Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers

Anna Siyanova-Chanturia, Kathy Conklin and Norbert Schmitt

Second Language Research published online 21 February 2011
DOI: 10.1177/0267658310382068

second
language
research

Second Language Research
1–22

DOI: 10.1177/0267658310382068
slr.sagepub.com
SAGE

Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers

Anna Siyanova-Chanturia
University of Modena and Reggio Emilia, Modena, Italy
University of Nottingham, UK

Kathy Conklin and Norbert Schmitt
University of Nottingham, UK

Abstract
Using eye-tracking, we investigate on-line processing of idioms in a biasing story context by native and non-native speakers of English. The stimuli are idioms used figuratively (*at the end of the day* – 'eventually'), literally (*at the end of the day* – 'in the evening'), and novel phrases (*at the end of the war*). Native speaker results indicate a processing advantage for idioms over novel phrases, as evidenced by fewer and shorter fixations. Further, no processing advantage is found for figurative idiom uses over literal ones in a full idiom analysis or in a recognition point analysis. Contrary to native speaker results, non-native findings suggest that L2 speakers process idioms at a similar speed to novel phrases. Further, figurative uses are processed more slowly than literal ones. Importantly, the recognition point analysis allows us to establish where non-natives slow down when processing the figurative meaning.

Keywords
language comprehension, figurative and literal language, mental lexicon, disambiguating context, recognition point

Corresponding author:
Anna Siyanova-Chanturia, Department of Biomedical Sciences, University of Modena and Reggio Emilia, Via Campi 287, Modena 41100, Italy
Email: anna.siyanova@unimore.it

I Introduction

The use of formulaic language, such as collocations, idioms, speech formulae, and binomials is regarded as an essential element of native-like communication (Pawley and Syder, 1983; Langacker, 1987; Sinclair, 1991; Cowie, 1998; Wray, 2002; Tomasello, 2003). Although figures vary as to how much of formulaicity is present in discourse, it is believed that about four formulaic sequences are produced in every minute of spoken discourse (Pollio et al., 1977; Glucksberg, 1989).

A number of researchers have proposed that, due to their frequency of occurrence and hence familiarity, formulaic language is processed differently from novel language (Langacker, 1987; Van Lancker and Kempler, 1987; Wray, 2002; Jurafsky, 2003). For example, Sosa and MacFarlane (2002) had participants monitor for *of* in two-word collocations (*sort of*) varying in frequency. They found that reaction times to *of* in high frequency phrases were significantly slower than in low frequency ones, indicating that frequent phrases were treated as unitary entities, which hindered access to their individual components. While the preposition *of* was identical across all conditions, lexical properties of the constituent words were not controlled for. Thus, one must be cautious about drawing strong conclusions from the study. In another study, Mondini et al. (2002) investigated the production of two-word compounds (*natura morta* 'still life') and matched novel combinations (*natura bella* 'beautiful nature') by two aphasic patients. Mondini and colleagues found that their participants performed significantly better on compounds than on novel noun–adjective combinations. This was taken to indicate that for novel phrases, participants retrieved the adjective and noun separately and then applied agreement rules. Compounds, on the other hand, were retrieved as a unit and, therefore, no morphosyntactic operations were necessary. Unfortunately, it is difficult to draw any far-reaching conclusions, because the study only investigated two brain-damaged participants.

These studies and other similar ones (Van Lancker et al., 1981; Bannard and Matthews, 2008; Arnon and Snider, 2010), imply that there are differences in the way people process novel language and formulaic speech. At the very least, the latter seems to enjoy faster processing. In the current study, we focus on one particular type of formulaic speech – idioms – in order to better understand how they are processed on-line in a first and second language.

II Idiom processing in native speakers

Since the 1970s idioms have received a fair amount of attention. One of the reasons idioms have been widely studied is that many of them allow for two distinct interpretations: figurative and literal.[1] Much of the research on idiom comprehension in native speakers has addressed the following issues: (1) activation of idioms' figurative vs. literal meanings, and (2) processing of idiomatic expressions vs. novel phrases. With regards to the former, a number of models have been proposed (Bobrow and Bell, 1973; Swinney and Cutler, 1979; Cacciari and Tabossi, 1988). One of the influential models of idiom comprehension is the 'lexical representation hypothesis' by Swinney and Cutler (1979). They propose that idioms are represented in the mental lexicon much like morphologically complex words are. They argue that the computation of the literal meaning and the retrieval of the figurative one are initiated simultaneously, as soon as the first word of the

expression is encountered. However, because computation of the literal meaning is more time-consuming than the retrieval of the figurative one, Swinney and Cutler argue that the latter meaning should become activated first.

Another prominent theory of idiom processing puts forward the idea an 'idiomatic key', which refers to the place where the expression becomes recognizable as idiomatic (Cacciari and Tabossi, 1988). According to this theory – dubbed the 'configuration hypothesis' – the individual words and their literal meanings are activated until the 'key' has been reached. Once the idiomatic key is reached, the idiomatic configuration emerges and the figurative meaning is accessed, while the literal meaning is rejected as no longer viable. However, it is important to note that Cacciari and Tabossi (1988), as well as Tabossi and Zardon (1993), point out that the above holds true only in the absence of a biasing context, which would prepare the reader for either figurative or literal rendering.

The above models make predictions with respect to figurative and literal idiom interpretations. In addition, researchers have looked at the processing of idioms vs. novel phrases. For example, Swinney and Cutler (1979) found that idioms (*break the ice*) were processed more quickly than non-idiomatic phrases (*break the cup*). The findings of Gibbs (1980), Van Lancker et al. (1981), Gibbs and Gonzales (1985), and Tabossi et al. (2009) also suggest that idioms enjoy faster processing (in comprehension, as well as production) than matched novel strings. The 'idiom decomposition hypothesis' (Gibbs et al., 1989) addresses when the figurative meaning of conventional language is activated in relation to novel language. This theory postulates that idiom processing is highly dependent on whether an idiom is decomposable (the meanings of idiom components are related to the overall figurative meaning) or non-decomposable (there is no obvious link between the meanings of idiom components and the overall figurative meaning). Gibbs et al. argue that only in the case of decomposable items (*pop the question*) should idioms be faster than their novel control phrases (*ask the question*), because their individual components contribute to the idiom's figurative meaning. For non-decomposable idioms (*kick the bucket*), where no such link exists, no processing advantage should be observed for idioms over their novel matches (*fill the bucket*). It is noteworthy, however, that the above proposition is not supported by Tabossi et al.'s (2009) results. In a semantic judgement task, their participants were found to be equally fast at judging decomposable and non-decomposable idioms and their corresponding novel controls, suggesting that whether an idiom's constituents are, or are not, related to the idiom's overall figurative meaning does not affect its processing.

Overall, despite the differences that exist among idiom theories, it is clear that in native speakers, idioms are processed faster than literal language. What remains an open question is whether the literal or figurative meaning of an idiom is activated more quickly. Using an eye-tracking paradigm and two types of analysis, the present investigation examines idiom comprehension with respect to, first, idiom vs. novel phrase comprehension, and, second, figurative vs. literal meaning processing by native and non-native speakers.

III Idiom processing in non-native speakers

Similar to studies with native speakers, one of the issues in the literature on non-native speakers is whether there is a difference in the processing of literal vs. figurative meanings. Van Lancker-Sidtis (2003) looked at whether prosodic cues helped native and

proficient non-native speakers distinguish between the two idiom interpretations. Participants listened to sentences that contained idioms used either figuratively or literally and then had to identify the intended meaning. Results suggested that prosodic cues enabled native participants to successfully differentiate between idioms used figuratively and literally, whereas even highly proficient non-natives were unable to do so.

In a cross-modal priming study by Cieslicka (2006), non-native participants listened to neutral sentences containing familiar idioms (*George wanted to bury the hatchet soon after Susan left*), and performed a lexical decision on one of four targets: a word related to the idiom's figurative meaning (*forgive*), its control (*gesture*), a word related to its literal meaning (*axe*), or its control (*ace*). Faster response times to targets related to the literal meaning than to ones related to the figurative one suggest that literal idiom interpretations are activated prior to figurative ones. Thus, according to Cieslicka, in non-native idiom comprehension, the literal meaning enjoys a processing advantage over the figurative meaning. However, perhaps, it is not surprising that upon hearing the word *hatchet* there is a strong facilitation for the word *axe*, since the two words are strongly semantically related.

Underwood et al. (2004) used an eye-movement paradigm to investigate the on-line processing of idioms. They compared fixation count and fixation durations for the terminal word of an idiomatic phrase (*honesty is the best policy*) and a sentence containing the same lexical item (*it seems that his policy of …*). A significant processing advantage was found for native participants; with fewer and shorter fixations made in the idiom condition compared to the novel one. For the non-native speaker group, no such differences were observed. Although informative in terms of idiomatic vs. novel language processing, the study does not deal with the issue of idiomatic vs. literal meaning processing.

Finally, Conklin and Schmitt (2008) conducted a self-paced moving-window reading experiment to investigate idiom comprehension by native and proficient non-native speakers when a highly biasing story context preceded the idiom. It was found that idioms were read more quickly than novel phrases by both groups of participants. Further, they observed no processing differences between figurative and literal meaning processing for either natives or non-natives. Because the same pattern of results was observed in both participant groups, the authors concluded that idiom comprehension in non-native speakers is similar to that in native speakers. One downside of the study is that a within-subject design was used, which meant that each participant read both figurative and literal uses, as well as the novel phrase.

Research on whether non-natives process idioms faster than matched novel strings is mixed. However, unlike native speakers, the above findings are suggestive of the fact that even highly proficient speakers may experience difficulties processing idioms used figuratively. In the current experiment, we will further investigate idiom processing in a non-native population.

IV The present study

In the study, we monitored eye movements as native and non-native speakers read a series of stories. The stories contained one of the following: an idiom used figuratively (*at the end of the day* – 'eventually'), an idiom used literally (*at the end of the day* – 'in the evening'), or a novel phrase (*at the end of the war*).

1 Hypotheses and predictions

Based on previous findings in the literature, we have a set of predictions regarding the processing of idioms used figuratively and literally, and novel phrases. With regards to the native speaker group, we hypothesized that, first, our participants should show a processing advantage for idioms over matched novel phrases, as previous research showed that familiar expressions are read faster than novel strings. Second, we hypothesized that native participants should read the idioms more quickly when they are used figuratively than when they are used literally, as these idioms occur more frequently in their figurative form than literal one.

Although we did not explicitly manipulate contextual constraints (i.e. presence or absence of a biasing context), it was possible that our preceding disambiguating story context would influence the pattern of results. Biasing context has been found to influence the processing of ambiguous lexical items (Rayner et al., 1999; Vu et al., 2003). In their discussion of the 'reordered access model' of lexical access, Rayner et al. (1999) hold that context can potentially boost the activation of the less frequent meaning of an ambiguous word such that it becomes available to interpretive processes before the more frequent meaning. Thus, if context supports the higher frequency figurative meaning of an idiom, it should be processed quickly. If the context supports the low frequency literal meaning of an idiom, activation of this meaning will be boosted, and thus it will also be processed quickly.

If non-native speakers process idioms in a similar manner to natives, we would expect non-natives to exhibit a similar processing advantage for idioms over novel strings of language. If, however, no such advantage for idioms is found, this will imply that non-native speakers' processing of idioms differs from that of natives. A further question is whether non-natives demonstrate a processing advantage for the more frequent literal interpretation of an idiom in the presence of a highly constraining context. Previous research has shown that, in the absence of such a context, non-native speakers experience difficulties comprehending figurative meanings (Matlock and Heredia, 2002; Cieslicka, 2006). We will, thus, be able to discuss our results in light of those reported in literature.

2 Participants

Thirty-six native and 36 non-native speakers of English took part in the study. All participants were students at the University of Nottingham. Native speakers received course credit for their participation, while the non-natives were paid a small fee. The non-natives had learned English in a classroom setting and came from a wide range of language backgrounds. Prior to commencing their degree at the University of Nottingham, they met the minimum language requirements (IELTS 6.0 or TOEFL 550). It is important to note that these are minimum entry requirements, and many students would have had a considerably higher score than this. Since the exam, they had, on average, nearly two years of intensive language contact in the UK in a university setting. Finally, all students were successfully completing undergraduate or postgraduate degrees at the time when the study was conducted. Thus, these participants can be classified as highly proficient

Table 1 English language proficiency for non-native speakers: means (minimums and maximums in parenthesis), $n = 36$

Age (years of age)	Time in UK (months)	1st contact (years of age)	Speaking*	Reading*	Writing*	Comprehension*
22.5 (18–31)	20 months (3–72 months)	7 (4–12)	3.7 (3–5)	4.1 (3–5)	3.7 (3–5)	3.9 (3–5)

Notes: * 1 = very poor; 2 = weak; 3 = ok; 4 = good; 5 = excellent

speakers. A description of the non-native participants, as well as their self-rating of English language proficiency is summarized in Table 1.

3 Materials

The idioms used in the study were chosen using the following criteria. First, they had to be frequent English expressions (the frequency of the idioms used in the experiment is discussed in Norming Study 1 below). Second, it was necessary for the idioms to be able to be used figuratively, as well as literally, and sound natural in both conditions. Third, matched novel phrases had to be constructed, which could be as close to the idiom (in the form) as possible. Following the above three criteria, 53 candidates were provisionally selected, which were then subjected to the following norming procedures.

a Norming study 1: To ensure that our non-native participants knew the idioms in the experiment, we compiled a test with 77 idioms. Fifty-three of these idioms were those selected earlier on, whereas the remaining 24 were low frequency unusual filler idioms (e.g. *egg on your face*). This test was given to a group of 20 non-native participants who were students at the University of Nottingham, UK. The participants were asked to indicate how familiar they were with the idioms by rating their knowledge on a four-point scale, ranging from 1 'I don't know the idiom' to 4 'I know the idiom'. On the basis of the results obtained, 21 idioms (out of the previously selected 53 idioms) were shown to be well known to the participants and hence were selected for the study (they received an average rating of 3.5).

As can be seen in Appendix 1, most of the selected idioms were frequent. It is, however, worth noting that, due to idiom length, idiom frequency cannot be compared with that of single words; 50 occurrences in the British National Corpus (BNC) would be deemed high frequency for an idiom, but not for a word. A few idioms that were somewhat infrequent were, nevertheless, found to be well known to non-native speakers according to the rating study (e.g. *you can't judge a book by its cover*). Finally, a BNC search showed that all of the 21 idioms appear figuratively much more frequently than they do literally (the ratio of approximately 83% to 17%).[2]

b Norming study 2: Previous research suggests that the status of an idiom as decomposable or non-decomposable plays a role in its processing (Gibbs and Nayak, 1989; Gibbs et al., 1989; Titone and Connine, 1999; however, for an opposing view, see Tabossi et al., 2009). Following the procedure established by Gibbs and Nayak (1989), we asked

14 native speakers (who did not participate in the on-line reading experiment) to judge whether the individual components of the idiom made some unique contribution to the phrase's figurative meaning. Out of the 21 idioms, 12 were judged as decomposable and nine as non-decomposable (Appendix 1). We then ensured that each of the three presentation lists contained exactly four decomposable and three non-decomposable idioms (for more information on presentation lists, see below).

c Norming study 3: Because of the contention of Cacciari and Tabossi (1988) that idioms have an idiomatic key, one of the aims of the study was to explore idiom processing before and after the idiomatic key, or the recognition point. To determine the point at which our expressions are recognized as idiomatic, we created five versions of a sentence completion task, which included 65 sentence fragments presented out of context, 21 of which were target idioms while the rest were novel distracters. For example, for the idiom *leave a bad taste in your mouth*, Version 1 contained the shortest fragment: 'leave'. Version 2 had a slightly longer fragment: 'leave a bad'. Version 3 had longer still – 'leave a bad taste' – and so on. The test was given to 50 native speakers of English (10 participants per version), who were asked to complete the phrases. Because McFalls and Schwanenflugel (2002) considered a sentence to be high constraint if at least 70% of participants completed it correctly, we adopted the same threshold for our recognition point. It is noteworthy that out of the 21 idioms used in the study, seven did not reach the threshold of 70% before the final word of the idiomatic expression and were hence excluded from the recognition point analysis (but were present in the full idiom analysis). The mean probability of the remaining 14 items to be completed idiomatically was found to be 86.5% with the completion range being 70% and 100% (Appendix 1).

Finally, a short story was written for each target. Because idioms used literally and figuratively, as well as novel phrases, had different meanings, they required different stories in order to ensure the context was sufficiently biasing (Appendix 2).

4 Apparatus and procedure

Stories were presented across three presentation lists. Each list contained 21 items: seven idioms used figuratively, seven idioms used literally, and seven novel phrases. Decomposable and non-decomposable idioms were evenly distributed across the three presentation lists. We also ensured that no participant saw more than one version of the same phrase.

The participants were asked to read the stories quickly but for comprehension and were advised that each story would be followed by a comprehension question. Following this, a nine-point grid calibration procedure was completed. The first three trials were always practice trials. The eye-tracker was calibrated at least four times during the experiment. The stories were presented in a pseudo-randomized order. Before each trial, a fixation point appeared in the middle of the screen. After participants fixated it and a calibration check was done, a story appeared on the screen. Once participants finished reading each story, they pressed a key to proceed to the comprehension question. Eye movements were monitored using an EyeLink I eye-tracker.

After the experiment, all non-native participants were asked to rate their knowledge of the 21 idioms on a four-point scale, resulting in the same familiarity rating as in the

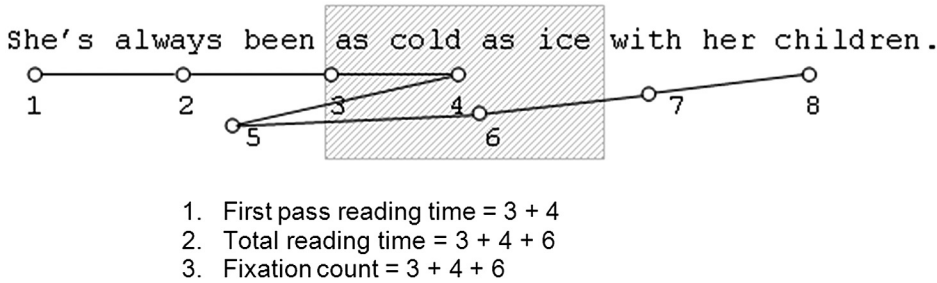

Figure 1 Hypothetical eye movement record: Shaded area represents the region of interest

norming study described above. We can thus be certain that all target idioms were well known to our non-native participants.

5 Analyses and results

Prior to the analysis, all trials where track loss occurred were removed. The missing data accounted for about 0.2% of the data. The participants had no difficulty answering comprehension questions, with an overall accuracy rate of 91.4% for native speakers and 90.4% for non-native speakers. One non-native participant was excluded from the analysis due to a high number of incorrect answers (a third were incorrect). For each target, we examined the following measures (exemplified in Figure 1):

1. First pass reading time: the sum of all fixation durations made within a region of interest before exiting either to the left or to the right (also known as 'gaze duration').
2. Total reading time: the sum of all fixation durations made within a region of interest.
3. Fixation count: the number of all fixations made within a region of interest.

It is generally assumed that early measures (first pass reading time) are sensitive to early processes in the comprehension of a text, such as early integration of information. Late measures (total reading time and fixation count) are believed to be sensitive to later processes associated with comprehension of a text, such as information re-analysis and discourse integration (Rayner et al., 1989; Paterson et al., 1999). The data were analysed in two different ways: analysis of the entire phrase (full idiom analysis), and analysis with regards to the idiom's recognition point (analysis before and after the recognition point).

6 Full idiom analysis

In the full idiom analysis, we analysed reading times and fixations for the entire phrase. The data were analysed using repeated measures ANOVAs treating participants and items as random variables. The dependent variables were mean total reading time, first pass reading time, and fixation count.

Table 2 Native and non-native fixation durations (in milliseconds) and fixation count for figurative and literal idiom uses, and novel phrases in the full idiom analysis with Standard Error (SE) in parenthesis

	Figurative	Literal	Novel
First pass reading time:			
Natives	447 (25)	454 (21)	497 (30)
Non-natives	743 (44)	705 (38)	720 (42)
Total reading time:			
Natives	514 (32)	507 (25)	628 (37)
Non-natives	937 (52)	817 (37)	880 (44)
Fixation count:			
Natives	2.8 (0.2)	2.7 (0.1)	3.2 (0.1)
Non-natives	4.2 (0.1)	3.7 (0.1)	3.9 (0.1)

Table 3 Analyses of variance and planned comparisons with participants (F_1) and items (F_2) as random variables for native speakers in the full idiom analysis

	By participants			By items		
	df	F_1	p	df	F_2	p
First pass reading time:						
Phrase type	2,70	2.0	ns	2,40	1.4	ns
Total reading time:						
Phrase type	2,70	9.6	**	2,40	6.3	**
Figurative vs. novel	1,35	14.1	**	1,20	8.0	*
Literal vs. novel	1,35	12.8	**	1,20	10.1	**
Figurative vs. literal	1,35	.05	ns	1,20	.032	ns
Fixation count:						
Phrase type	2,70	5.6	**	2,40	3.5	*
Figurative vs. novel	1,35	8.2	*	1,20	4.3	*
Literal vs. novel	1,35	8.8	**	1,20	6.9	*
Figurative vs. literal	1,35	.16	ns	1,20	.09	ns

Notes: * $p \leq .05$, ** $p \leq .005$, ns = non-significant

7 *Native speakers*

Means for the three eye-tracking measures can be found in Table 2 and the statistical comparisons in Table 3. No significant main effect of Phrase Type was found in the early measure, first pass reading time, and thus no planned comparisons were conducted. There was a significant main effect of Phrase Type by participants and items for total reading time and fixation count. Planned comparisons for the two late measures revealed that idioms used figuratively and literally were read significantly faster and elicited fewer fixations than novel phrases. No significant difference was found in the figurative vs. literal comparison in either of the late measures.

Table 4 Analyses of variance and planned comparisons with participants (F_1) and items (F_2) as random variables for non-native speakers in the full idiom analysis

	By participants			By items		
	df	F_1	p	df	F_2	p
First pass reading time:						
Phrase type	2,68	.45	ns	2,40	.29	ns
Total reading time:						
Phrase type	2,68	3.5	*	2,40	3.4	*
Figurative vs. novel	1,34	1.3	ns	1,20	0.7	ns
Literal vs. novel	1,34	2.4	ns	1,20	3.8	ns
Figurative vs. literal	1,34	6.7	*	1,20	8.4	*
Fixation count:						
Phrase type	2,68	4.2	*	2,40	3.5	*
Figurative vs. novel	1,34	2.3	ns	1,20	1.2	ns
Literal vs. novel	1,34	2.0	ns	1,20	1.0	ns
Figurative vs. literal	1,34	8.0	*	1,20	11.8	**

Notes: * $p \leq .05$, ns = non-significant

Overall, these results indicate a processing advantage for idiomatic expressions over novel strings. More importantly, they show that the two meanings – literal and figurative – are processed with a similar speed. One interesting finding is that the difference between idioms and novel language emerged late. We will come back to this in the general discussion.

8 Non-native speakers

Means for the three eye-tracking measures can be seen in Table 2 and statistical comparisons in Table 4. There was no significant main effect of Phrase Type in the first pass reading time analysis. Because no significant main effect was found, no further comparisons were conducted. However, in the total reading time and fixation count analyses, a significant main effect of Phrase Type was observed across participants and items. Planned comparisons revealed no differences in figurative vs. novel, or literal vs. novel processing, suggesting that both meanings of idioms were read at the speed comparable with novel language. More importantly, planned comparisons showed that an idiom's figurative meaning was processed significantly slower than the literal one, even though the context supported this meaning.

The above results indicate that for non-native speakers, novel phrases are not processed any slower than figurative or literal meanings of idioms. Further, there is clear evidence that the figurative meaning of an idiom is processed more slowly than the literal one. Interestingly, similar to native speakers, the above differences were observed in the late but not early measures.

9 Comparing native and non-native speaker performance

In order to assess the role of proficiency on idiom processing more directly, we conducted further ANOVAs on combined native and non-native data with Participant

Table 5 Analyses of variance and planned comparisons with participants (F_1) and items (F_2) as random variables with participants (proficiency) as a between factor in the full idiom analysis

	By participants			By items		
	df	F_1	p	df	F_2	p
First pass reading time:						
Phrase type	2,138	.718	ns	2,80	.601	ns
Proficiency	1,69	40.6	**	1,40	177.0	**
Phrase type × proficiency	2,138	1.2	ns	2,80	.728	ns
Total reading time:						
Phrase type	2,138	6.0	**	2,80	5.3	*
Proficiency	1,69	53.0	**	1,40	93.0	**
Phrase type × proficiency	2,138	5.1	*	2,80	4.0	*
Fixation count:						
Phrase type	2,138	5.0	*	2,80	3.8	*
Proficiency	1,69	39.0	**	1,40	54.9	**
Phrase type × proficiency	2,138	4.7	*	2,80	3.2	*

Notes: * significant at $p \leq .05$, ** significant at $p \leq .005$, ns = non-significant

Proficiency as a between factor (see Table 5). Overall non-native speakers took significantly more time to read all three types of stimuli, which was evidenced by a highly significant main effect of Proficiency in both early and late measures. It is hardly surprising that native speakers are faster readers than non-native speakers. Importantly, we observed a significant main effect of Phrase Type, as well as a significant interaction between Phrase Type and Proficiency in the two late measures. This suggests that not only are non-native speakers overall slower than natives, but that the nature of their processing differs. Namely, where native speakers tend to slow down (reading novel strings compared to idioms), non-native speakers do not. On the other hand, where non-natives show a significant processing cost (figurative renderings vs. literal ones), natives do not.

10 Recognition point analysis

The recognition point analysis was performed on two idiom portions, before and after the recognition point. Because our novel phrases had no recognition point, they were not included in the recognition point analysis. Thus, the analyses reported below include two types of stimuli: idioms used figuratively and literally.[3]

11 Native speakers

The data were analysed using repeated measures ANOVAs. The mean total reading time, first pass reading time, and fixation count for idioms used figuratively and literally before and after the recognition point are given in Table 6. Statistical comparisons are illustrated in Table 7.

None of the three measures we looked at showed any processing differences in figurative vs. literal idiom interpretations before or after the recognition point. This finding

Table 6 Native and non-native fixation durations (in milliseconds) and fixation count for figurative and literal idiom uses in the recognition point analysis with Standard Error (SE) in parenthesis

	Before recognition point		After recognition point	
	Figurative	Literal	Figurative	Literal
First pass reading time:				
Natives	284 (19)	270 (13)	220 (8.6)	221 (11)
Non-natives	413 (27)	399 (22)	385 (20)	375 (15)
Total reading time:				
Natives	299 (22)	299 (15)	200 (12)	214 (13)
Non-natives	526 (38)	444 (22)	424 (27)	371 (17)
Fixation count:				
Natives	1.6 (0.10)	1.5 (0.07)	1.1 (0.06)	1.2 (0.05)
Non-natives	2.3 (0.12)	2.0 (0.09)	1.9 (0.11)	1.7 (0.06)

Table 7 Analyses of variance and planned comparisons with participants (F_1) and items (F_2) as random variables for native speakers before and after the recognition point

	By participants			By items		
	df	F_1	p	df	F_2	p
First pass reading time: Figurative vs. literal:						
Before	1,35	.492	ns	1,13	.832	ns
After	1,35	.003	ns	1,13	.048	ns
Total reading time: Figurative vs. literal:						
Before	1,35	.000	ns	1,13	.178	ns
After	1,35	.952	ns	1,13	.404	ns
Fixation count: Figurative vs. literal:						
Before	1,35	.558	ns	1,13	.627	ns
After	1,35	.744	ns	1,13	.267	ns

Note: ns = non-significant

does not support the proposition of Cacciari and Tabossi (1988). However, this is not surprising given that in the current study, the context biased the reader to the upcoming idiom interpretation, while Cacciari and Tabossi's theory makes predictions for idioms outside of a biasing context.

Overall, it is clear that the native speaker group processes the two idiom meanings in a very similar way. In this, these findings replicate those obtained in the full idiom analysis discussed above. Crucially, we observed that there was no speed-up for the figurative rendering after the recognition point, where the difference was most likely to occur.

12 Non-native speakers

In the full idiom analysis, we established that non-native speakers slow down when reading idioms' figurative meanings. The aim of the recognition point analysis was to determine

Table 8 Analyses of variance and planned comparisons with participants (F_1) and items (F_2) as random variables for non-native speakers before and after the recognition point

	By participants			By items		
	df	F_1	p	df	F_2	p
First pass reading time: Figurative vs. literal:						
Before	1,34	.243	ns	1,13	.485	ns
After	1,34	.230	ns	1,13	.007	ns
Total reading time: Figurative vs. literal:						
Before	1,34	5.9	*	1,13	6.4	*
After	1,34	5.2	*	1,13	2.5	ns
Fixation count: Figurative vs. literal:						
Before	1,34	6.8	*	1,13	6.4	*
After	1,34	3.5	=.07	1,13	1.1	ns

Notes: * $p \leq .05$, ns = non-significant

where this occurred. Mean reading times before and after the recognition point can be found in Table 6 and statistical comparisons in Table 8.

Similar to the full idiom analysis, the early measure revealed no reliable differences before or after the recognition point for figurative and literal uses of idioms. Before the recognition point, both late measures showed that figurative uses were read more slowly than literal ones. After the recognition point, the only significant difference observed was that for the total reading time measure in the analysis by participants but not items. Fixation count data suggested marginally significant differences between the two idiom meanings in the analysis by participants but not items.

Taken together, the recognition point results confirmed what was suggested previously in the analysis of the entire idiom; namely the idiom's figurative meaning incurs a processing cost when compared to its literal equivalent. Importantly, with the help of the recognition point analysis, it was possible to establish that non-native speakers make reliably more and longer fixations when reading the figurative meaning before the recognition point has been reached.

13 Comparing native and non-native speaker performance

In order to explore the role of participant proficiency in a more direct way, we conducted further ANOVAs with Participant Proficiency as a between factor on combined native and non-native data (see Table 9). This analysis showed that, for all three measures, native speaker processing was significantly faster than that for non-natives before and after the recognition point, as evidenced by a significant main effect of Proficiency. Again, it is hardly surprising that native speakers read more quickly than non-native speakers. What is of greater interest is that before the recognition point, both late measures showed a significant main effect of Phrase Type. After the recognition point, no significant main effect of Phrase Type was found in any of the measures. The interaction between Phrase Type and Proficiency in both *F1* and *F2* analyses was not found to be significant.

Table 9 Analyses of variance and planned comparisons with participants (F_1) and items (F_2) as random variables with participants (Proficiency) as a between factor in the recognition point analysis

	By participants			By items		
	df	F_1	p	df	F_2	p
First pass reading time: Before:						
Phrase type	1,69	1.1	ns	1,26	1.2	ns
Proficiency	1,69	34.8	**	1,26	96.0	**
Phrase type × proficiency	1,69	0.0	ns	1,26	0.0	ns
First pass reading time: After:						
Phrase type	1,69	1.400	ns	1,26	0.00	ns
Proficiency	1,69	72.600	**	1,26	138.80	**
Phrase type × proficiency	1,69	.377	ns	1,26	.024	ns
Total reading time: Before:						
Phrase type	1,69	5.0	*	1,26	4.4	*
Proficiency	1,69	46.8	**	1,26	91.7	**
Phrase type × proficiency	1,69	2.9	ns	1,26	2.2	ns
Total reading time: After:						
Phrase type	1,69	3.1	ns	1,26	1.1	ns
Proficiency	1,69	69.8	**	1,26	187.4	**
Phrase type × proficiency	1,69	3.8	*	1,26	2.9	ns
Fixation count: Before:						
Phrase type	1,69	7.2	*	1,26	5.0	*
Proficiency	1,69	31.4	**	1,26	36.0	**
Phrase type × proficiency	1,69	2.0	ns	1,26	1.0	ns
Fixation count: After:						
Phrase type	1,69	1.8	ns	1,26	.388	ns
Proficiency	1,69	49.3	**	1,26	71.5	**
Phrase type × proficiency	1,69	4.2	*	1,26	1.3	ns

Notes: * significant at $p \leq .05$, ** significant at $p \leq .005$, ns = non-significant

V General discussion

In this study, we looked at the processing by native and proficient non-native speakers of idioms in a story context that encouraged either a figurative or literal interpretation. This was compared to the processing of matched novel phrases. The study had four aims. First, in native speakers, we wanted to confirm previous findings that idioms are processed faster than matched novel phrases. Second, we wanted to explore whether there are any processing differences between figurative and literal idiom renderings encountered in a biasing story context. Our third goal was to compare the processing of idioms' literal and figurative meanings before and after the recognition point. Finally, we aimed to compare idiom comprehension in a first and second language.

Although in native speakers, we did not observe significant differences in the early measure, we found a processing advantage for idioms like *at the end of the day* over novel phrases such as *at the end of the war* in the two late measures. This indicates that

compared to novel phrases, idiomatic expressions are read faster and require less re-reading and re-analysis. This finding is highly compatible with the existing research, which suggests facilitation for idioms vs. novel phrases (Swinney and Cutler, 1979; Gibbs, 1980; Gibbs and Gonzales, 1985; Tabossi et al., 2009).

Much of the idiom research in the past has focused on how the figurative and literal meanings available in ambiguous idioms are activated in relation to each other. For native speakers, none of the measures, early or late, showed a processing advantage for figurative idiom uses over their literal equivalents. This suggests that the preceding disambiguating context was sufficient to resolve the ambiguity that may have arisen during processing. Because none of the theories of idiom processing make specific claims about the effect of a biasing context, it is difficult to compare our studies to the existing models. The fact that for native speakers, no differences were observed in terms of fixation duration or fixation count for the literal and figurative meanings indicates that, in a biasing context, the activation of both meanings occurs comparably quickly. As this pattern of activation is in part driven by the presence of a preceding disambiguating context, further study is needed to investigate how idiom activation is modulated by the presence or absence of the disambiguating region. It is possible that when disambiguating context is not provided, the figurative meaning will be activated prior to the literal one because it is more frequent. With regards to the recognition point analysis, after the recognition point (i.e. after the expression has been recognized as idiomatic), it seemed probable that the figurative interpretation might be read more quickly than the literal one. However, the results clearly indicated that before, as well as after, the recognition point both idiom uses were read with the same speed by the native participants.

With respect to contextual constraints, one particular study is of relevance. Colombo (1993) investigated the role of context in the activation of figurative and literal idiom meanings. In a series of lexical decision tasks, it was found that the idiomatic meaning of an ambiguous idiom became activated only following the context that biased the figurative interpretation. In the absence of a figurative-biasing context (when a neutral or literally-biasing context was provided), only literal computations were observed. It is important to note that the two meanings of the idioms used in Colombo's study were equally frequent. The question, however, remains about how context modulates the activation of figurative and literal meanings when one meaning is more frequent than the other.

Due to a range of findings in the literature on non-native speakers, it was unclear whether, like natives, they would too process idioms faster than novel language. We found that, unlike native speakers, the non-native group's processing of idioms and novel phrases was very similar. Both early and late measures showed that idioms were processed with the same speed as novel phrases; no significant differences were found in the figurative vs. novel or literal vs. novel comparisons. These processing differences suggest that idioms are not represented in the mental lexicon of a non-native speaker in the same way they are represented in the lexicon of a native speaker. Our non-native results are in contrast with those of Conklin and Schmitt (2008), who found that both figurative and literal meanings had a robust processing advantage over novel phrases. However, as has already been mentioned, the difference in the results between the two studies may be due to limitations with Conklin and Schmitt's experimental design. Our non-native speaker results seem to be in agreement with those reported in Underwood et al. (2004).

Similar to the present study, they did not observe any processing advantage for idioms over novel phrases for non-native speakers. However, because Underwood et al. only measured reading times for the terminal word of idioms and novel phrases, their results provide an incomplete picture of idiom processing. Our analyses of the whole phrase – as well as before and after the recognition point – provide a broader picture of idiom processing and indicate where re-reading and re-analysis occur.

As has been discussed throughout, one of the key issues in idiom processing is when the two meanings of idioms are activated relative to each other. The eye-tracking measures reveal that in non-native speakers, figurative meanings required more re-reading and re-analysis than literal ones. These findings are in line with those reported in Cieslicka (2006), who observed that literal meanings were activated prior to figurative ones by her non-native participants. Our findings also support those of Matlock and Heredia (2002), who looked at the processing of phrasal verbs with a figurative meaning (*Paul went over the exam with his students*) vs. identical verb-preposition combinations used literally (*Paul went over the bridge with his bicycle*). They found that natives accessed idiomatic phrasal verbs more quickly than identical verb-preposition combinations used literally. For the non-native group, on the other hand, no differences were observed in reading times for phrasal verbs used figuratively vs. verb-preposition combinations used literally.

The main rationale behind the recognition point analysis was to find out where, in the course of idiom comprehension, the processing cost associated with the figurative meaning is greatest for non-native speakers. Both late measures showed that non-natives spent significantly more time reading the figurative meaning of an idiom than the literal one before the recognition point. After the recognition point, the total reading time and fixation count measures revealed a significant and marginally significant processing cost, respectively, for the figurative meaning. However, because this was observed in the analysis by participants but not items, further work is needed to ascertain if the figurative meaning continues to cause processing difficulty after the recognition point has been reached.

It is important to consider why non-native speakers require more processing effort when reading the figurative meaning of idioms even in the presence of a biasing context. Researchers agree that a fundamental task in second language vocabulary acquisition is building connections between a form and meaning (Van Patten et al., 2004; Schmitt, 2008). If our non-native speaker group has not yet developed strong form–meaning connections between an idiom and its figurative meaning, they will not show the same pattern of idiom processing as the native speaker group. Let us consider the idiom *at the end of the day*. The non-native speakers undoubtedly have connections between the individual lexical items and their meanings. The post-experiment test also clearly showed that our participants knew that these items occur together in the idiomatic phrase *at the end of the day*. Our finding of slow reading times for *at the end of the day* when used figuratively, thus, suggests that the link between the idiom and the meaning ‘eventually’ is not as strong as the link between the form and the meaning of the individual lexical items. As a result, the figurative meaning ‘eventually’ is not activated as quickly as the literal ‘in the evening’. Thus, the meaning with the highest level of activation is the incorrect one in contexts where the figurative meaning is the appropriate interpretation.

In order to better understand the overall pattern of results observed in the current study, we need to consider three factors: frequency, predictability, and context. We will look at each of these in turn. First, an idiom's figurative meaning is almost always more frequent than its literal counterpart. As evidenced by the BNC, all of the idioms used in the present study appear figuratively more frequently than they do literally (approximately 83% vs. 17%).[2] Researchers have suggested that (at least in native speakers) idioms are more readily understood figuratively than literally because they occur figuratively with much higher frequency in everyday discourse (Gibbs, 1986; Popiel and McRae, 1988; Van Lancker-Sidtis, 2003). Thus, based on a frequency account alone, processing should be faster for the figurative use of an idiom than for its literal counterpart. However, the results of the full idiom analysis as well as the recognition point analysis showed that this was not the case.

The second factor that may be implicated is predictability. Idioms – or at the minimum the words after the recognition point – can be considered to be highly predictable, as indicated by their high cloze probability in Norming Study 3. What this means is that readers can predict *day* after having seen *at the end of the*. On a predictability account, idioms used both figuratively and literally should be processed faster than novel language. However, because an idiom's completion is equally predictable in both literal and figurative phrases, this factor cannot be used to hypothesize which meaning of an idiom should be activated more quickly.

It is clear from the above discussion that frequency and predictability cannot satisfactorily account for the pattern of results observed in our study, and therefore a third factor – context – may be implicated. A number of word recognition studies have shown that if a preceding context creates strong enough expectancies, then the processing of the low frequency form of a word that has multiple interpretations may be processed equally as fast as its high frequency equivalent (Vu et al., 1998; Martin et al., 1999). In a self-paced reading task, Martin et al. (1999) showed that in the presence of a strongly biased context, reading times for the less frequent meaning of a homophone (*bulb* – 'the root of a plant') did not differ from those of the more frequent meaning (*bulb* – 'light bulb'). In the absence of a strongly biasing context, the more frequent meanings were read faster than less frequent ones. Because the story contexts used in our study biased readers towards either literal or figurative renderings, we believe that our results suggest that in native speakers, if a preceding context is strong enough, it facilitates the processing of the less frequent literal form. Non-native speakers, on the other hand, were not able to use context to effectively boost what should be the more frequent meaning.

Finally, one last issue merits attention. Although it is apparent that the patterns of idiom activation in native and non-native speakers are rather different, there appears to be one thing that the two groups have in common: the absence of any significant differences in the early measure (when the differences *are* significant in the late measures). As such, this finding has important implications in terms of the nature of eye-tracking measures and their significance for multi-word sequences. It appears that, upon initial reading (i.e. during the first pass reading time), figurative and literal idiom uses, as well as novel strings, are all read in a comparable way. However, there seems to be a need for a reader to exit the region of interest (to the left or to the right) and then come back to it, resulting in

significant differences across conditions in late measures, because some items require longer re-reading and/or re-analysis than others. It is thus possible that when reading longer strings of language (as opposed to single words), early measures may not be sensitive to potential differences. Previous research has shown that length manipulations for individual words affect early measures differently from later ones (Hyona, 1993; Rayner and Well, 1996). Unfortunately, the present study does not allow us to be more certain or specific with regards to the differences between early and late measures. However, the fact that this trend was apparent in both natives and non-natives (whose idiom comprehension was otherwise found to be rather distinct) does suggest that early eye-tracking measures may not be adequate for investigating long multi-word sequences.

To conclude, our analyses of figurative and literal idiom uses, as well as novel phrases, reveal a number of findings. First, proficient non-native speakers do not process idioms more quickly than novel phrases. Crucially, non-natives require more time to retrieve figurative senses of idioms than literal ones, even when the context biases the reader towards the figurative interpretation. This slow-down is largely evident before the recognition point. With respect to native speakers, the present study further confirms previous findings that idioms are read faster than novel language. Finally, in the presence of a preceding disambiguating context, native speakers do not process the low frequency literal meaning of an idiom any differently from the high frequency figurative one.

Acknowledgements

This research was supported by an ESRC grant PTA-030-2006-00151 to Anna Siyanova-Chanturia and a British Academy Grant LRG-42426 to Kathy Conklin. We would like to thank Walter van Heuven for his support in conducting the experiment and for his very helpful discussions. We are also grateful to SR Research Ltd (Canada) for their help with the Experiment Builder and Data Viewer software.

Notes

1 Not all idioms have a literal interpretation (e.g. *shoot the breeze*). Such idioms were not investigated here.

2 Output from the BNC was rated as being either figurative or literal. For four idioms that have more than 100 occurrences in the BNC, we only looked at the first 100 of them (e.g. *on the other hand* appeared 5,311 times), while for the rest of the idioms (those that have fewer than 100 occurrences), we looked at every instance.

3 The recognition point identification was done out of context, and was thus taken to be the same both for literal and figurative meanings. It is possible that when the biasing context precedes an idiom, the recognition point shifts forward. If the recognition point did shift, it should have shifted for both meanings. However, for the current idioms, it is unlikely that the recognition point would shift forward, as only one or two content words would remain, leaving many likely completions (e.g. *at the* is unlikely to be predictive of *at the end of the day*).

References

Arnon I and Snider N (2010) More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language* 62: 67–82.

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: 241–48.

Bobrow S and Bell S (1973) On catching on to idiomatic expressions. *Memory and Cognition* 1: 343–46.

Cacciari C and Tabossi P (1988) The comprehension of idioms. *Journal of Memory and Language* 27: 668–83.

Cieslicka A (2006) Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research* 22: 115–44.

Colombo L (1993) The comprehension of ambiguous idioms in context. In: Cacciari C and Tabossi P (eds) *Idioms: Processing, structure, and interpretation.* Hillsdale, NJ: Erlbaum, 163–200.

Conklin K and Schmitt N (2008) Formulaic sequences: Are they processed more quickly than nonformulaic language by native and non-native speakers? *Applied Linguistics* 29: 72–89.

Cowie AP (1998) Phraseology: Theory, analysis, and applications. Oxford: Oxford University Press.

Gibbs R (1986) Skating on thin ice: Literal meaning and understanding of idioms in conversation. *Discourse Processes* 9: 17–30.

Gibbs R (1980) Spilling the beans on understanding and memory for idioms in conversation. *Memory and Cognition* 8: 449–56.

Gibbs R and Gonzales G (1985) Syntactic frozenness in processing and remembering idioms. *Cognition* 20: 243–59.

Gibbs R and Nayak N (1989) Psycholinguistic studies on the syntactic behaviour of idioms. *Cognitive Psychology* 21: 100–38.

Gibbs R, Nayak N, and Cutting C (1989) How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language* 28: 576–93.

Glucksberg S (1989) Metaphors in conversation: how are they understood? Why are they used? *Metaphor and Symbolic Activity* 4: 125–43.

Hyona J (1993) Effects of thematic and lexical priming on readers' eye movements. *Scandinavian Journal of Psychology* 34: 293–304.

Jurafsky D (2003) Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In: Bod R, Hay J, and Jannedy S (eds) *Probabilistic linguistics*. Cambridge: The MIT Press, 39–95.

Langacker R (1987) *Foundations of cognitive grammar: Volume 1*. Stanford: Stanford University Press.

Martin C, Vu H, Kellas G, and Metcalf K (1999) Strength of discourse context as a determinant of the subordinate bias effect. *The Quarterly Journal of Experimental Psychology* 52A: 813–39.

Matlock T and Heredia R (2002) Understanding phrasal verbs in monolinguals and bilinguals. In: Heredia R and Altarriba J (eds) *Bilingual sentence processing*. Amsterdam: Elsevier, 251–74.

McFalls E and Schwanenflugel P (2002) The influence of contextual constraints on recall for words within sentences. *American Journal of Psychology* 115: 67–88.

Mondini S, Jarema G, Luzzatti C, Burani C, and Semenza C (2002) Why is 'red cross' different from 'yellow cross'?: A neurophysiological study of non-adjective agreement within Italian compounds. *Brain and Language* 81: 621–34.

Paterson K, Liversedge S, and Underwood G (1999) The influence of focus operators on syntactic processing of short relative clause sentences. *The Quarterly Journal of Experimental Psychology* 52A: 717–37.

Pawley A and Syder F (1983) Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In: Richards JC and Schmidt RW (eds) *Language and Communication*. London, New York: Longman, 191–226.

Pollio H, Barlow J, Fine H, and Pollio M (1977) *Psychology and the poetics of growth: Figurative language in psychology, psychotherapy, and education*. Hillside, NJ: Lawrence Erlbaum.

Popiel S and McRae K (1988) The figurative and literal uses of idioms, or all idioms are not used equally. *Journal of Psycholinguistics Research* 17: 475–87.

Rayner K and Well A (1996) Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin and Review* 3: 504–09.

Rayner K, Binder K, and Duffy S (1999) Contextual strength and the subordinate bias effect: Comment on Martin, Vu, and Metcalf. *The Quarterly Journal of Experimental Psychology* 52A: 841–52.

Rayner K, Sereno S, Morris R, Schmauder A, and Clifton C (1989) Eye movements and on-line language comprehension processes. *Language and Cognitive Processes* 4: 21–50.

Schmitt N (2008) Instructed second language vocabulary learning. *Language Teaching Research* 12: 329–63.

Sinclair J (1991) *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sosa A and MacFarlane J (2002) Evidence for frequency-based constituents in the mental lexicon: Collocations involving the word *of*. *Brain and Language* 83: 227–36.

Swinney D and Cutler A (1979) The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behaviour* 18: 523–34.

Tabossi P, Fanari R, and Wolf K (2009) Why are idioms recognized fast? *Memory and Cognition* 37: 529–40.

Tabossi P and Zardon F (1993) The activation of idiomatic meaning in spoken language comprehension. In: Cacciari C and Tabossi P (eds) Idioms: Processing, structure, and interpretation. Hillsdale, NJ: Lawrence Erlbaum, 145–61.

Titone D and Connine C (1999) One the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics* 31: 1655–74.

Tomasello M (2003) Constructing a language: A usage-based theory of language acquisition. Cambridge/London: Harvard University Press.

Underwood G, Schmitt N, and Galpin A (2004) The eyes have it: An eye-movement study into the processing of formulaic sequences. In: Schmitt N (ed.) *Formulaic Sequences*. Amsterdam: John Benjamins, 153–72.

Van Lancker-Sidtis D (2003) Auditory recognition of idioms by first and second speakers of English. *Applied Psycholinguistics* 24: 45–57.

Van Lancker D and Kempler D (1987) Comprehension of familiar phrases by left- but not by right-hemisphere damaged patients. *Brain and Language* 32: 265–77.

Van Lancker D, Canter G, and Terbeek D (1981) Disambiguation of ditropic sentences: Acoustic and phonetic cues. *Journal of Speech and Hearing Research* 24: 330–35.

Van Patten B, Williams J, and Rott S (2004) Form–meaning connections in second language acquisition. In: Van Patten B, Williams J, and Rott S (eds) *Form–meaning connections in second language acquisition*. Mawah, NJ: Lawrence Erlbaum. 1–26.

Vu H, Kellas G, Paul S (1998) Sources of sentence constraint on lexical ambiguity resolution. *Memory and Cognition* 26: 979–1001.

Vu H, Kellas G, Petersen E, and Metcalf K (2003) Situation-evoking stimuli, domain of reference, and the incremental interpretation of lexical ambiguity. *Memory and Cognition* 31: 1302–15.

Wray A (2002) *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Appendix 1 Idiom frequencies and their recognition point shown by a slash

	Idioms and their recognition point	Frequency[a]	Percentage[b]	Decomposability[c]
1.	a breath | of fresh air	89	80	D
2.	a piece | of cake	70	70	N
3.	add fuel to | the fire	14	80	D
4.	as cold | as ice	24	90	D
5.	at the end | of the day	760	90	D
6.	kill two | birds with one stone	36	90	D
7.	leave a bad taste | in your mouth	13	90	D
8.	left in | the dark	17	20	D
9.	cut a long | story short	39	80	D
10.	not my cup | of tea	19	90	N
11.	on the other | hand	5311	100	D
12.	pain in | the neck	36	60	D
13.	put your foot | down	112	30	N
14.	ring a | bell	75	50	N
15.	see which way the | wind is blowing	23	60	N
16.	sick and | tired	58	90	N
17.	the other side of | the coin	63	20	D
18.	tie the | knot	48	90	N
19.	twist someone's | arm	36	90	N
20.	under your | nose	104	30	N
21.	you can't judge | a book by its cover	11	80	D

Notes: [a] Total frequencies were taken from the BNC (British National Corpus) and are given per 100 million words. The frequency includes permissible variants 'tie the knot', 'tied the knot', etc.). [b] The percentage of correct completions. [c] D = judged as decomposable, N = judged as non-decomposable

Appendix 2 Examples of stories and comprehension questions that followed them

Figurative

I had my younger brother and my sister-in-law over for dinner yesterday. They both have their degrees from Cambridge, whereas most of the people they work with have theirs from less well-known overseas and British universities. Personally, I think you can have the highest degree from the best university in the world, but at the end of the day it's your contribution to the society that matters, and not the name of the university you went to at all. Sadly, they didn't agree with me.
Did they both study at Cambridge?

Literal

After my second year at university, I moved house. When I started packing, I realized that I had a lot more stuff than I had when I moved in as a first-year student. The house I was moving to was next door to the house I was moving from, which was very handy. However, I still had to carry most of my stuff in small boxes from my old room to the new one. I had to make at least 50 trips so at the end of the day I was absolutely exhausted.

I'm hoping to stay at this house for at least another two years. I really don't want to move any more.
Is she planning to move again soon?

Novel

One of my granddads was an army officer for most of his life. Despite being an army guy, he's always been a very humane and kind person. He is also a very artistic and creative person. For example, one of his hobbies is writing poetry. He's a retired man now who served in Vietnam and who's been through many things in his life, so he's got plenty of things to write about. I know that at the end of the war he went on to teach students at the Military Academy. That was something he found particularly challenging but also rewarding in many respects.
Did he take part in the war?