are processed faster across the frequency range. The meta-analysis revealed a direct relation between frequency of occurrence and processing latency: the more often a phrase had been experienced, the faster it was processed.

The current findings are hard to accommodate within a strong dual-system model like the words-and-rules model (Pinker 1999) where frequency effects are taken as a marker of lexical storage. Compositional units (regular words or compositional phrases) are not expected to display whole-form frequency effects because they are not stored as such. One way to explain these effects is to allow for compositional forms to be stored. This is the solution adopted by Ullman and Wallenski (2005) to account for the frequency effects found for regularly-inflected words. Our current findings would require this model to extend the lexicon dramatically to include many (if not all) compositional phrases. It is no longer clear what, if anything remains outside the lexicon, thus undermining the distinction between the mental lexicon and grammar that these models depend on.

The distinction could also be maintained if compositional forms could be both 'stored' and 'generated'. Sometimes phrases would be stored (resulting in frequency effects) and other times generated. This solution runs into an equally difficult problem. It is not clear when speakers use each type of phrase or how this can be tested empirically. The results are also not easy to accommodate within a 'weak' dual-system model that posits a unique status for very frequent forms, for there was no indication of a clear difference between very frequent and low frequency phrases. Frequency effects were found across the continuum. Using a frequency threshold as a determiner of storage is problematic because speakers cannot know a priori which phrases will become frequent enough to merit storage. Whatever information is maintained for very frequent phrases must have once been registered for all phrases. This information could be discarded at later stages of learning, but this seems improbable.

The results are most compatible, however, with single-system models where frequency is expected to affect all linguistic forms in a similar way. Compositional phrases showed whole-form frequency effects like those displayed by simple and inflected words.

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 136)

In the next section, we look more closely at the postulated distinction between compositional and non-compositional forms from another perspective: We ask whether idioms, often thought to be stored in the lexicon in dual-system models, are processed differently from compositional phrases. We do this by conducting a syntactic priming experiment to see if, and to what degree, idiomatic and non-idiomatic phrases activate the syntactic structure they occur in.

3. Priming from idiomatic and non-idiomatic datives

In this section, we report an experiment that shows that idiomatic datives prime their syntactic structure (make it more likely to be repeated) just as well as non-idioms do. But first, we take up the status of idioms and non-idioms in dual-system and single-system models.

3.1. Introduction

The processing of idioms further blurs the distinction between 'stored' and 'computed' material. Idioms are often seen as prototypical candidates for 'storage' in dual-system models because of their non-compositional character (Pinker 1999; Jackendoff 1995). Take a prototypical example like *kick the bucket*: the meaning of this phrase (at least synchronically) is radically different from what would be expected given typical uses of *kick* and *bucket*. The meaning 'die' arises from the idiosyncratic interpretation of this particular combination of lexical items. More generally, idioms cannot be transparently derived from their parts. To deal with this, idioms are assumed to be stored in the lexicon as a single entry that contains their special lexical, semantic and syntactic features (e.g. Jackendoff 1997). Such a view suggests that idiomatic and non-idiomatic phrases are generated, processed, and retrieved differently. Since idioms are stored together with their structural information, they should not undergo the same syntactic processes as non-idiomatic phrases. Much of the research on idioms assumes a dual-system view in which idioms are stored in the lexicon, with the debate centering on the degree to which they have internal structure (how compositional they are, e.g., Nunberg, Sag and Wasow 1994) and on the kind and amount of syntactic information represented in their lexical entries (e.g., Cutting and Bock 1997; Sprenger et al. 2006).

In an influential article, Nunberg, Sag, and Wasow (1994) assumed a dual-system model but argued that many idioms are compositional, thereby

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 137)

limiting the number of idiomatic expressions that need to be stored. They pointed out that many structures with highly metaphorical or idiomatic meaning derive that meaning from a metaphorical sense of the words involved, which is then computed by regular syntactic processes. In their account, a phrase like *pull strings* derives its idiomatic meaning from a metaphorical use of *strings*, meaning something like "connections", but the structure is otherwise compositional. On the other hand, *kick the bucket* is an idiomatic construction in their theory since there is no metaphorical sense of *kick* or *bucket* that yields the meaning 'die'. This view clearly predicts differences between the syntactic processes involved in idiomatic and non-idiomatic constructions.

Single-system models acknowledge that while there is a lot of idiosyncrasy peculiar to idioms, they share a lot of structural similarities with other syntactically compositional structures. They therefore blur the distinction between storage and computation by allowing for redundancy in linguistic representation. Goldberg (2006) pointed out that regular linguistic patterns are often instantiated by exemplars that are highly idiomatic. For example, an idiom like *go kicking and screaming* <*path*> is structurally an exemplar of the general pattern *go*VP*ing*. Expressions like *the bigger they come, the harder they fall*, and *the more the merrier*, are instances of the more general *the* X*er the* Y*er* pattern, even though they have their own special features as well. Idioms are tokens of more general (and regular) patterns in addition to being tokens of their own more specific patterns. Given this analysis, and assuming that all linguistic material is processed by a similar mechanism, a single-system account would predict (1) that idioms should maintain links with the more general (and regular) patterns they are instances off, and (2) that, in doing so, they will be similar to non-idiomatic expressions.

3.2. How to distinguish idioms from non-idioms?

Dual-system models treat idioms and non-idioms as qualitatively different entities, but such a distinction is not easy to operationalize. One key issue in the study of idioms is how they should be defined: what makes something an idiom? Such a definition is hard to come by since idiomatic phrases seem to fall on a continuum of compositionality (how transparent their meaning is given their parts) and flexibility (how flexible they are in terms of the lexical items used, number, tense, etc.), with both factors contributing to their perceived status as an idiom (Jackendoff 1997; Nunberg, Sag, & Wasow 1994; Wulff 2008).

Empirically, many findings highlight the inherent complexity in classifying idioms. Idiomatic expressions don't fall neatly into compositional and non-compositional. Idioms differ in their degree of compositionality. Some idioms get their idiomatic meaning more from individual words than others. For example, changing a phrase like *kick the bucket* to *kick the pail* doesn't evoke the idiomatic meaning of 'die'. But a similar change from *pop the question* to *pop the request* still retains some of the idiomatic meaning of 'propose marriage' (Gibbs & Nayak 1989). That is, even phrases that seem highly non-compositional show some degree of reliance on their parts for meaning. Idioms also fall on a continuum with regard to their flexibility: whether the idiomatic meaning is retained in different syntactic constructions (e.g., passivization), morphological realizations (e.g. change of person, number, etc.), and lexical substitutions. Flexibility also seems to be a matter of degree, and is affected by various factors (Wulffe, 2008). Idiomatic phrases can be more flexible in one dimension than another. For example, the idiomatic phrase *throw in the towel* cannot be passivized (*the towel was thrown in* does not mean 'quit') but the idiomatic meaning is retained when the verb is substituted with *toss*. In sum, neither flexibility nor compositionality provide a clear-cut way to distinguish between idioms and non-idioms; idiomaticity seems to a gradient notion and not a categorical one.

In light of these findings (and because resolving this quandary is beyond the scope of the current chapter), we adopt a working definition of idiomatic phrases based on semantic compositionality taken from Nunberg et al. (1994), also adopted by Konopka and Bock (2009): An utterance is idiomatic to the degree that its meaning is not predictable from any regular sense of the words involved. Importantly, our definition of an idiom is gradient, not categorical. Our claims should be understood accordingly: in this chapter, we are interested in investigating parallels in the processing of phrases differing in their degree of idiomaticity.

Several results reveal such parallels between the processing of idiomatic and non-idiomatic phrases. Comprehension and production findings show that literal word meanings are activated during idiom processing (Cacciari and Tabossi 1988; Cutting and Bock 1997). Sprenger et al. (2006) showed that idioms can prime and be primed by words that appear in them (e.g. *hit the road* primes *road*), suggesting that like compositional phrases, they have internal structure. Konopka and Bock (2009) showed that idiomatic and non-idiomatic phrasal verbs (e.g. *pull off a robbery*) can prime particle placement (whether the particle appears before or after the direct object) in non-idiomatic phrases that have no lexical overlap (e.g. *knocked over*

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 139)

the vase can prime *pull off a robbery*, see section 3.3). Using acceptability judgments of familiar and invented idioms, Tabossi, Wolf, and Koterle (2009) suggested that the syntax of idioms is governed by syntactic and pragmatic principles qualitatively similar to those that govern non-idiomatic language.

In this section, we add to these studies by providing further evidence that idiomatic and non-idiomatic datives prime their syntactic construction to a similar degree. Such a finding (1) enhances the idea that idioms have internal syntactic structure and (2) undermines the possibility of distinguishing empirically between idiomatic and non-idiomatic forms, a distinction predicted by dual-system, but not single-system, models.

3.3. Using syntactic priming to compare idiomatic and non-idiomatic phrases

Several methodologies have been applied to idiom processing. Syntactic priming is particularly interesting because it offers insight into the representational similarity of structures. In syntactic priming, syntactic structures are re-used by speakers, as in the following dialogue from the Switchboard corpus (Godfrey 1992):

(1) I don't feel we should *loan them money*...

I wish our leaders were really seeking the Lord on these things, and if we feel led to *give a country money* to help them, fine...

The speaker first chooses the Double Object (DO) dative construction *loan them money*, even though the Prepositional Object (PO) construction is possible (*loan money to them*). Later, when the speaker produces another dative, they again choose the DO alternate, possibly because of priming from the previously produced dative. Priming was first commented on by sociolinguists (Sankoff & LaBerge 1978; Poplack 1980; Weiner & Labov 1983; Estival 1985), but experimental psychologists (Bock 1986; Pickering & Branigan 1998) have since argued strongly for its role in illuminating representations in language processing. In the psychological literature, priming is seen as a general process (i.e. occurring in both production and comprehension) where the processing of a stimulus (the 'target') is facilitated if a similar stimulus (the 'prime') has just been processed. This facilitation is greater the more similar the prime and the target, and in fact only occurs if they are similar along some cognitive dimension. As Branigan *et al* (1995) argue, this is why priming can illuminate the mental representation of linguistic knowledge, because if people's behavior is sensitive to this similarity, it indicates that the two structures share a cognitive representation on some dimension. Thus, by exploring the dimensions of similarity experimentally between primes, one may gain insight into the mental representations of the relevant stimuli. The dependence of priming on similarity is important: an utterance should prime a construction (i.e. make it more likely to be repeated, or more easily comprehended) only if it is perceived to be an instance of that construction. Priming thus becomes an important diagnostic for determining whether idioms are instances of the more abstract (and regular) constructions they appear in.

Single-system and dual-system models make different predictions about the priming of idioms. 'Strong' dual-system models (Jackendoff 1997) argue that idioms are stored separately from superficially similar structures with similar word orders. They would predict that idioms should not prime superficially similar structures. For example, a compositional phrase like *give the child some food* is an instance of the double-object dative pattern. However, an idiom like *give the child a lift* is stored separately, and is not a token of the double-object dative pattern. An idiom should therefore not be able to prime a compositional structure like the double object dative because it is not structurally similar. 'Weaker' dual-system models may allow idioms to have internal syntactic structure (Chang, Dell, and Bock 2006), but would probably predict that idioms would prime less than non-idiomatic phases because the link to the construction is weaker. In single-system models, idiomatic and non-idiomatic phrases are represented in the same way (Goldberg 2006). Idioms therefore have internal structure: to the extent that two structures share features like lexical items, argument order, and syntactic construction, they should prime one another. Therefore an idiomatic dative should prime a non-idiomatic one.

3.4. Previous work with priming

The first experiment to examine semantic compositionality and priming was Konopka and Bock (2009). They did a production priming study of the verb-particle alternation where a particle precedes or follows the object NP (e.g., *A celebrity threw in the first ball.* vs. *A celebrity threw the first ball in.*). The task was to repeat a sentence that had been presented rapidly, one word at a time, in the center of a screen. People sometimes mis-repeat the target sentence and use the other alternant instead of the original. Konopka and Bock measured whether this tendency to mis-remember increased when the other alternant was primed by appearing in the previous sentence, and indeed they found priming in this alternation. They went on to manipulate the idiomaticity of the prime sentence, as determined by a norming task where idiomaticity was defined as the extent to which the meaning of the sentence deviated from that expected given the "dictionary definitions" of the words in the sentence. In this way, they took into account the points of Nunberg, Sag, and Wasow (1994), by defining idioms as *constructions* with idiomatic meaning, not structures that derive their metaphorical or idiomatic interpretation from the metaphorical or idiomatic senses of the words. They found that idiomaticity had no effect on priming: idiomatic verb-particle constructions (*The teenager shot off his mouth.*) were just as likely to be repeated as non-idiomatic ones were (*Judy snapped on her earrings.*). They also looked at the effect of flexibility (whether the structure can appear in the second alternant) on how likely the structure is to be repeated. This is related to the hypothesis that flexibility is correlated with semantic compositionality and hence the idiomaticity of the construction (Jackendoff 1997; Nunberg *et al.* 1994). They manipulated flexibility independently (along with idiomaticity) and found a main effect of flexibility in that frozen structures (e.g., *The crooked salesman couldn't take the customer in*) were less likely to be repeated than flexible structures (*The graduating senior sent his application in*), but found no interaction with idiomaticity. Their findings showed that idiomaticity does not affect whether a structure primes, suggesting similarity in the syntactic processes associated with idiomatic and non-idiomatic structures.

Given that this is the only experiment to date that has examined compositionality and production priming, we wanted to look more closely at the effect of idiomaticity (as measured by semantic compositionality) on priming. We conducted another experiment using a different methodology and a different syntactic alternation. We wanted to use a method closer to natural production where participants have more freedom in what they produce, so we chose a sentence completion task, where participants complete sentence fragments. And we used the dative alternation where the double object (DO) structure (*The mother gave the hungry baby some food*) alternates with the prepositional object structure (*The mother gave some food to the hungry baby*), because this construction lends itself well to completion tasks (Pickering and Branigan 1998).

Strong dual-system models predict that idioms should not prime the structure that they occur in, or at least prime it less, because they are represented in a fundamentally different way from a superficially similar compositional structure. Single-system models predict that idioms should prime their structure just as well as non-idioms do, if they share similarities like lexical items and argument order.

3.5. Syntactic Priming Experiment

We did a production priming experiment of the dative alternation that manipulated the idiomaticity of the prime.

3.5.1. Method

3.5.1.1. Participants

Thirty-five students (mean age 20 years) from the University of Rochester participated in the study. All were native English speakers and were paid $7.50 in return for their participation.

3.5.1.2. Procedure

We used a sentence completion task (Pickering and Branigan 1998) to assess production priming. Participants saw partial sentences (one at a time) on a screen, and were instructed to complete them in the most sensible way, succinctly, without using pronouns, and to type the entire sentence (not just their additional material) into the input box. Participants were told that if a word or phrase appeared in parentheses after the fragment, they should use that material in the completed sentence. This ensured that the desired recipient and theme were used. Participants saw sequences of prime and target sentences, with fillers appearing between each prime-target sequence. Prime sentences contained enough material to force participants to complete them with the desired alternation: for the DO condition, the sentence fragment included the recipient (e.g. "The mother gave the hungry baby (some food)"), and for the PO condition, the sentence fragment included the theme and the preposition '*to*' (e.g. "The mother gave some food to (the hungry baby)"). The target sentence fragment contained only a subject NP and a dative verb (e.g. "The flight attendant gave"), and could be completed with either alternative. The experiment was conducted using Linger (developed by Douglas Rhode, http://tedlab.mit.edu/~dr/Linger).

3.5.1.3. Materials

The experiment contained 24 items, with each item appearing in two conditions that varied in prime construction (Double Object vs. Prepositional Object). Idiomaticity was manipulated between items (based on a measure of idiomaticity derived via the norming experiment described in the next section). Our choice of theme determined whether the utterance was idiomatic or not. The two item variants (DO or PO) were followed by the same sentence fragment to elicit the target. A sample item is illustrated in Table 2:

Table 2. Example materials for the priming experiment

Prime:		
Higher idiomaticity	DO:	The lifeguard gave the surfer (the creeps)
	PO:	The lifeguard gave the creeps to (the surfer)
Lower idiomaticity	DO:	The mother gave the hungry baby (some food)
	PO:	The mother gave some food to (the hungry baby)
Target:		The flight attendant gave

More and less idiomatic datives were extracted from the British National Corpus (BNC, the automatically parsed version of Roland et al. 2007). The BNC corpus was used because of its size and availability in at least an automatically parsed form, in order to have a sufficient number of idioms and their frequencies. We first extracted all dative sentences where the verb-theme combinations were of sufficient frequency (over 10 times in the BNC corpus). We selected more and less idiomatic verb-theme combinations that we then normed for idiomaticity (using the definition of Nunberg, Sag, and Wasow 1994 discussed above) and alternation bias (how likely they are to appear in either construction). The norming was done by American English speakers to fit the language experience of the American participants who participated in the priming experiment. We provide more details about this in the next section. The vast majority of dative idioms in the corpus involved the verb '*give*', so much so that 12 idiomatic items could not be constructed with a reasonable variety of verbs. We therefore decided to use only *give* in all the primes and targets (but obviously not the fillers). Given that '*give*' is used in 80% of datives in spoken language (Bresnan et al. 2007), we assume that participants would not notice the high frequency of '*give*' in the experiment. As a further precaution, we presented only half the items, counterbalanced, to each subject, so they would not see too many tokens of '*give*'. We also selected the items with respect to how they were scored on norming tasks measured idiomaticity and flexibility (whether there was a strong bias towards PO and DO), as described below. The materials were presented by Linger in 8 randomized lists using a latin-square design, and each participant saw only one of the two variants of each item.

3.5.2. Norms

3.5.2.1. Idiomaticity

The idiomaticity of each variant was determined using a rating task performed over the web (on Amazon Mechanical Turk, www.mturk.com).

Participants were asked to judge the idiomaticity of each item. Idiomaticity was defined just as in Konopka and Bock: how predictable the meaning of the sentence is given the "dictionary definitions" of the words involved. Participants rated idiomaticity on a 1–7 scale, with 7 being highly idiomatic and 1 being highly non-idiomatic. All items were presented in the DO alternant. 10 ratings were collected for each stimulus (2 conditions per item, for 40 stimuli). Because participants on Mechanical Turk do not have to complete the entire experiment (and often do not), 40 people participated, with the only restriction on participation being that their IP address be from the United States (so that they will resemble the test population). Each item consisted of one task page in the Mechanical Turk interface, with a filler occurring before each experimental item. Participants were paid $0.02 for each stimulus completed.

We analyzed the norming task, and found that there was a significant difference in idiomaticity judgments between the "idiom" and "non-idiom" items ($t(31) = 24$, $p < .001$): "idiom" items had a mean of 4.2 (range 3.7–4.8), and "non-idiom" items a mean of 2.2 (range 1.1–2.7).

3.5.2.2. Alternation bias

We performed a further norming experiment to determine the bias of each of the 24 items towards the DO or PO construction. We did this for two reasons: First, Konopka and Bock showed that idiomaticity and flexibility were independent factors, and we wanted to manipulate idiomaticity independently of flexibility, so we ensured that all items were flexible, they could occur plausibly in both alternations. Second, since Konopka and Bock found an effect of alternation bias on priming such that structures that do not alternate also prime less, we wanted to be able to add item-bias as another factor in our analysis.

Participants were asked to compare the acceptability of the PO alternant versus the DO one using magnitude estimation (Bard et al. 1996). One alternant was set as a baseline (with a score of 100), and participants were asked to judge how many times more or less acceptable the other alternant was by comparison. Which alternant was presented as the baseline was randomized, and only one condition was presented per item per participant, with the condition selected at random. Each item consisted of one task page in the Mechanical Turk interface, with a filler occurring before each experimental item. Some fillers included what we thought would be non-alternating datives (extremely biased towards PO or DO) as a comparison (e.g. "*The captain gave the old sailor the willies.*"). Participants were paid $0.02 for each stimulus completed. 24 judgments were collected per item (for an average of 12 per stimulus), and 133 people participated, restricted to United States IP addresses. The norming results confirmed that all items were indeed variable (all experimental stimuli fell within 2 standard deviations of the mean log odds with respect to their alternation bias).

3.5.3. Fillers

Each experimental item (prime-target pair) was separated by at least 2 fillers, with a total of 42 fillers, and the first 4 were part of a practice block. Half of the fillers used intransitive verbs, one quarter used monotransitive verbs in the simple past tense in order to elicit the active voice, and one quarter used monotransitive verbs in the passive participle and with the preposition '*by*' in order to elicit a passive. Active and passive fillers were presented in order to distract from the dative alternations being elicited in the main experiment and to mask the true object of study.

3.6. Results

Each response from the participants on the experimental stimuli (both primes and targets) was coded by the first author for construction (DO, PO, or non-dative). Some participants produced fewer than 20% datives in the target ($n = 3$), or fewer than 20% of one alternant ($n = 11$), so they were excluded for producing insufficient variation. This left 21 subjects for the analysis. Prime-target pairs where the prime and the target were not both completed with a dative were also excluded ($n = 68$), leaving 100 tokens for the analysis.

The data were analyzed with mixed-model logistic regression (for more details on analyzing categorical data with such models, see Jaeger 2008). The dependent variable was whether the prime construction was repeated in the target (1 = repeated, 0 = not repeated). A positive and significant model intercept would indicate priming: it would show that prime construction affected the target construction. The independent variable was a categorical variable representing the idiomaticity of the prime. We ran a mixed-effect model with idiomaticity as a fixed effect. The model also included a random effect of subject and another random effect of item that modeled whether the primes had the same subject and recipient (16 levels, these were sometimes repeated in order to produce more natural stimuli). There was a significant effect of priming ($B = 0.63$, $p < .005$), indicating that the alternant produced in the prime was likely to be repeated in the target. There was no main effect of idiomaticity ($B = -0.03$, $p > .8$).

We also tested an effect of a continuous covariate of idiomaticity derived from the norming data because covariates have been argued to have increased power over arbitrarily defined categorical factors (Baayen 2008, p. 237). There was still no effect of idiomaticity ($B = 0.03$, $p > .8$). We also tested for main effects or interactions with alternation bias (again derived from the norming data), because Konopka and Bock found an effect of construction flexibility on priming. We found no effects or interactions with construction flexibility; however, our materials were designed to have less variability along this dimension, which may explain the difference between our results and Konopka and Bock's.

The general priming effect is illustrated in Figure 2. The y-axis shows the proportion of primes repeated for each condition on the x-axis. For ease of visualization, the idiomaticity factor is shown as a categorical variable. The effect of priming is clear in that all conditions have a repetition rate of greater than 50%: the construction is more likely to be repeated than not. Another way to quantify the effect is to see if the proportion of one construction (say PO) is higher after primes of the same construction. PO constructions were produced more often after PO primes (53%) than after DO primes (26%), whether the prime was idiomatic or not (62% PO after an idiomatic PO prime, and 43% PO after a non-idiomatic PO prime; this difference is not statistically significant).

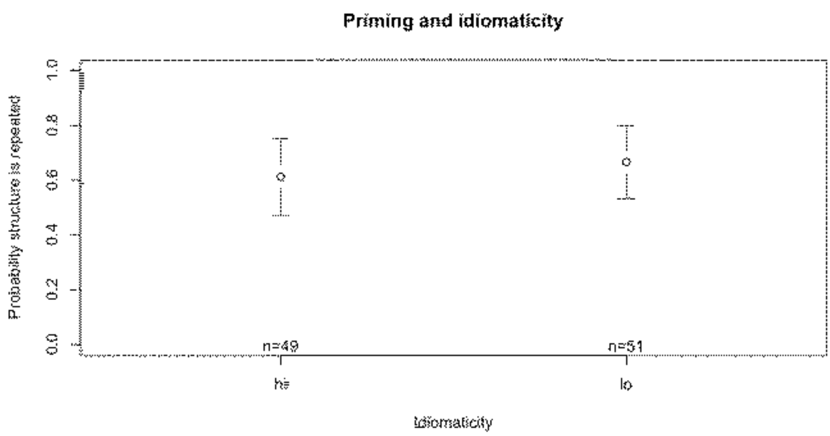

Figure 2. Proportion of prime structures repeated in the target for idiomatic and non-idiomatic primes. The bars represent 95% confidence intervals

The results of the experiment show that idiomaticity (as defined and manipulated in this experiment) does not affect priming. These findings are consistent with Konopka and Bock's finding that flexibility is the determinant of priming behavior, not compositionality.

3.7. Discussion

These results, along with those found earlier, suggest that idioms do have internal syntactic structure, and in some respects, involve syntactic processes similar to those used with non-idiomatic expressions. Semantic compositionality does not seem to determine whether an utterance is, or is not, an instance of a construction. Given that structures that belong to the same construction are more likely to prime than those that belong to two superficially similar constructions (supposedly stored separately because of the non-compositional meaning of one of them), the priming results indicate a shared construction for idiomatic and non-idiomatic datives. Despite the fact that the meaning of *The lifeguard gave the surfer the creeps* is harder to derive from its parts than the meaning of *The mother gave the hungry baby some food*, the former is still perceived to be an instance of the Double Object/Prepositional Object construction.

One important limitation of the current results is that the idiomatic forms we used (since we were limited to verbs that could alternate) were not as non-compositional as in previous work on idiom processing. However, our idiomatic items were still judged as significantly less compositional than the non-idiomatic ones in the norming study. Our items were also limited in that they all used only one verb (*give*). This could have been a problem if our main question was about the generality of priming (which has been reported with many verb types), but since we were concerned with the similarity in priming between more and less idiomatic items, the repetition of the verb becomes less of an issue. We will return to both these issues in the General Discussion.

The priming results are more consistent with a single-system model in which compositional and non-compositional phrases are processed in a qualitatively similar way, and where idioms have internal structure. This is in fact quite similar to Konopka and Bock's 'structural' model of idioms where "their internal structure is accessible to and undergoes the type of generalized syntactic processing involved in both production and comprehension" (pp. 4). There is another argument to make against using semantic compositionality as a determinant of lexical storage, as in dual-system models (Goldberg 2006; Wray 2002). As we noted earlier, it is not easy to determine whether a phrase is compositional or not, since compositionality is more a matter of degree than a binary distinction. Moreover, from the perspective of the child learner who has yet to home in on the regularities of the language, all linguistic input starts out being idiosyncratic and 'irregular' to some degree. Starting out, a child cannot know that *dogs* is regular but *teeth* is not. To extract patterns of regularity, the child first has to have access to multiple stored tokens, both regular and irregular.

This is not to say that idiomatic expressions are not special – speakers have to acquire knowledge about their idiosyncratic semantic (and sometimes syntactic) features to be able to properly produce and comprehend them. But such knowledge may be learned via the same mechanisms used to derive meaning and structure from compositional forms. Put differently, the unique meanings (and norms of use) for idiomatic expressions can be learned from experience without blocking out what they have in common with other, more compositional phrases.

Idiomatic and non-idiomatic datives in the same construction share structural features like argument order and the presence or absence of the preposition '*to*', so they prime that structure to the same degree. Similarity is also the primary factor that drives generalization in single-system models like connectionist and exemplar models, and the finding that idioms can indeed prime argues for a model with one representational mechanism rather than a model with two separate mechanisms for compositional and non-compositional forms.

4. General Discussion

Dual-system models often use two criteria to differentiate between those structures that are stored and those that are computed: compositionality and frequency. We have presented two experiments that show that linguistic structures are processed in qualitatively the same way regardless of where they fall on the frequency and compositionality continua. In Arnon and Snider (2010), we showed that compositional 4-word phrases are responded to more quickly the more frequent they are. This is evidence that language users have knowledge about the frequency of phrases this size, just as they have knowledge of the frequency of words (regular and irregular, Alegre and Gordon 1999; Baayen et al. 1997; Taft 1979). In this respect, 'stored' elements seem no different from 'computed' ones. Importantly, we also showed that they are sensitive to frequency across the continuum: there is no threshold beyond which phrases are attended to. High and low frequency phrases are processed in a qualitatively similar way: their processing is affected by a continuous measure of frequency.

We also presented a priming experiment that showed that compositionality is unable to differentiate stored and computed representations. Both non-compositional and compositional dative structures prime their construction, and do so to the same degree. One limitation of our results is that the idiomatic phrases in our experiment were somewhat compositional: they were not judged at the far end of the scale in the norming experiment (though they were still judged as significantly less compositional than the non-idiomatic phrases), and they were not as strongly idiomatic as in previous experiments (Konopka & Bock, 2008; Sprenger et al. 2006). While it is possible that more highly idiomatic phrases would prime less, thereby showing their diminished internal structure, such a result has not been found to date. Even studies using more idiomatic phrases than ours (Konopka & Bock, 2008; Sprenger et al. 2006) still find evidence for internal structure in idioms and strong parallels with compositional phrases. It is possible that "stored" and "computed" forms can be empirically distinguished on the basis of other measures, in particular, flexibility (which was not manipulated in the current study). But given the multitude of components that make up flexibility (morphological, syntactic and lexical, Wulff, 2008), and given the fact that it seems to have an effect on priming regardless of idiomaticity (Konopka & Bock, 2008), it is unlikely to provide a clear empirical criteria for distinguishing "stored" and "computed" forms. In future work, we would like to investigate further parallels in the processing of more and less flexible forms as well.

It is also possible that this experiment primed verb specific representations because the same verb was used in the prime and target (Gries &Wulff, 2005). However, even assuming that we are activating subtypes of the *give* dative constructions, we still manipulate compositionality within this set of constructions, so our results are not confounded by repeating *give* in prime and target. Idiomatic and non-idiomatic phrases primed (this construction) to the same degree.

The priming experiment results affirm the finding of Konopka and Bock (2009): compositionality does not affect how much that structure persists and is re-used in later processing. The priming result is also consistent with a 'weak' dual-system model (like that proposed by Konopka and Bock 2009), where there is a distinction between idioms and purely compositional phrases, but idioms are formed using 'regular' syntactic processes. However, such an account leaves little of the original concep-

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 150)

tion of idioms as holistic lexical entries. Taken together with the frequency results, our findings are more consistent with the redundancy in storage predicted by single-system models.

These results support one of the fundamental tenets of single-system models: the similarity between 'stored' and 'computed' forms. In such models, similarity between structures and the frequency of those structures determines the extent to which they generalize. This is unlike dual-system models, which have two separate mechanisms, storage and computation, drawing on different representational bases. The criteria that have been argued to distinguish these two types of structures, semantic compositionality and frequency, are challenged by the current findings Neither serves as a clear empirical criterion distinguishing 'stored' from 'computed' forms. Non-compositional forms still appeared to have internal structure, and there was no evidence for a threshold beyond which frequency affected processing: more frequent structures were processed more easily across the continuum. These findings echo those in the morphological literature showing parallels in the processing of regular and irregular forms (Alegre and Gordon 1999; Baayen et al. 1997; Baayen 2006; Taft 1979).

The difficulty in finding a clear criterion for inclusion in the lexicon has led Elman (2009) to the radical solution of "lexical knowledge without a lexicon". Elman reviews numerous studies detailing the rich information language users have about verbs (from the agents they appear with to the discourse situations they evoke), and the rapid way this information is used in online processing. To explain the ready availability of such detailed, situation-specific lexical information in online processing, Elman suggests that "either the lexicon must be expanded to include factors that do not plausibly seem to belong there; or else virtually all information about word meaning is removed, leaving the lexicon impoverished". He argues for a third alternative, an emergentist model in which linguistic knowledge is viewed as a constantly changing dynamic system and where the lexicon doesn't contain fixed units but dynamic patterns. We propose that phrasal frequency effects and idiom priming effects similarly require a model that transcends traditional notions of the lexicon.

One possibility, in line with exemplar models of language (Bod 1998; Goldinger 1996; Johnson 1997; Pierrehumbert 2001, 2006) is to implement the representations produced by the exemplar-based syntactic models of Bod (1998, 2006) in a spreading-activation network, as proposed in Snider (2008). In the model that Bod presents, syntactic productivity is achieved by starting with arbitrarily large linguistic units and deducing syntactic structure from similarity and statistical inference. The resulting lexicon has

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 151)

structurally analyzed chunks of different grain-sizes, which are necessarily redundant, along with a mechanism for generating larger structures out of them. The processing of units is influenced by the probability of the smaller units used to form them (Bod 2006). Implementing these representations in a spreading-activation network (Snider 2008) will result in patterns of varying levels of abstraction (from fully realized strings of words, to fully abstract constructions) that are linked to each other, and whose activation is related, among other factors, to frequency of occurrence.

Multi-word phrases can be represented naturally in this model, and be linked to the words and smaller strings they consist of. For example, the phrase *don't have to worry* would be linked to *don't*, *have*, *to*, and *worry* as well as *don't have*, *to worry*, and so on. Multi-word phrases, including idioms, are also linked to the more abstract units they are instances of: verb-phrases, constructions, etc. (so *give the old sailor a lift* is linked to the DO construction as well as its own idiom). The same would apply to all phrases, regardless of their semantic compositionality or frequency, and would lead to complementary representations at different grain sizes.

Adopting a single-system model of linguistic representation has many additional implications for language processing and learning. In comprehension, processing should take advantage of such knowledge of the likelihood of generalizations at many levels of abstraction and semantic compositionality. There is already evidence that processing is affected by expectations at many levels: the frequency of words in specific syntactic structures (verb-subcategorization biases, Clifton, Frazier, and Connine 1984; Garnsey, Pearlmutter, Myers, and Lotocky 1997; MacDonald, Pearlmutter, and Seidenberg 1994), co-occurrence relations between verbs and specific arguments (Trueswell, Tanenhaus, and Garnsey 1994); as well as the overall frequency of syntactic structure (e.g. main clause vs. reduced relative, Frazier and Fodor 1978). Representing the connections between similar structures at differing levels of semantic compositionality may play a role in the processing of metaphorical language and conventional expressions that are essential for fluent communication (Pawley and Syder 1983). Production models would have to take into account the possibility of selecting whole phrases from storage, rather than from the two levels (lexicon and grammar) of current models (and there is growing evidence that production is sensitive to fine-grained expectations, Jaeger, in press; Jurafsky et al. 2001; Gahl and Garnsey 2004; Tily et al. 2009) Phrasal storage also has implications for learning, especially if representational knowledge arises by generalizing over tokens of stored experience. Using larger units may aid in extracting grammatical regularities (e.g., using

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 152)

frequent frames to learn about grammatical categories, Mintz 2003), and not doing so may be one of the factors that hinders adult language learning (Arnon and Ramscar 2009).

5. Conclusion

In this chapter we have presented findings that challenge the distinction between 'stored' and 'computed' forms by (1) undermining the empirical criteria used to distinguish between them, and (2) demonstrating parallels in the processing of words and phrases (frequency effects), and idiomatic and non-idiomatic phrases (priming). Frequency, while often thought to be a marker of lexical storage, affects the processing of compositional phrases. Idioms, while often thought to be holistically stored, show priming of their construction just like non-idioms. Together, these findings highlight the utility of models that deal with all linguistic experience in a qualitatively similar fashion, and allow for experience to influence the learning, representation and processing of all linguistic patterns.

References

Alegre, M. and Gordon, P.
1999 Frequency effects and the representational status of regular inflections. *Journal of Memory and Language* 40: 41–61. Elsevier.

Arnon, I. and Ramscar, M.
2009 Order-of-acquisition affects what gets learned. In N.A. Taatgen and H. van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (pp. 2112–2117). Cognitive Science Society.

Baayen, R.H., T. Dijkstra and R. Schreuder
1997 Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language* 37: 94–117. Elsevier.

Baayen, H.
2008 *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R.* Cambridge University Press.

Baayen, R., Feldman, L. and Schreuder, R.
2006 Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language* 55: 290–313.

Bannard, C. and Matthews, D.
2008 Stored word sequences in language learning: the effect of familiarity on children's repetition of four-word combinations. *Psychological Science* 19: 241–248.
Bard, Ellen, Dan Robertson & Antonella Sorace
1996 Magnitude estimation of linguistic acceptability. *Language* 72.1: 32–68.
Barlow, G.M. and Kemmer, S.E.
2000 Usage-Based Models of Language. Stanford, CSLI.
Bates, E. and MacWhinney, B.
1989 Functionalism and the competition model. In *The crosslinguistic study of sentence processing* 3–73.
Bell, A., Brenier, J., Gregory, M., Girand, C. and Jurafsky, D.
2009 Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60: 92–111.
Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory and M., Gildea, D.
2003 Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the Acoustical Society of America* 113: 1001–1024.
Bock, J.K.
1986 Syntactic persistence in language production. *Cognitive Psychology* 18: 355–387.
Bod, R., J. Hay and S. Jannedy
2003 *Probabilistic linguistics*. MIT Press.
Bod, R.
1998 *Beyond Grammar: An Experience-based Theory of Language*. Center for the Study of Language and Information, CA.
Bod, R.
2006 An all-subtrees approach to unsupervised parsing. Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the ACL, 865–872.
Branigan, H.P., M.J. Pickering, S.P. Liversedge, A.J. Stewart and T.P. Urbach
1995 Syntactic priming: Investigating the mental representation of language. *Journal of Psycholinguistic Research* 24: 489–506.
Bresnan, J., Cueni, A., Nikitina, T. and Baayen, H.
2007 Predicting the dative alternation. *Proceedings of the Royal Netherlands Academy of Science Workshop on Foundations of Interpretation.* Amsterdam.
Bybee, J. and P. Hopper
2001 *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.
Bybee, J.
1998 *The emergent lexicon*. In Chicago Linguistic Society, 34: 421–435.

Bybee, J.
2006 From usage to grammar: The minds response to repetition. *Language*, 82, 711–733.
Bybee, J. and McClelland, J.
2005 Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review* 22: 381–410.
Bybee, J. and Scheibman, J.
1999 The effect of usage on degrees of constituency: The reduction of don't in English. *Linguistics* 37: 575–596.
Cacciari, C. and Tabossi, P.
1988 The comprehension of idioms. *Journal of Memory and Language* 27: 668–683.
Chang, F., Dell, G.S. and Bock, J.K.
2006 Becoming syntactic. *Psychological Review* 113: 234–272.
Chomsky, N.
1957 *Syntactic Structures*. The Hague: Mouton.
Chomsky, N.
1965 *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.
Chomsky, N.
1981 *Lectures on Government and Binding*. Dordrecht: Foris Publications.
Chomsky, N.
1995 *The Minimalist Program*. Cambridge, MA: The MIT Press.
Cieri, C., Miller, D. and Walker, K.
2004 The Fisher Corpus: a Resource for the Next Generations of Speech-to-Text. *In Proceedings of LREC 2004*.
Clifton, C., Frazier, L. and Connine, C.
1984 Lexical expectations in sentence comprehension. *Journal of Verbal Learning and Verbal Behavior* 23: 696–708.
Cutting, J.C. and Bock, K.
1997 That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition* 25: 57–71.
Elman, J.L.
1991 Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning* 7: 195–225.
Elman, J.L.
2009 On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science* 33: 547–582.
Estival, D.
1985 Syntactic priming of the passive in English. *Text* 5.7–21.
Fodor, J.
1983 *The modularity of mind*. Cambridge, MA: MIT Press.

Frazier, L. and Fodor, J.D.
1978 The sausage machine: A new two-stage parsing model. *Cognition* 66: 291–325.

Gahl, S. and S.M. Garnsey
2004 Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language* 80: 748–775.

Garnsey, S., Pearlmutter, N., Myers, E. and Lotocky, M.
1997 The Contributions of Verb Bias and Plausibility to the Comprehension of Temporarily Ambiguous Sentences. *Journal of Memory and Language* 37: 58–93.

Gibbs, R. and N. Nayak
1989 Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology* 21: 100–138.

Godfrey, J.J., E.C. Holliman and J. McDaniel
1992 Switchboard: Telephone speech corpus for research and development. In IEEE ICASSP, 517–520.

Goldberg, A.E.
1995 *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.

Goldberg, A.
2006 *Constructions at Work: The Nature of Generalization in Language*. Oxford University Press.

Goldinger, S.D.
1996 Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22: 1166–1183.

Gries, Stefan Th. and Stefanie Wulff
2005 Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. *Annual Review of Cognitive Linguistics* 3: 182–200.

Gregory, M.L., Raymond, W.D., Bell, A., Fosler-Lussier, E., Hay, D., Pierrehumbert, J. and Beckman, M.
2004 Speech perception, well-formedness, and the statistics of the lexicon. *Papers in Laboratory Phonology* VI: 58–74.

Jackendoff, R.
1995 The boundaries of the lexicon, in M. Everaert, E.-J. van der Linden, A. Schenk, and R. Schreuder, (Eds.), *Idioms: Structural and Psychological Perspectives*, 133–165. Hillsdale, NJ: Erlbaum.

Jackendoff, R.
1997 *Architecture of the language faculty*. Cambridge, MA: MIT Press.

Jackendoff, R.
2002 *Foundations of language*. Oxford University Press, New York.

Jaeger, T.F.
2008 Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59:4, 434–446.

Jaeger, T.F.
in press Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*.

Johnson, K.
1997 Speech perception without speaker normalization. In K. Johnson and Mullenix (Eds.) *Talker Variability in Speech Processing*. San Diego, Academic Press, 145–166.

Jurafsky, D., Bell, A., Gregory, M. and Raymond, W.
2001 Probabilistic relations between words: Evidence from reduction in lexical production. In: J. Bybee and P. Hopper, P. (Eds.), *Frequency and the Emergence of Linguistic Structure*. John Benjamins, Amsterdam, 229–254.

Kaplan, R.M. and Bresnan, J.
1982 Lexical-Functional grammar. In J. Bresnan (Ed.), *The mental representation of grammatical relation*, 173–281. Cambridge, MA: MIT Press.

Konopka, A.E. and Bock, J.K.
2009 Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology* 58: 68–101.

Langacker, R.
1987 *Foundations of Cognitive Grammar*. Stanford University Press.

Langacker, R.
1988 A usage-based model. *Topics in Cognitive Linguistics* 50: 127–163.

Levy, R. and Jaeger, T.F.
2007 Speakers optimize information density through syntactic reduction. In: B. Schlokopf, J. Platt, and T. Hoffman (Eds.), *Advances in neural information processing systems* (NIPS), 19, MIT Press, Cambridge, MA, 849–856.

MacDonald, M.C., Pearlmutter, N.J. and Seidenberg, M.S.
1994 The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101: 676–703.

Marchman, V.A.
1993 Constraints on plasticity in a connectionist model of the English past tense. *Journal of Cognitive Neuroscience* 5: 215–234.

McDonald, S.A., and Shillcock, R.C.
2003 Eye-movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science* 14: 648–652.

Mintz, T.H.
2003 Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* 90: 91–117.

Monsell, S.
1991 The nature and locus of word frequency effects in reading. In D. Besner and G.W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*, 148–197. Hillsdale, NJ: Erlbaum.

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 157)

Morton, J.
1969 The interaction of information in word recognition. *Psychological Review* 76: 165–78.

Newmeyer, F.
2003 Grammar is Grammar and Usage is Usage. *Language* 79: 682–707.

Nunberg, G., Wasow, T. and Sag, I.A.
1994 Idioms. *Language* 70: 491–538.

Pawley, A. and Syder, F.H.
1980 Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J.C. Richards & R. Schmidt (Eds.), *Communicative competence*, 191–225. London: Longmans.

Pickering, M.J. and H.P. Branigan
1998 The Representation of Verbs: Evidence from Syntactic Priming in Language Production. *Journal of Memory and Language* 39: 633–651.

Pierrehumbert, J.
2001 Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee and P. Hopper (eds.) *Frequency Effects and the Emergence of Lexical Structure*. John Benjamins, Amsterdam, 137–157.

Pierrehumbert, J.
2006 The next toolkit. *Journal of Phonetics* 34: 516–530.

Pinker, S. and Ullman, M.T.
2002 The past and future of the past tense. *Trends in Cognitive Sciences* 6: 456–463.

Pinker, S.
1991 Rules of Language. *Science* 253: 530–535.

Pinker, S.
1999 *Words and Rules: The Ingredients of Language*. Basic Books.

Pinker, S. and Prince, A.
1988 On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28: 73–193.

Plunkett, K. and Marchman, V.
1991 U-shaped learning and frequency effects in a multi-layered perception: Implications for child language acquisition. *Cognition* 38: 43–102.

Plunkett, K. and Marchman, V.
1993 From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition* 48: 21–69.

Pollard, C.J. and Sag, I.A.
1994 *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.

Poplack, S.
1980 The notion of the plural in Puerto Rican Spanish: Competing constraints on/s/deletion. *Locating language in time and space*, 55–86.

Prince, A. and Pinker, S.
1988 Rules and connections in human language. *Trends in Neurosciences* 11: 195–202.

Reali, F. and Christiansen, M.
2007 Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language* 57: 1–23.

Roland, D., Dick, F. and Elman, J.L.
2007 Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language* 57: 348–379.

Rumelhart, D. and McClelland, J.
1986 On learning the past tenses of English verbs. In: D. Rumelhart and J. McClelland, J. (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. MIT Press, 216–271.

Sankoff, D. and S. Laberge
1978 Statistical dependence among successive occurrences of a variable in discourse. *Linguistic Variation: Methods and Models*, 119–126.

Seidenberg, M.S.
1994 Language and connectionism: The developing interface. *Cognition* 50: 385–401.

Snider, N.
2008 *An exemplar model of syntactic priming*. Unpublished dissertation, Stanford University, Stanford, CA.

Sprenger, S., Levelt W.J.M. and Kempen, G.
2006 Lexical access during the production of idiomatic phrases, *Journal of Memory and Language* 54: 161–184.

Taft, M.
1979 Recognition of affixed words and the word-frequency effect. *Memory and Cognition* 7: 263–72.

Tily, H., Gahl, S., Arnon, I., Kothari, A., Snider, N. and Bresnan, J.
2009 Pronunciation reflects syntactic probabilities: Evidence from spontaneous speech. *Language and Cognition* 1:2.

Tomasello, M.
2003 *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.

Trueswell, J.C., Tanenhaus, M.K. and Garnsey, S.M.
1994 Semantic Influences on Parsing: The use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 33: 285–318.

Ullman, M.T.
2001 A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience* 2: 717–726.

Ullman, M.T., Corkin, S., Coppola, M., Hickok, M., Growdon, J.H., Koroshetz, W.J. and Pinker, S.
1997 A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience* 9: 266–276.

Ullman, M.T., Pancheva, R., Love, T., Yee, E., Swinney, D. and Hickok, G.
2005 Neural correlates of lexicon and grammar: Evidence from the production, reading, and judgment of inflection in aphasia. *Brain and Language* 93: 185–238.

Ullman, M.T. and Wallenski, M.
2005 Reply: Moving past the past tense. *Brain and Language* 93: 248–252.

Underwood, G., Schmitt, N. and Galpin, A.
2004 The eyes have it: An eye-movement study into the processing of formulaic sequences. In: N. Schmitt (Ed.) *Formulaic Sequences: Acquisition, Processing, and Use*. John Benjamins, Amsterdam, 153–172.

Weiner, E.J. and William Labov
1983 Constraints on the agentless passive. *Journal of Linguistics* 19: 29–58.

Wray, A.
2002 *Formulaic Language and the Lexicon*. Cambridge University Press.

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

Appendix A: experimental idioms for study 2

The columns contain the following information: item number, condition, prime structure, mean idiomaticity rating, prime stimulus, target stimulus.

Item	Condition	Prime	Idiom-aticity	Prime stimulus	Target stimulus
1	nonidiom	do	2	The racing driver gave the helpful mechanic (a job)	The patient gave
	nonidiom	po		The racing driver gave a job to (helpful mechanic)	The patient gave
2	idiom	do	3.7	The efficient secretary gave the grumpy businessman (a look)	The little girl gave
	idiom	po		The efficient secretary gave a look to (grumpy businessman)	The little girl gave
3	nonidiom	do	2	The famous journalist gave the fashion designer (her address)	The diver gave
	nonidiom	po		The famous journalist gave her address to (fashion designer)	The diver gave
4	idiom	do	3.7	The blackmailer gave the sleazy journalist (control)	The lonely sailor gave
	idiom	po		The blackmailer gave control to (sleazy journalist)	The lonely sailor gave
5	nonidiom	do	2.7	The millionaire gave the struggling artist (some advice)	The explorer gave
	nonidiom	po		The millionaire gave some advice to (struggling artist)	The explorer gave
6	idiom	do	4.6	The mother gave the hungry baby (a boost)	The flight attendant gave
	idiom	po		The mother gave a boost to (hungry baby)	The flight attendant gave
7	nonidiom	do	2.5	The researcher gave the experienced surgeon (some information)	The man gave
	nonidiom	po		The researcher gave some information to (experienced surgeon)	The man gave
8	idiom	do	4.4	The cheerful engineer gave the architect (an edge)	The teacher gave
	idiom	po		The cheerful engineer gave an edge to (architect)	The teacher gave

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 161)

9	nonidiom	do	1.1	The mother gave the hungry baby (some food)	The flight attendant gave
	nonidiom	po		The mother gave some food to (hungry baby)	The flight attendant gave
10	idiom	do	4.5	The famous journalist gave the fashion designer (a hand)	The diver gave
	idiom	po		The famous journalist gave a hand to (fashion designer)	The diver gave
11	nonidiom	do	2.2	The lifeguard gave the surfer (a list)	The inventor gave
	nonidiom	po		The lifeguard gave a list to (surfer)	The inventor gave
12	idiom	do	4.3	The spy gave the double agent (trouble)	The consultant gave
	idiom	po		The spy gave trouble to (double agent)	The consultant gave
13	nonidiom	do	2.4	The grandmother gave the little girl (some money)	The tennis fan gave
	nonidiom	po		The grandmother gave some money to (little girl)	The tennis fan gave
14	idiom	do	3.8	The woman gave the new neighbor (credit)	The librarian gave
	idiom	po		The woman gave credit to (new neighbor)	The librarian gave
15	nonidiom	do	2.3	The kind teacher gave the youngster (directions)	The private detective gave
	nonidiom	po		The kind teacher gave directions to (youngster)	The private detective gave
16	idiom	do	4.2	The lifeguard gave the surfer (five)	The inventor gave
	idiom	po		The lifeguard gave five to (surfer)	The inventor gave
17	nonidiom	do	2.6	The wedding planner gave the guests (a picture)	The pharmacist gave
	nonidiom	po		The wedding planner gave a picture to (guests)	The pharmacist gave
18	idiom	do	4.1	The car salesman gave the couple (some thought)	The park ranger gave
	idiom	po		The car salesman gave some thought to (couple)	The park ranger gave
19	nonidiom	do	2.6	The manager gave the secretary (instruction)	The boyfriend gave
	nonidiom	po		The manager gave instruction to (secretary)	The boyfriend gave

20	idiom	do	4.8	The lifeguard gave the surfer (the creeps)	The inventor gave
	idiom	po		The lifeguard gave the creeps to (surfer)	The inventor gave
21	nonidiom	do	2	The car salesman gave the couple (a job)	The park ranger gave
	nonidiom	po		The car salesman gave a job to (couple)	The park ranger gave
22	idiom	do	3.7	The manager gave the secretary (a shot)	The boyfriend gave
	idiom	po		The manager gave a shot to (secretary)	The boyfriend gave
23	nonidiom	do	2.1	The efficient secretary gave the grumpy businessman (an answer)	The little girl gave
	nonidiom	po		The efficient secretary gave an answer to (grumpy businessman)	The little girl gave
24	idiom	do	4.6	The captain gave the old sailor (a lift)	The bus driver gave
	idiom	po		The captain gave a lift to (old sailor)	The bus driver gave

(V9 21/3/12 13:49) WDG-LCB (155mm×230mm) TimesNRMT 1374 Divjak pp. 127–164 1374 Divjak_04_snider (p. 163)