Author Accepted Manuscript version of the paper by

Milburn, E., Vulchanova, M., & Vulchanov, V. (2021). Collocational frequency and context effects on idiom processing in advanced L2 speakers. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale*, 75(2), 169–174. https://doi.org/10.1037/cep0000214

Collocational frequency and context effects on idiom processing in advanced L2 speakers

Evelyn Milburn[1], Mila, Vulchanova[1], & Valentin Vulchanov[1]

Norwegian University of Science and Technology[1]

Draft Date: November 26, 2020

Correspondence to:

Evelyn Milburn

NTNU Dragvoll

7049 Trondheim

Trøndelag, Norway

evelyn.a.milburn@ntnu.no

ABSTRACT

Multiword expressions have attracted attention recently following suggestions that they are acquired chunk-wise by children in the first language, while adults learning a second language may focus more on individual words within an expression. This is of particular interest for the acquisition of idioms, which are multiword expressions in which the literal meanings of the component words do not (always) directly contribute to overall phrasal meaning, resulting in a figurative interpretation. Figurative meaning access is speeded both by idiom-internal characteristics, like higher collocational frequency, and idiom-external characteristics, like supportive contexts. We examined the relationship between the collocational frequency of idioms' component words and the context in which an idiom is embedded. In a visual world eye-tracking study, advanced non-native English speakers heard incomplete English phrases embedded within contexts that biased either literal or idiomatic continuations and saw images representing literal or figurative completions, or distractor images. Participants showed higher looks to figurative completions that were at odds with contextual bias, suggesting that integrating frequency information in context in adult L2 users may be overridden when a phrase is figurative. However, higher-proficiency participants showed more successful suppression of inappropriate figurative continuations. These results suggest that idiom conventionality when compared to literal phrases may be a stronger driver of predictive looks than collocational frequency or contextual bias alone, and that sensitivity to contextual fit when processing idioms may still be developing even among very advanced L2 users.

Keywords: Figurative Language, Idioms, Language Acquisition, Bilingualism, Eyetracking

Public Significance Statement: Adult second language users often have difficulty with figurative language such as idioms, possibly reflecting difficulty accessing conventionalized multiword sequences. The main finding was that knowledge about when an idiom fits the context may still be developing even among very advanced adult users. However, as users' proficiency grows, they are more able to process figurative phrases holistically instead of word-by-word.

INTRODUCTION

Increasing evidence demonstrates that multiword expressions constitute a large part of everyday language. These expressions—also called multiword chunks, fixed expressions, lexical bundles, or formulaic sequences, among others—are familiar sequences of words that occur with high frequency in language and which may be stored and retrieved as single units during comprehension (Appel & Trofimovich, 2017; Carrol & Conklin, 2020). They can be common binomials like *fish and chips*, speech routines like *don't worry about it*, or idioms like *take the wheel*, to name just a few. Although methods vary for estimating how much of spoken and written language is formulaic, estimates suggest that multiword expressions comprise one-third to one-half of all language (Conklin & Schmitt, 2012), and serve a wide variety of pragmatic functions, such as facilitating social interaction, organizing discourse, or efficiently expressing technical information. Using multiword expressions also confers cognitive benefits: comprehenders are sensitive to the frequency of multiword expressions, such that higher frequency phrases are processed faster (Arnon & Snider, 2010; Siyanova-Chanturia et al., 2011) and remembered more accurately (Tremblay et al., 2011). Using multiword expressions is additionally thought to aid fluency and speed processing by reducing load on working memory (Conklin & Schmitt, 2012).

This recent focus on multiword expressions as distinct units of language raises the question of how they are acquired and how their trajectory of acquisition may vary depending on characteristics of the learner. Adult users seem to use multiword expressions in their second and third languages in a distinctly nonnative way, producing fewer multiword expressions than do native speakers, relying on a small number of multiword expressions while ignoring the rest, and treating multiword expressions as more flexible than they actually are (Arnon & Christiansen, 2017).

In particular, second language learners often struggle with idioms. Idioms are conventionalized multiword expressions in which the literal meanings of the expression's component words do not (always) directly contribute to the meaning of the expression, resulting in a figurative or non-literal interpretation (Gibbs, 1980; Glucksberg, 1991). However, many idioms have well-formed literal meanings as well. For example, a phrase like *turn the tables* figuratively means "to move from a weaker position to a stronger one", although it can also be interpreted literally as "to rotate a table". Which of these possible interpretations is ultimately selected depends on a number of factors, particularly including the context in which the idiom is embedded (Titone & Connine, 1999).

This tension between literal and figurative processing may underlie the difficulty adult second language learners frequently have with learning idioms. Eye-tracking studies of reading show that L2 learners may have difficulty accessing figurative meanings. Thus, L2 learners processed idioms at similar speeds to novel compositional phrases, contrasted with faster figurative processing by native speakers (Siyanova-Chanturia, Conklin, & Schmitt, 2011; although see Conklin & Schmitt, 2008, for figurative processing advantages in proficient learners). Interestingly, this contrasts with findings that adult L2 learners may be sensitive to the frequency of common binomials (Siyanova-Chanturia et al., 2011). Taken together with research on L2 idiom comprehension, this suggests a picture in which adult L2 learners can recognize that an idiom is present, but may have difficulty switching from literal to figurative processing despite recognizing the necessity. Similar problems with suppressing the literal interpretation have been identified in figurative language processing in deficit populations (e.g., autism spectrum disorders, Vulchanova & Vulchanov, 2018; Vulchanova et al., 2019a).

A number of different factors can scaffold figurative processing. One strong influence on processing is the context in which an idiom is embedded (for an overview, see Vulchanova et al., 2019b). Contexts biasing literal and figurative interpretations can push idiom interpretation towards one meaning or the other, sometimes without any processing slowdown, at least in native speakers. For example, in a visual world eye-tracking study, Holsinger (2013) found that participants looked more at figurative probes when idioms appeared in figurative context, and at literal probes when idioms appeared in literal contexts. This suggests that context was driving interpretation towards a context-congruent meaning without resulting in processing disruption. Similar results have been found in eye-tracking during reading. Milburn and Warren (2019) preceded idioms with figuratively- and literally-biasing contexts, finding no differences in eye movement measures regardless of context. This suggests that participants were able to use context to immediately select the correct meaning of the idiom with little or no interference from the unbiased meaning. Context can therefore be a strong scaffold for choosing a particular idiom meaning among alternatives.

Critically, context is an external influence on idiom processing: it is not a part of the idiom and can be varied without changing the idiom itself. However, internal characteristics of an idiom also influence processing. In particular, idiom predictability – operationalized as the cloze probability of the final word in an idiom (Vulchanova et al., 2019b) – may aid idiom recognition and retrieval. Using this metric in a norming study, Bulkes and Tanner (2017) found that predictability positively correlated with idiom length: longer idioms provide more information to the comprehender, resulting in easier idiomatic completions. This finding has been borne out in experimental research. Fanari et al. (2010) found that the length of an idiomatic string was critical for how fast idiomatic meanings were computed: idiomatic

meanings were available at the offset of long idioms, but not short ones, as indexed by a lexical decision task. They suggested that participants were able to recognize longer idioms earlier than shorter idioms, speeding access of the figurative meaning and resulting in a fully active figurative meaning at idiom offset. In a more recent study, Cacciari and Corradini (2015) conducted a self-paced reading task in which predictable and unpredictable idioms were followed by literal and figurative disambiguating regions. They found faster reading times, indicating easier processing, when predictable idioms were followed by figurative continuations compared to literal: when an idiom was highly predictable the figurative meaning became active earlier in the idiom, and therefore was easier to integrate when the figurative continuation was reached. Conversely, they found faster reading times when unpredictable idioms were followed by literal continuations compared to figurative. They concluded that unpredictable idioms were recognized as idiomatic much later, potentially not until after the entire idiom had been read. This means that processing of an unpredictable idiom was primarily literal, with the idiomatic meaning only being available for a very short time before the disambiguating region was reached. When continuations were literal, the briefly-active figurative meaning was easy to discard in favor of the much stronger literal meaning. When continuations were figurative, in contrast, the strongly-active literal meaning was difficult to discard in favor of the figurative disambiguation, resulting in processing slow-downs.

The current study asked whether advanced L2 users primarily relied on supportive context - an idiom-external influence - or predictability - an idiom-internal influence - to retrieve and produce idioms. Using the visual world paradigm, advanced non-native English speakers listened to sentences consisting of a biasing context clause and a final clause with the last word missing. This final clause could be completed to form either literal or figurative phrases; whether

literal or figurative was appropriate depended on the context clause. We also included a neutral context condition. We predicted that participants would look more towards idiomatic completions when given an idiomatic context, and towards literal completions when given a literal context (Holsinger, 2013; Milburn & Warren, 2019). However, given previous findings in L2 idiom retrieval and recognition, it was also possible to find looks to inappropriate continuations depending on context. We hypothesized that looks to the figurative completion when context was literal would indicate that idiom predictability is a stronger driver of processing than contextual bias.

METHODS

PARTICIPANTS

48 participants ranging in age from 18 to 35 participated in the study. Participants began learning English at an average age of 6.06 years (self-report), were native speakers of Norwegian who had grown up in Norway, had normal or corrected-to-normal vision, and did not participate in the stimulus norming described below. Ethics approval was received from the Norsk Senter for Forskningsdata (Norwegian Center for Research Data).

MATERIALS

Visual stimuli consisted of 21 arrays of four images arranged in a grid (see Supplemental Materials for a prototypical image array). For example, one prototypical array consisted of a table, a car, a key, and a flask. Images were selected from either the MultiPic databank (Duñabeitia et al., 2018) or the Bank of Standardized Stimuli (Brodeur et al., 2014).

Each array was accompanied by one of three possible sentence fragments, consisting of a biasing context clause and a final clause with the last word missing; participants were instructed to click on the image that best completed the sentence (see Kessler et al., 2020, for a similar

design). Final clauses could be completed to form either literal or figurative phrases (for example, “turn the” can be completed literally as “turn the car” or figuratively as “turn the tables”). We will refer to this final clause as the “idiom region”, although idiomatic interpretation was not always required. The words before the idiom region, referred to as the pre-idiom region, were the same across conditions, although the length of this region varied between items. Context clauses biased either a figurative completion or a literal completion, or were unbiased (1a-c; pre-idiom region underlined; idiom region bold). The unbiased condition usually consisted of the pre-idiom region and the idiom region only.

1a. (figurative bias) To get even with his enemies, Chia-Ming **turned the** ____

1b. (literal bias) To get to his destination, Chia-Ming **turned the** ____

1c. (unbiased) Chia-Ming **turned the** ____

All images in the array were possible continuations of the sentence but varied in how compatible they were with the biasing context. One image in the array—the table in Figure 1, referred to as the *figurative target*—was most compatible with the figuratively-biasing context (1a): the idiom “turn the tables” means to move from a weak position to a strong one. A second image—the car in Figure 1, referred to as the *literal target*—was most compatible with the literally-biasing context (1b). Finally, two images in the array—the key and the flask, referred to as the *high-cloze* and *low-cloze distractors*—were most compatible with the unbiased context (1c; although all images were possible continuations in this condition). See the Supplemental Materials for a list of all linguistic stimuli. In order to control for the role of frequency (see Vulchanova et al., 2019 for recommendations and a theoretical discussion), the figurative and literal targets were matched in collocation frequency using the Corpus of Contemporary

American English (COCA; Davies, 2009). See the Supplemental Materials for a description of the matching procedure.

NORMING

Stimuli were normed to ensure that idioms were familiar, that sentences were similarly natural across conditions, that picture stimuli were identifiable, and that contexts appropriately constrained selection of the final word. See Supplemental Materials for full norming details.

PROCEDURE

INDIVIDUAL DIFFERENCE MEASURES

Participants completed two assessments of English L2 skills. The first was LexTALE (Lemhöfer & Broersma, 2012), a test of vocabulary knowledge administered as an un-speeded lexical decision task. Participants also completed a 15-question multiple choice grammar test conducted through Exam English (examenglish.com). See the Supplemental Materials for details.

EYE-TRACKING

Participants' eyes were tracked using a desktop-mounted Eyelink 1000 tracker (SR Research Ltd., Toronto, Ontario, Canada). See Supplemental Materials for full eye-tracking details.

RESULTS

ACCURACY

Accuracy results can be seen in the Supplemental Materials.

EYE-TRACKING

Because we were interested in comparing gazes between different images within a condition rather than between conditions, we analyzed each condition separately. We compared looks to the Figurative Target to looks to the three other images in separate comparisons within each condition using centered treatment coding. In each comparison, the non-Figurative Target image

was treated as the baseline. Graphs of fixation proportions for all objects in the three conditions across the entire analysis window can be seen in Figures 2A-C.

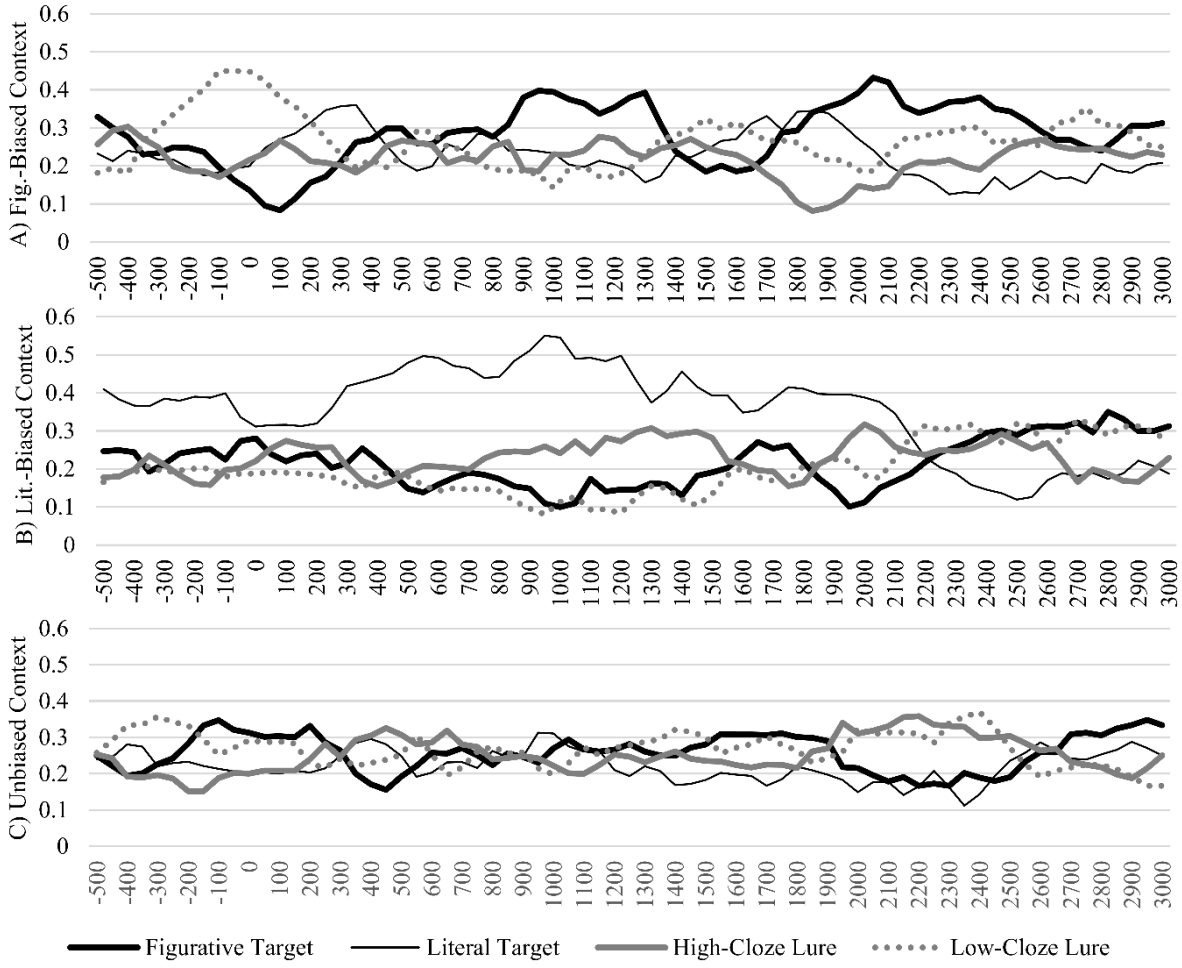

Figure 2A-2C: Fixation proportions (Y-axis) for all objects in the three conditions across the entire analysis window, measured from idiom onset.

Gaze data for these three comparisons were analyzed during a time window that began 500 ms before idiom onset and ended 3000 ms post-idiom onset. Eye gaze data were aggregated in nonoverlapping 50 ms time bins with 50 samples per bin using the littlelisteners package (Mahr, ver. 0.0.0.9000). Fixations preceding or following a blink were removed from analysis. Data were analyzed using Growth Curve Analysis (Mirman, 2016) in the R statistical computing

package (R Development Core Team, 2013; ver. 3.0.1). For model specifications and full model results, see the Supplemental Materials.

Figuratively-Biased Condition

Participants looked at the Figurative Target marginally more than the Literal Target across the time window (β=.06; SE=.03; t=1.97; p=.052). Looks to the Figurative Target increased marginally faster than did looks to the Literal Target (β=.44; SE=.25; t=1.74; p=.08). Participants looked at the Figurative Target more than the High-Cloze Distractor across the time window (β=.07; SE=.03; t=2.25; p<.05). There were no other significant effects.

Literally-Biased Condition

Participants looked at the Literal Target more than the Figurative Target across the time window (β=-.14; SE=.03; t=-5.25; p<.05). Looks to the Figurative Target increased more sharply than did looks to the Literal Target (β=.76; SE=.29; t=2.60; p<.05). Finally, looks to the Figurative Target compared to the Literal Target (β=.95; SE=.24; t=4.04; p<.05) and the High-Cloze Distractor (β=.53; SE=.19; t=2.80; p<.05) formed a steeper parabolic curve with increased looks at the beginning and end of the time window.

To investigate whether competition from the Figurative Target varied depending on participants' overall English proficiency, we computed the effect size of Figurative Target vs Literal Target and Figurative Target vs High-Cloze Distractor looks on the linear and quadratic time terms for each participant and correlated this with LexTALE score (Mirman, 2016). Participants with higher LexTALE scores tended to show shallower parabolic looking curves to the Figurative Target compared to the Literal Target (Pearson's r=-.20) and the High-Cloze Distractor (Pearson's r=-.24).

Unbiased Condition

There were no significant effects.

DISCUSSION

The current study asked whether idiom-internal or -external forces were stronger drivers of cued idiom retrieval for advanced L2 users. There were three primary findings. First, L2 users were successfully able to use context to choose appropriate literal or idiomatic continuations, as indicated both by their accuracy scores and by preferential looking towards the appropriate targets. Second, top-down idiom knowledge seemed to be the strongest driver of processing over and above contextual fit or cloze probability, as indicated by the growth curve analyses of the eye movement data. Finally, individual characteristics of users, such as L2 proficiency, may influence how easily they are able to suppress inappropriate figurative continuations when they are presented visually, although this result is exploratory and worthy of a more in-depth investigation.

L2 users successfully used context to choose appropriate literal or idiomatic continuations, although figurative processing appeared weaker than literal processing. In the Figuratively-Biased condition participants fixated the appropriate Figurative Target over the frequency-matched, but contextually inappropriate, Literal Target marginally more across the selected time window, and looks to the Figurative Target additionally increased marginally faster than did looks to the Literal Target. These marginal results point to potentially weakened figurative processing, and, in all likelihood, weakly consolidated and/or integrated L2 idiom knowledge. However, participants also fixated the Figurative Target more than the High-Cloze Lure across the time window, showing that they were able to easily ignore continuations that were locally high-probability, but contextually inappropriate. In the Literally-Biased condition,

participants fixated the Literal Target over the Figurative Target across the time window, again demonstrating appropriate use of context.

Top-down idiom knowledge seemed to drive processing over either contextual fit or cloze probability. This was particularly evident when examining looks to the Figurative Target in the Literally-Biased condition. The Figurative Target was a strong lure in this condition, as indexed by the steeper increase in looks to the Figurative Target compared to the Literal Target (note that this reflects the slope of the line of gaze proportions, not the absolute value). This suggests that participants were considering the Figurative Target as a viable continuation even when context biased a literal continuation, possibly indicating that the advanced users in this study, having overcome the bias towards literal/compositional processing seen in adult L2 learners (Siyanova-Chanturia et al., 2011), are now overextending use of L2 idioms into contextually inappropriate situations. When viewed together with the gaze patterns in the Figuratively-Biased condition, it appears that L2 users may have relatively stable representations of figurative expressions, but are unsure of their contextual fit, and are therefore less able to ignore contextually-inappropriate figurative continuations when explicitly presented. These results may suggest less stable integration rather than consolidation of L2 idiom knowledge (cf. Gaskell & Dumay, 2003). Of particular note is the finding of a steeper parabolic looking curve for the Figurative Target compared to the High-Cloze Lure in the Literally-Biased condition: both the High-Cloze Lure and the Figurative Target were contextually inappropriate, but varied in their local appropriateness, with the High-Cloze Lure being a much better choice based solely on cloze probability. If bottom-up sensitivity to frequency information was driving looks, we would expect more looks to the High-Cloze Lure, which was more locally appropriate than the Figurative Target. These results suggest that as proficiency increases, L2 users become more

able to process idioms as multiword sequences and bypass word-by-word literal processing. However, they are still less likely to integrate this knowledge with contextual constraints.

Finally, higher proficiency seemed to be associated with better suppression of inappropriate figurative continuations, although this analysis was exploratory and is worth a complete study to investigate fully. Participants with higher LexTALE scores tended to show shallower parabolic looking curves to the Figurative Target compared to the Literal Target and High-Cloze Lure in the Literally-Biased condition. This indicates that the Figurative Target was less competitive both with the contextually-appropriate literal continuation and the contextually-inappropriate, but locally appropriate high-cloze continuation, for participants with higher English proficiency. Critically, the participants in this study were very advanced English learners. It is noteworthy, then, that variability in their English proficiency nonetheless was somewhat predictive of gaze patterns. This finding is congruent with the finding that native and nonnative speakers focused on different things when explaining the meanings of unfamiliar idiomatic phrases (Wray et al., 2016): native speakers tended to invoke the context, whereas L2 learners frequently referred to the meanings of single words. However, L2 learners were more likely to use context and analogy to explain meanings as their knowledge of the language increased, pointing to improved holistic processing of the unfamiliar phrases and less focus on single words at the expense of phrasal meaning. Similarly, the more proficient participants in the present study seemed to focus more on the contextually-appropriate Literal Targets in the Literally-Biased condition, with less competition from the Figurative Target.

The design of the current study, because it presents idiom continuations visually, necessarily elicits literal activation of individual idiom components that might not be activated during normal idiom comprehension. Kessler et al (2020) addressed this issue in a similar study:

they played idioms with the final words missing and asked participants to select the appropriate continuation from a set of words displayed on a screen. They found increased looks to words that were semantically related to the correct idiom-final word compared to unrelated words (let the cat out of the BAG/BASKET/ARM/STOMACH), and pointed out that these looks could have been an epiphenomenal effect of the visual presentation. However, when they played participants the same idioms and continuations in an ERP experiment, they found graded N400 responses to the related and unrelated words, suggesting that the increased looks to the related words in the eyetracking were likely not entirely a result of bottom-up spread from sensory input. This highlights the importance of using converging experimental paradigms to unpick the processes underway during idiom comprehension.

The present study examined the relationship between the collocational frequency of idioms' component words and the context in which an idiom is embedded in English as a second language. The results suggest that idiom conventionality may be a stronger driver of predictive looks than collocational frequency or contextual bias alone, and that sensitivity to contextual fit when processing idioms may still be developing even among very advanced L2 users. Such results are also consistent with research on word learning in the L1 suggesting that item consolidation is the initial stage, whereby items are stored (and can be recognized), with lexical integration and associations with other items in the rest of the lexicon (including competitors) representing a second and final stage (e.g. Gaskell & Dumay, 2003). The current study also confirms that different types of behavioral measures, such as performance accuracy, and on-line measures, like gaze behavior, can reflect different stages of language processing and should be used in combination (e.g. Roberts & Siyanova-Chanturia, 2013).

References

Appel, R., & Trofimovich, P. (2017). Transitional probability predicts native and non-native use of formulaic sequences: Transitional probabilities and formulaic sequences. *International Journal of Applied Linguistics*, *27*(1), 24–43.

Arnon, I., & Christiansen, M. H. (2017). The Role of Multiword Building Blocks in Explaining L1-L2 Differences. *Topics in Cognitive Science*, *9*(3), 621–636.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*(1), 67–82.

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos. *PLoS ONE*, *9*(9), e106953.

Bulkes, N. Z., & Tanner, D. (2017). "Going to town": Large-scale norming and statistical analysis of 870 American English idioms. *Behavior Research Methods*, *49*(2), 772–783.

Cacciari, C., & Corradini, P. (2015). Literal analysis and idiom retrieval in ambiguous idioms processing: A reading-time study. *Journal of Cognitive Psychology*, *27*(7), 797–811.

Carrol, G., & Conklin, K. (2020). Is All Formulaic Language Created Equal? Unpacking the Processing Advantage for Different Types of Formulaic Sequences. *Language and Speech*, 63(1), 95–122.

Carrol, G., Littlemore, J., & Dowens, M. G. (2018). Of false friends and familiar foes: Comparing native and non-native understanding of figurative phrases. *Ling.*, *204*, 21-44.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers?. *App. linguistics*, *29*(1), 72-89.

Conklin, K., & Schmitt, N. (2012). The Processing of Formulaic Language. *Annual Review of Applied Linguistics*, *32*, 45–61.

Davies, M. (2009). The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *Int. journal of corpus linguistics*, *14*(2), 159-190.

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816.

Fanari, R., Cacciari, C., & Tabossi, P. (2010). The role of idiom length and context in spoken idiom comprehension. *European Journal of Cognitive Psychology*, *22*(3), 321–334.

Gaskell, M. Gareth & Nicolas Dumay. 2003. Lexical competition and the acquisition of novel words. *Cognition,* 8 (2). 105–132.

Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, *8*(2), 149–156.

Glucksberg, S. (1991). Beyond Literal Meanings: The Psychology of Allusion. *Psych. Sci*, *2*(3), 146–152.

Holsinger, E. (2013). Representing Idioms: Syntactic and Contextual Effects on Idiom Processing. *Language and Speech*, *56*(3), 373–394.

Kessler, R., Weber, A., & Friedrich, C. K. (2020). Activation of Literal Word Meanings in Idioms: Evidence from Eye-tracking and ERP Experiments. *Language and Speech*.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343.

McCauley, S. M., & Christiansen, M. H. (2017). Computational Investigations of Multiword Chunks in Language Learning. *Topics in Cognitive Science*, *9*(3), 637–652.

Milburn, E., & Warren, T. (2019). Idioms show effects of meaning relatedness and dominance similar to those seen for ambiguous words. *Psych. bulletin & review, 26(2), 591-598.*

Mirman, D. (2016). *Growth curve analysis and visualization using R*. CRC press.

Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition*, *35*(2), 213-235.

Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. J. B. (2011). Seeing a phrase "time and again" matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 776–784.

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 Lang. Res.*, *27*(2), 251-272.

Titone, D. A., & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics*, *31*(12), 1655–1674.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Lan. Lear.*, *61*(2), 569–613.

Vulchanova, M., & Vulchanov, V. (2018). Figurative language processing in autism. The interactive mind: Language, vision and attention (pp. 67–76).

Vulchanova, M., Milburn, E., Vulchanov, V., & Baggio, G. (2019b). Boon or Burden? The Role of Compositional Meaning in Figurative Language Processing and Acquisition. *Journal of Logic, Language and Information, 28(2), 359-387.*

Wray, A., Bell, H., & Jones, K. (2016). How native and non-native speakers of English interpret unfamiliar formulaic sequences. *European Journal of English Studies*, *20*(1), 47–63.