

# CONTEXT MATTERS, FIGURATIVELY, FOR L2 READERS: EVIDENCE FROM SELF-PACED READING

Sara D. Beck and Andrea Weber

University of Tübingen  
[sara.beck@uni-tuebingen.de](mailto:sara.beck@uni-tuebingen.de), [andrea.weber@uni-tuebingen.de](mailto:andrea.weber@uni-tuebingen.de)

## ABSTRACT

In a self-paced reading study, we investigated the extent to which non-native speakers use biasing context in idiom processing, and whether idiom literalness limits these effects as it does in native speakers. Idioms with a high potential for literal interpretations (e.g., *break the ice*) and a low potential for literal interpretations (e.g., *lose one's cool*) were embedded into sentences biasing literal or figurative readings and followed by resolutions that were either congruent or incongruent with these expectations. Context significantly impacted figurative interpretations, but not literal ones. While the reading time patterns replicate those of native speakers, literalness did not significantly interact with contextual effects as in L1 readers, particularly in how it impacts literal readings. The results highlight differences between L1 and L2 competencies.

**Keywords:** non-native speakers, idioms, context

## 1. INTRODUCTION

Idioms require speakers to be able to interpret both a literal and a figurative interpretation with ease in real-time processing. Although some studies have provided evidence that non-native (L2) speakers can process idioms similarly to native (L1) speakers, [5], others report findings that vary [17]. Not only must L2 speakers be familiar with an idiom in order to interpret it figuratively [18], but speakers must also learn to deal with additional complexities such as linguistic context in order to decide whether a literal or figurative interpretation is appropriate. In addition to context, idioms vary greatly in their potential for a literal interpretation (literalness). An idiom like *break the ice*, clearly interpretable in both a literal and a figurative (i.e., to relieve social tensions) sense, has high-literalness compared to a low-literalness idiom like *lose one's cool*, which has a less clear literal sense compared to its figurative one (to lose control of one's emotions). While L1 experience with the language and individual idioms has given these speakers the tools to integrate

these cues into online processing strategies, it's unclear whether L2 speakers also have these same abilities.

In a self-paced reading study, Beck and Weber [4] investigated the effects of biasing context in L1 idiom processing and asked whether idiom literalness limited contextual effects. Following the basic design of a study on highly literal phrasal verbs by Holsinger and Kaiser [12], high- and low-literalness idioms were embedded into sentences with contexts biasing either a literal or a figurative interpretation, and followed by resolutions that were congruent or incongruent with these expectations (e.g., [The new schoolboy/the chilly Eskimo] just wanted to *break the ice* [with his peers/on the lake] ...). They found that reading times for figurative and literal interpretations were faster where context and resolutions were congruent in high-literalness idioms. By contrast, low-literalness idioms showed only a benefit for figurative interpretations compared to literal ones, regardless of context. Thus, for native speakers, context is an important cue for idioms with a higher potential for literal interpretations (i.e., high-literalness), and a figurative interpretation is more salient for low-literalness idioms. This outcome suggests that L1 experience deems whether or not context should be considered in meaning integration.

The current study is a replication of [4] for highly-proficient L2 readers. In particular it asks: (1) how sensitive L2 readers are to linguistic context in idiomatic processing, and (2) whether idiom literalness limits these effects. In answering these questions, this study looks into possible differences in L1 and L2 idiom processing and their abilities to integrate linguistic and lexical cues (i.e., context and literalness) in online processing.

### 1.1. Context and Literalness Effects

As in L1 research, studies have shown that context is beneficial to L2 figurative meaning activation, e.g. [8], [14]. However, several studies have suggested that, even in the presence of biasing context, literal meaning is activated more quickly than figurative meaning, e.g. [7], [17], and possibly prioritized due to its salience, e.g. [11]. In line with L1 research, e.g. [12], Cieřlicka, Heredia, and

Olivares [9] compared the effects of context in an eye-tracking study on dominant and non-dominant bilinguals in high-literality idioms. They found that biasing context impacted the fixations and reading times of figurative and literal interpretations, but determined that non-dominant bilinguals did seem to show a literal preference, in contrast with dominant bilinguals. As the current study investigates high-proficiency L2 speakers, they may behave like non-dominant bilinguals and prefer literal readings regardless of context. This result would indicate a clear difference from our L1 speakers.

Literality has been investigated in only a limited number of L2 studies, e.g. [7], and so far no differences between the idiom types have been found to impact L2 processing. However, the scarcity of such studies does not allow for clear predictions. Thus, following the L1 results from [4], we expect that L2 readers are either a) too inexperienced with idioms and do not show any differences between the idiom types or b) able to take literality as a salient cue that limits the impact of context.

### 3. SELF-PACED READING STUDY

We conducted a phrase-by-phrase self-paced reading study with three manipulations: idiom *literality* (high- or low-literality idioms), *biasing context* (literal or figurative bias), and *resolution-type* (literal or figurative resolution). We expected our results to highlight possible differences in L1 and L2 abilities to integrate context and use idiom literality in their online processing strategies.

#### 3.1. Method

##### 3.1.1. Participants

Forty-seven native speakers of German (40 female; average age of 24.48, SD = 3.01) were paid to take part in the study. All were highly proficient in English (5 years minimum education), with an average self-reported proficiency of 6.05 (SD = 0.68) on a 7-point scale.

##### 3.1.2. Materials

The experiment consisted of 22 target and 78 filler trials. Target trials began with a biasing context (noun phrase + relative clause) followed by the infinitive form of the idiom, then by a resolution (prepositional phrase) congruent or incongruent with contextual expectations, and ended with two additional short phrases shared across all conditions. Phrases were controlled for letter length and average word frequencies, with only minimal

differences between conditions (see analysis section for more information). See example in Table 1.

**Table 1:** Example stimuli.

Biasing Subject		Idiom	Resolution	Resolution+1	
[The new schoolboy ...]	just wanted	<i>break the ice</i>	[with his peers]	<i>sooner than later</i>	...
[The chilly Eskimo ...]	to		[on the lake]		

Eleven low- and 11-high literality idioms were included. All idioms were taken from [4]; were pre-tested and normed [6] based on familiarity (L1 and L2), literality, length, and constituent frequency; had the same syntactic structure (to-infinitive verb + determiner + noun); and were short and unpredictable (i.e., the final word is necessary in order for idiom recognition where no context is present). Any minor differences were accounted for by including these variables in the analysis.

Sentences were also normed for strength of biasing context and plausibility of the sentences. While there were differences between items [see 4], the norming values for each item were included as fixed variables in the final analysis, where warranted. [All stimuli are listed here.](#)

Three regions of interest were identified for our analysis: the Idiom, the Resolution, and the following phrase (Resolution + 1), labeled in Table 1. Following [4], we expected the effects of (mis-) matching contexts to show up in the Resolution + 1 region. Since this region is identical across conditions for each idiom, observed differences in reading times cannot be attributed to differences in lexical items or syntactic structure.

##### 3.1.3. Procedure

The experiment was programmed and executed using E-Prime [15], and a subsequent idiom recognition test (to ensure participants knew the idioms) and a language background questionnaire were completed using Adobe Acrobat. Participants sat in a lab setting and wore noise-cancelling headphones. Responses were recorded with the spacebar of a keyboard.

The study began with four practice trials and included a self-timed break at the half-way point. A standard moving-window phrase-by-phrase presentation was used in which each phrase was masked by hyphens corresponding in length to the phrase to be presented. Only one phrase was seen at a time. Phrase-by-phrase rather than word-by-word was used in order to better mimic natural reading patterns and avoid a forced incremental processing pattern, e.g. [13], which might directly affect the questions at hand. 30 of the 90 reading

trials (including all target trials) were followed by multiple choice comprehension questions, and 10 were followed by arithmetic problems.

### 3.2. Results

We used R [16] and lme4 [3] for linear mixed effects analyses of the relationships between *biasing context* (figurative, literal), *resolution type* (figurative, literal), and *idiom literality* (high-, low-literality) on reading times in three areas of interest: *idiom*, *resolution*, and *resolution + 1* regions. Only trials with correctly answered comprehension questions were analyzed. One participant was excluded for poor performance on the comprehension questions (56%). Two idioms (*play the field* and *turn the tide*) were also excluded as they were unknown to more than half of participants. Based on a visual inspection of all reading times, outliers beyond overall minimum and maximums were excluded per section as reading times less than 200ms (all regions) and greater than 1600ms (*idiom*, 4.4%), 1800ms (*resolution*, 4.4%), and 1700ms (*resolution + 1*, 3.8%), respectively. See Table 2 for means in each region.

**Table 2.** Mean Reading Times (ms), all regions.

Region	Resolution type	Figurative Context		Literal Context	
		high-lit.	low-lit.	high-lit.	low-lit.
		RTs	RTs	RTs	RTs
Idiom	-	770	740	771	765
Resolution	figurative	781	818	773	874
	literal	767	804	754	794
Resolution+1	figurative	685	738	752	770
	literal	763	827	716	802

Each region was analyzed individually on normalized log-transformed values. With the exception of the idiom section (details below), independent measures and fixed factors, numerically centered around 0, were coded and included as follows: *biasing context* (figurative: 0.5, literal: -0.5), *resolution type* (figurative: 0.5, literal: -0.5), and *idiom literality* (high-literality: 0.5, low-literality: -0.5). Theoretically relevant factors were kept in the models, and items and participants were included as random factors with random slopes. A maximally justified random effects structure was determined for each region by step-wise selection and model comparison, e.g. [2] consulting RePsychLing [1]. Additionally, fixed factors of *trial order*, *region length* (number of letters), *average lexical frequency*, and norming values for *idiom familiarity*, *plausibility*, and *strength of context* were also included in full models, and they were eliminated from

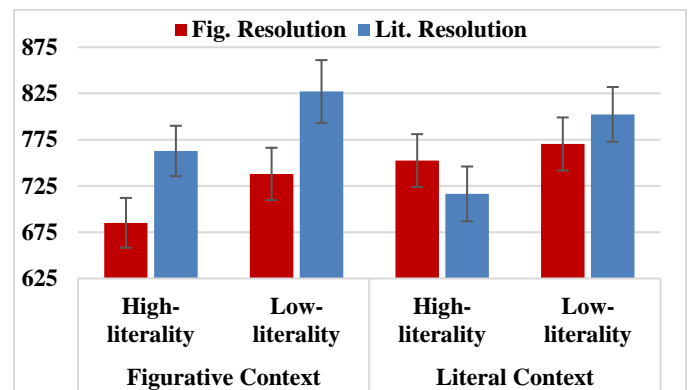
the models if backward step-wise selection showed that they did not improve the model. The analyses for each section are discussed in the following, with a focus on the resolution + 1. [Model outputs can be found here.](#)

#### 3.2.1. Analysis/Discussion

The model for reading times for the idiom region did not include *resolution type* or the norming factor for *plausibility* as the ending had not yet been encountered at this point. Significant effects of *region length* ( $\beta = .031, t = 2.33, p < .05$ ) and *trial order* ( $\beta = -.06, t = -5.48, p < 0.001$ ) suggest that longer idioms were read more slowly and reading times improved during the experiment. These effects are consistent throughout the successive regions but did not contribute to the overall findings other than to justify their inclusion in the models. They will therefore not be discussed further in the subsequent regions.

In the *resolution* region, significant effects of *region length* ( $\beta = .028, t = 2.20, p < .05$ ), *trial order* ( $\beta = -.05, t = -5.01, p < 0.001$ ), and *literality* ( $\beta = -.07, t = -2.70, p < 0.05$ ) were present. The effect of literality shows the pattern that resolutions following high-literality idioms were read more quickly than those following low-literality idioms. This is both consistent with L1 results and in line with research showing faster reading times for idioms with higher literality, e.g. [10].

**Figure 1:** Resolution+1 mean reading times (ms).



Finally, in the *resolution+1* region, main effects were found again for *region length* ( $\beta = .055, t = 3.08, p < .01$ ), *trial order* ( $\beta = -.05, t = -4.99, p < 0.001$ ), and *literality* ( $\beta = -.08, t = -2.33, p < 0.05$ ). Additionally, a marginal effect of *resolution-type* ( $\beta = -.04, t = -1.89, p = 0.07$ ) and a significant interaction of *biasing context* by *resolution-type* ( $\beta = -.11, t = -2.65, p < 0.05$ ) were found. In order to better interpret the data in the presence of this interaction, the data were split by biasing context for further analysis.

The additional analyses revealed that, for figurative contexts, figurative endings were significantly faster than literal ones ( $\beta = -.10, t = -4.07, p < 0.001$ ), but this effect is not present for literal contexts and endings. The result for figurative contexts reflects the L1 data in [4], but the literal context analysis does not. A close look at Figure 1, however, indicates that the same pattern exists, in which congruently-resolving phrases for high-literality idioms appear faster than incongruent, whereas for low-literality idioms, figurative resolutions are preferred overall. While this result is not significant in the L2 data here, the smaller effect size is likely due to a lower proficiency, a conclusion in line with the differences shown in [9].

#### 4. CONCLUSION

The current study asked (1) how sensitive L2 readers are to linguistic context in idiomatic processing, and (2) whether idiom literality limits these effects. We conclude that L2 readers do integrate linguistic context in idiomatic processing, particularly for figurative interpretations. For literal interpretations, our results are inconclusive, nor did we find significant evidence that literality impacts this process. However, based on the patterns present in low-literality idioms, and its similarities to previous L1 evidence [4], we conclude that this lack of finding may be due to proficiency. Furthermore, we would expect that as proficiency increases, so too will language experience, both which should be reflected in processing patterns.

These results are in line with previous L2 research that indicates that context matters for figurative interpretations [17], but we maintain that our results do not provide evidence for a literal priority that is distinct from L1 processing [7]. Rather, like L1 readers, experience with the language can indicate when context is more or less relevant in reaching an interpretation. Though, more studies with direct L1 and L2 comparisons are needed to investigate these claims.

#### REFERENCES

1. Baayen, R. H., Bates, D., Kliegl, R., & Vasishth, S. (2015). RePsychLing: Data sets from Psychology and Linguistics experiments.
2. Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. <https://arxiv.org/abs/1506.04967>
3. Bates, D., Maechler, M., Bolker, B., & Walker S. (2015). lme4: Linear mixed-effects models using Eigen and S4: R Package. Retrieved from <http://CRAN.R-project.org/package=lme4>
4. Beck, S. D., & Weber, A. (submitted). Context and Literality in Idiom Processing: Evidence from Self-paced Reading.
5. Beck, S. D., & Weber, A. (2016a). Bilingual and Monolingual Idiom Processing Is Cut from the Same Cloth: The Role of the L1 in Literal and Figurative Meaning Activation. *Frontiers in Psychology*, 7, 1350. <https://doi.org/10.3389/fpsyg.2016.01350>
6. Beck, S. D., & Weber, A. (2016b). English-German Database of Idiom Norms (DIN). [Data file].
7. Cieřlicka, A. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research*, 22, 115–144. <https://doi.org/10.1191/0267658306sr2630a>
8. Cieřlicka, A., & Heredia, R. (2011). Hemispheric asymmetries in processing L1 and L2 idioms: effects of salience and context. *Brain and Language*, 116, 136–150. <https://doi.org/10.1016/j.bandl.2010.09.007>
9. Cieřlicka, A., Heredia, R., & Olivares, M. (2014). It's All in the Eyes: How Language Dominance, Salience, and Context Affect Eye Movements During Idiomatic Language processing. In M. Pawlak & L. Aronin (Eds.), *Second language learning and teaching. Essential topics in applied linguistics and multilingualism*, (pp. 21–41). Cham [u.a.]: Springer.
10. Cronk, B., & Schweigert, W. (1992). The comprehension of idioms: The effects of familiarity, literalness, and usage. *Applied Psycholinguistics*, 13, 131–146. <https://doi.org/10.1017/S0142716400005531>
11. Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 8(3), 183–206. <https://doi.org/10.1515/cogl.1997.8.3.183>
12. Holsinger, E., & Kaiser, E. (2013). Processing (non)compositional expressions: mistakes and recovery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 866–878. <https://doi.org/10.1037/a0030410>
13. Jegerski, J. (2014). Self-Paced Reading. In B. VanPatten & J. Jegerski (Eds.), *Second language acquisition research series. Research methods in second language psycholinguistics* (pp. 20–49). New York: Routledge.
14. Liontas, J. (2002). Context and idiom understanding in second languages. *EUROSLA Yearbook* 2, 155–185. <https://doi.org/10.1075/eurosla.2.11lio>
15. Psychology Software Tools, I. (2013). E-Prime. Pittsburg, PA. Retrieved from <http://www.pstnet.com>
16. R Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for statistical computing. Retrieved from <http://www.R-project.org/>
17. Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, 29, 72–89. <https://doi.org/10.1177/0267658310382068>
18. Titone, D. A., & Libben, M. R. (2014). Time-dependent effects on decomposability, familiarity and literal plausibility on idiom meaning activation: A cross-modal priming investigation. *The Mental Lexicon*, 9, 473–496. <https://doi.org/10.1075/ml.9.3.05tit>