

Understanding Idiomatic Variation

Kristina Geeraert¹, R. Harald Baayen^{2,3} and John Newman^{3,4}

¹Department of Linguistics, KU Leuven, Belgium

²Department of Linguistics, University of Tübingen, Germany

³Department of Linguistics, University of Alberta, Canada

⁴School of Languages, Literatures, Cultures & Linguistics, Monash University, Australia

kristina.geeraert@kuleuven.be, harald.baayen@uni-tuebingen.de,
john.newman@ualberta.ca

Abstract

This study investigates the processing of idiomatic variants through an eye-tracking experiment. Four types of idiom variants were included, in addition to the canonical form and the literal meaning. Results suggest that modifications to idioms, modulo obvious effects of length differences, are not more difficult to process than the canonical forms themselves. This fits with recent corpus findings.

1 Introduction

Idioms have traditionally been regarded as multiword units whose meaning can not be derived from the meaning of its parts (Bobrow and Bell, 1973). This has led some researchers to claim that idioms are semantically opaque, that their structure is syntactically fixed, and they are stored whole as a ‘large word’. Thus, research investigating how idioms are understood has focused predominantly on the canonical form and how it differed from a literal paraphrase (Swinney and Cutler, 1979; Gibbs, 1980; Cacciari and Tabossi, 1988; Titone and Connine, 1999).

Recent corpus-based research however has shown that idioms can in fact occur with a range of variation (Moon, 1998; Barlow, 2000; Langlotz, 2006; Schröder, 2013). Idioms can undergo syntactic variation (e.g. *they have really bitten the bullet this time* and *her new-found reputation was a bubble that would burst*), be lexically varied (e.g. *throw/toss in the towel*, *miss the boat/bus*), truncated (e.g. [*he who pays the piper*] *calls the tune*), and even modified with adverbials or adjectives (e.g. *spill royal beans*, *pulling political strings*, *make rapid headway*). This variation can even occur with nondecomposable idioms (Duffley, 2013), such as *kick the bucket* (e.g. *no buckets have been kicked*, *when his parents kick their*

gold-plated bucket, and *my phone kicked the pail last week*). These studies have illustrated that idioms are not nearly as fixed or rigid in form as previously assumed.

Few studies have investigated idiomatic variation from an experimental perspective. Gibbs and colleagues (Gibbs et al., 1989; Gibbs and Nayak, 1989) explored lexical and syntactic variation of decomposable and nondecomposable idioms using a semantic similarity rating task. They found that decomposable idioms (i.e. idioms whose constituents contribute to the meaning of the whole, as in *pop the question*) were rated as more similar in meaning to a literal paraphrase than were nondecomposable idioms, or idioms whose constituents do not contribute meaning to the whole (e.g. *kick the bucket*). However, nondecomposable idioms can be modified in context while retaining their idiomatic meaning, as was demonstrated by Duffley (2013). Moreover, replication studies do not return consistent results. The role of decomposability has not proven to be a reliable measure, with participants performing at chance when classifying idioms into decomposability categories (Titone and Connine, 1994b; Tabossi et al., 2008). In addition, decomposable and nondecomposable idioms are not always found to be significantly different (Tabossi et al., 2008). Finally, semantic similarity has been shown to be largely predicted by the same local contexts as observed in corpora (Miller and Charles, 1991), suggesting that the semantic similarity measure collected in these studies simply reflected how interchangeable the variant is with its paraphrase and did not accurately reflect the comprehension of these variants.

Meanwhile, McGlone et al. (1994) explored the semantic productivity of idiom variation. Variants in this study produced a new idiomatic meaning based on the original (e.g. *shatter the ice*, from *break the ice*, meaning ‘to break an uncomfortable or stiff social situation in one fell swoop’).

Using self-paced reading, they measured the reaction time for participants to read the final sentence of a story, which contained idioms, variants, or literal paraphrases. They found that participants were significantly faster at reading the canonical form of the idiom, but that the variants were read as fast as the literal paraphrases. They suggest that canonical forms of idioms are accessed whole, but that variants are processed like literal language and are therefore processed slower. While the results show that modified idioms can be understood in context, they did not control for the type of variation. They used instances of lexical variation (e.g. *shatter the ice*), quantification (e.g. *not spill a single bean*), and even hyperboles (e.g. *it's raining the whole kennel*). Based on their findings, it is uncertain whether some types of variation are easier to comprehend than others.

The current study explores the processing of several types of variation, as well as the literal meaning of the idiom, through an eye-tracking experiment. Two research questions are explored: (1) are variants processed differently from the canonical form; and (2) are variants processed differently from each other. The first question plans to determine whether variants are still processed differently from the canonical form when the type of variation is controlled for (e.g. is lexical variation more difficult to comprehend than the original idiom?). Second, by including several types of variation and controlling for them, a comparison can be made between the different types of variants (e.g. Are there processing differences between, say, lexical variation and partial forms of an expression?).

While this experiment was largely exploratory, we did have some predictions about the results. For example, formal idiom blends are typically regarded in the literature as 'errors' or 'slips of the tongue' (Fay, 1982; Cutting and Bock, 1997). We therefore hypothesized that idiom blends would be more difficult to process due to this 'error-like' nature. Meanwhile, some idioms can occur in "idiom sets" or "clusters", such as *shake/quake/quiver in one's boots* or *down the drain/chute/tube/toilet* (Moon, 1998). We hypothesized that lexical variation would not be more difficult to understand than the canonical form. Lastly, partial or truncated forms of an expression have words omitted and should be faster to read, whereas additional adjectives inserted into the expression should take addi-

tional time due to the presence of an extra word.

The remainder of the paper proceeds as follows: We first describe the design of the experiment and the materials used. Next, we present the results, focusing on two areas of interest: the idiom as a whole and the altered word within the idiom. Finally, we discuss our findings and how they fit into the larger discussion on idioms.

2 Methodology

2.1 Materials

Sixty idioms were extracted from the Oxford Dictionary of English Idioms (Ayto, 2009) and the Collins COBUILD Idioms Dictionary (Sinclair, 2011). The form listed in the dictionary was regarded as the canonical form. If more than one form was listed then the form most familiar to the first author was used, as she spoke the same variety as the participants in the study. These idioms varied in length and syntactic structure: 20 three-word idioms consisting of a verb and a noun phrase (i.e. V-NP, e.g. *rock the boat*); 20 four-word idioms consisting of a verb and a prepositional phrase (i.e. V-PP, e.g. *jump on the bandwagon*); and 20 five- or six-word idioms (10 each) consisting of a verb, noun phrase, and a prepositional phrase (i.e. V-NP-PP, e.g. *hear something through the grapevine*). Two contexts were created for each idiom: one literal and one figurative (e.g. *I used to pretend I could talk to plants, and I would hear things through the grapevine* = literal; and *I used to be a socialite, and I would hear things through the grapevine* = figurative). Both contexts had identical final clauses, with the idiom in sentence-final position. As syntactic variation is possible with idioms (Moon, 1998; Schröder, 2013), the contexts in this study were not restricted to the present tense.

These idioms were manipulated for four types of variation within the figurative context (i.e. the context was identical for all variants), in addition to the canonical form. First, lexical variation, where one of the lexical items within the expression was altered to a synonymous or near-synonymous word (e.g. *discover something through the grapevine*). Synonyms were selected based on their naturalness in the context to convey a similar meaning.¹ Second, partial form of the idiom, where only a portion of the idiom was

¹An online thesaurus (<http://www.thesaurus.com/>) was often utilized for synonymous words.

presented, usually a key word or words (e.g. *use the grapevine*). In order for the sentence to still be grammatically correct, pronouns or lexically-vague words replaced the missing elements of the expression, such as *it, them, things* for nouns, or *have, be, do, use* for verbs. Third, integrated concept, where an additional concept was integrated into the idiom (e.g. *hear something through the judgemental grapevine*). These additional concepts expanded or emphasized the figurative contexts in which the idiom occurred. Finally, formal idiom blend, where two idioms were blended together (e.g. *get wind through the grapevine – blending hear something through the grapevine with get wind of something*). Each experimental idiom (i.e. the 60 idioms selected) was paired with a non-experimental idiom for use in the idiom blend condition. These “blending” idioms were chosen for their intuitive plausibility, but controlled for their syntax and semantics (Cutting and Bock, 1997).

Half of the idioms had the beginning portion of the expression altered (verb), while the other half had alternations made to the final portion of the expression (noun). In total, there are six conditions: one in a literal context and five in a figurative context (i.e. one canonical form and four variants). The experiment utilized a Latin-square design, where every participant saw each idiom once in one of the six conditions. Six versions of the experiment were created, each one containing 10 idioms in each of the six conditions.

CONDITIONS:

1. **Literal Meaning** of the idiom in its canonical form (e.g. *While the guys were reshingling, they suddenly went through the roof.*)
2. **Canonical Form** of the idiom in a figurative context (e.g. *Although these were new stocks, they suddenly went through the roof.*)
3. **Lexical Variation** of the idiom in a figurative context (e.g. *Although these were new stocks, they suddenly went through the ceiling.*)
4. **Partial Form** of the idiom in a figurative context (e.g. *Although these were new stocks, they suddenly went through it.*)
5. **Integrated Concept** within the idiom in a figurative context (e.g. *Although these were new stocks, they suddenly went through the investment roof.*)
6. **Idiom Blend** of two idioms in a figurative context (e.g. *Although these were new stocks, they suddenly went through the charts.*)

Since the “blending idioms” only occurred in one condition (i.e. Idiom Blend), they were used as fillers in their canonical form in the other five versions of the experiment, occurring in either a figurative or literal context. Each blending idiom was excluded as a control in the version of the experiment where it occurred in the idiom blend condition in order to avoid a bias in the materials. Therefore, in each version of the experiment, 10 of the blending idioms occurred in the idiom blend condition, while the remaining 50 appeared in their canonical form as fillers. Of these fillers, 20 occurred in a figurative context and 30 occurred in a literal context. This was done to increase the number of literal contexts in the experiment so that they were not so underrepresented. In sum, each participant saw 110 items: 60 experimental idioms (10 in each of the six conditions) and 50 blending idioms as fillers.

Finally, six practice sentences were created using six “practice” idioms. These idioms all occurred in their canonical form. Three were in a figurative context and three in a literal context. These were the same for all participants.

2.2 Procedure

This experiment used the Eye-Link 1000, desk-top mounted video-based eye-tracking device, manufactured by SR Research. The eye-tracker sampled the pupil location and size at a rate of 1000Hz, and was calibrated using a 9-point calibration grid. Calibration occurred at the beginning of the experiment, after the practice, and again after every 22 sentences, for a total of five blocks. The computer screen resolution was set to 1920 x 1080 pixels.

The stimuli were presented in two parts. Participants first saw the “context clause” (e.g. *Although these were new stocks,*), followed by the “idiom clause” (e.g. *they suddenly went through the roof.*) on a separate screen. Each trial began with a fixation cross presented for 1,000 msec on the left side of a light-grey screen. Next, they saw the context clause, also on a light-grey background, in a bold, black, Courier New 30-point font. Every clause was displayed in full and fit on one line. To exit this screen, participants had to trigger an invisible boundary in the bottom right corner. A blank, light-grey screen was presented for 1,000 msec before the fixation cross preceding the idiom clause appeared. The sequence of screens for the idiom clause was identical to the context clause.

Ten percent of the stimuli were followed by a true/false comprehension question, which pertained to the immediately preceding sentence, and were presented randomly throughout the experiment. Participants pushed one of two buttons on a game controller to answer these questions, which were clearly labelled on the question screen. The experiment began with a practice session, which consisted of six practice sentences and three questions. These were the same for all participants, although their order varied.

All participants had normal or corrected-to-normal vision. The right eye of each participant was tracked. Participants sat approximately 85cm from the computer screen, with the camera placed on the desk about 35cm in front of the computer screen. The participants sat in a sound-proof booth, while the experimenter sat outside the booth, running the experiment. The lights were kept on. The experiment was self-paced and took about 45 minutes to complete. Each participant was given an opportunity for a short break halfway through the experiment.

After the participants had completed the eye-tracking portion, they were asked to indicate their knowledge of each expression in a separate task. Each idiom appeared on the computer screen, in its canonical form, in a black, bold, 22-point Courier New font, centered on a white background. Above the idiom was the question “Do you know this expression?”; below were two boxes, one labelled ‘yes’ and the other labelled ‘no’. Using the mouse, participants clicked on the appropriate box to respond. The mouse repositioned itself to the center of the screen on each trial.

At the end of this second task, participants were presented with a few additional questions, pertaining to their idiom usage (e.g. How often do you use these expressions?; Do you like using these expressions?). Below each question was a Visual Analogue Scale (VAS), which is a continuous graphical rating scale (Hayes and Patterson, 1921; Freyd, 1923; Funke and Reips, 2012). Participants responded by clicking the mouse anywhere along the VAS scale. The scale was labelled with a ‘thumbs-up’ image on the right for a positive response and a ‘thumbs-down’ image on the left for a negative one. Lastly, participants were asked to rate the acceptability of seven prescriptively ‘incorrect’ sentences, shown below, using the same VAS scale. These sentences

attempted to elicit a measure of the participant’s flexibility with language and non-standard usage.

Language Questions (LQs):

1. The only option the school board has is to lay off a large *amount* of people.
2. Slot machines are thought to be more *addicting* than table games.
3. The document had to be signed by both Susan and *I*.
4. While cleaning the kitchen, Sally looked up and saw a spider on the *roof*.
5. I thought it could *’ve went* either way.
6. She *could care* less what he had to say about it.
7. You have to balance your life, *irregardless* of what anybody thinks.

2.3 Participants

Sixty University of Alberta linguistics undergraduate students participated in this experiment. All were native speakers of Canadian English. There were 43 female and 17 male participants, ranging from 17–29 years of age. Four participants were left-handed. All participants were reimbursed for their time with course credit.

3 Results

The results were analyzed using mixed-effects linear regression, using the `lme4` package (Bates et al., 2014) in R (R Core Team, 2014). Only the results for the Total Fixation Duration (i.e. the total amount of time spent fixating on the Area Of Interest, or AOI) will be discussed. We focus on two AOIs: the idiom as a whole (i.e. the summed fixations on all words within the idiom) and the altered word within the idiom (i.e. the synonymous word in lexical variation, the additional word in the integrated concept, the semantically vague ‘replacement’ word in partial forms, and the word from another idiom in the idiom blend). The analyses focus on the 60 experimental idioms. Further information about this study and the results can be found in Geeraert (2016).

Ten predictor variables appeared significant in the models. `Condition` is a factor indicating the type of variation with which the idiom occurred (e.g. lexical variation, partial form). `Length` specifies the number of words within the idiom’s canonical form. `PortionAltered` is a factor specifying which part of the idiom (i.e. beginning/verb or ending/noun) was manipulated in the

variant. `Trial` is the standardized order of presentation of the stimuli in the experiment. As the stimuli was presented randomly, this order will be different for each participant.

`MeanVariationRating` is a standardized mean measure of acceptability for the particular idiom with a specific type of variation. This measure was collected in a separate experiment, where participants were asked to rate the acceptability of the variants in the same contexts. These ratings were included to determine if participants' preferences influence their ease of comprehension.

As the decomposability classification is unreliable (Titone and Connine, 1994b; Tabossi et al., 2008), two measures reflecting the semantic contribution of the constituents were utilized instead. `meanTransparencyRating` is a standardized average measure of transparency for the idiom's meaning as a whole. These ratings were collected in a separate experiment, where participants saw each idiom, along with its definition and an example, and were asked to rate how obvious was the meaning of the expression. The average rating for each idiom was included as a predictor to determine whether the overall transparency of the idiom influences speakers' processing of variants. `LSA.Score.Paraphrase` is a measure of similarity using Latent Semantic Analysis (LSA), between the words in the idiom and its paraphrase (e.g. *spill the beans* 'reveal a secret'). This score was obtained from a pairwise comparison of two texts (i.e. an idiom and its paraphrase), which compares the local contexts in order to obtain a value of similarity (Landauer et al., 1998).² This measure allows us to control for the idiom's compositionality. If the exact words in the idiom have little to do with the expression's meaning, then the LSA score will be small (e.g. *cut the mustard* – 'be acceptable' = 0.07). But if the words used share meaning or contribute to the idiom's meaning, then the LSA score will be larger (e.g. *stop something in its tracks* – 'stop something' = 0.87).

As idioms are multi-word expressions, multiple frequency measures were obtained: the frequency of the idiom, frequencies of the individual words, and all possible combinations of adjacent words (e.g. `word1` and `word2`; `word2` and `word3`; `word1` and `word2` and `word3`). To avoid collinearity, a Principal Components Analy-

²The LSA scores were obtained from the English Lexicon Project (Balota et al., 2007), available at <http://lsa.colorado.edu/>.

sis (PCA) was conducted on these frequency measures. Only the first Principal Component (henceforth `PC1.logFrequency`) is significant.

Several participant-related variables are also significant. `KnowIdiom` is a factor indicating the participant's knowledge of each idiom (i.e. 'yes' or 'no'). `Gender` is a factor specifying whether the participant is male or female. Finally, a second PCA was conducted on the rating responses for the seven Language Questions (LQs) above. Only `PC2.LQ` was significant. This variable is used to reflect the participant's flexibility with language usage.

3.1 Idiom as Area of Interest

The first model examines the summed fixation durations on the idiom as a whole. The fixed effects for this model are shown in Table 1. `Condition` occurs in two significant interactions; the first, between `Condition` and `KnowIdiom`, is shown in the left panel of Figure 1. The canonical form, and the majority of variants, show the same general pattern: shorter fixation durations on known idioms. These variants (except integrated concepts) are therefore shown in grey, as they do not significantly differ from the canonical form. Partial forms however show a different pattern. Fixation durations on this variant are relatively similar regardless of whether the participant is familiar with the expression or not; thus a facilitation effect for knowing the idiom is not observed as it is with the other variants. This particular variant is fixated upon less than the canonical form, likely due to it being shorter in length (i.e. fewer number of words). This is in line with longer fixations observed on integrated concepts – an additional word is integrated into the idiom, making it longer in length and requiring additional fixations.

The second interaction, shown in the second panel of Figure 1, is between `Condition` and `Length`. The general pattern observed here is that longer idioms show longer summed fixation durations, as expected, due to the increased number of words in the idiom. Lexical variation, formal idiom blends, and the literal meaning of the idiom are not significantly different from the canonical form (shown in grey). The other two variants show a pattern that is significantly different from the canonical form. Idioms with integrated concepts show a slight inhibitory effect of length, where an additional concept is more

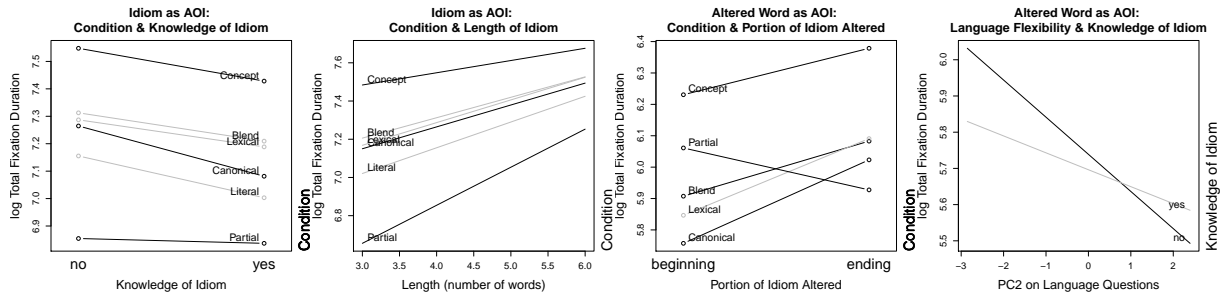


Figure 1: Interactions in the Mixed-Effects Linear Regression Models for the Summed Total Fixation Duration on the Whole Idiom and the Altered Word as an Area of Interest. Lines in grey represent factors levels which are not significantly different or slopes which are not significant.

difficult to integrate into shorter idioms (i.e. extra time is needed). However, partial forms of shorter idioms have even fewer words to fixate upon and therefore show considerably shorter fixation durations. In sum, durations on integrated concepts and partial forms are more comparable to the canonical form when the idiom is longer.³

Interestingly, the literal meaning of the idiom shows shorter fixation durations than the canonical form. These fixations are not significantly shorter ($t = -1.94$), but certainly trending towards significance. The literality of the expressions (Titone and Connine, 1994a) may be contributing to this result. Nevertheless, a general pattern is evident based on these two above interactions with `Condition`: variants of the same length as the canonical form are not processed significantly different from this canonical form.

The model presented in Table 1 also shows six main effects. Longer fixation durations are observed on the whole idiom if the beginning of the idiom (i.e. the verb) was altered. This is not dependent on the type of variation, but rather all variants are easier to process if the change comes later in the expression (see `PortionAltered`). This is a different result than that of Gibbs and colleagues (Gibbs et al., 1989; Gibbs and Nayak, 1989) who found no difference with ratings in whether the noun or verb was altered.

A significant main effect is also observed for `meanVariationRating`. Variants which re-

	Estimate	Std. Error	t-value
Intercept	6.71	0.09	75.97
Condition=Concept	0.49	0.10	5.04
Condition=Blend	0.08	0.10	0.75
Condition=Lexical	0.01	0.10	0.05
Condition=Literal	-0.19	0.10	-1.94
Condition=Partial	-0.75	0.16	-4.80
KnowIdiom=Yes	-0.18	0.04	-4.32
Length	0.11	0.02	6.76
PortionAltered=Ending	-0.06	0.02	-2.52
PC2.LQ	-0.07	0.03	-2.42
LSA.Score.Paraphrase	0.24	0.07	3.49
meanVariationRating	-0.06	0.01	-7.23
Gender=Male	-0.17	0.08	-2.17
Trial	-0.04	0.01	-3.78
I(KnowIdiom=Yes Condition=Concept)	0.06	0.05	1.16
I(KnowIdiom=Yes Condition=Blend)	0.08	0.06	1.42
I(KnowIdiom=Yes Condition=Lexical)	0.08	0.06	1.52
I(KnowIdiom=Yes Condition=Literal)	0.03	0.06	0.55
I(KnowIdiom=Yes Condition=Partial)	0.17	0.06	2.75
I(Length Condition=Concept)	-0.05	0.02	-2.62
I(Length Condition=Blend)	-0.01	0.02	-0.36
I(Length Condition=Lexical)	0.00	0.02	0.20
I(Length Condition=Literal)	0.02	0.02	1.04
I(Length Condition=Partial)	0.08	0.03	2.48

Table 1: Fixed Effects for the Idiom as AOI

ceived higher acceptability ratings are fixated on less long, suggesting preferred variants are easier to understand and interpret (or perhaps variants easier to interpret are preferred). Additionally, longer fixation durations appear on idioms which have higher LSA scores for the idiom’s paraphrase (i.e. `LSA.Score.Paraphrase`). This finding seems initially surprising, as previous analyses on the comprehension of idioms suggest that idioms are easier to understand when the individual components contribute meaning to the whole (Gibbs et al., 1989). However, the LSA scores indicate how similar the local contexts are between the idiom and its paraphrase (i.e. how interchangeable is the expression with its paraphrase). When the LSA score is high (i.e. the paraphrase is easily interchangeable) then looking time increases as the contexts are not distinctive for the idiom. But if the LSA score is low, then the idiom and its paraphrase are less interchangeable, making the context more distinctive and the idiom more predictable. Interestingly, `meanTransparencyRating` is

³`PC1.logFrequency` was also significant in the Idiom as AOI model. However, this variable is strongly correlated with `Length` ($r = -0.9$). This correlation is unsurprising given that `PC1.logFrequency` was created using adjacent co-occurrence frequencies. Model comparison shows that `Length` is the more significant predictor in this model, producing a considerably lower AIC value, and therefore was retained at the expense of `PC1.logFrequency`.

not significant. The degree to which the idiom is considered ‘obvious in meaning’ does not seem to influence the comprehension of idioms or variants.

A main effect was also observed for `PC2.LQ`, a latent variable representing the participants’ flexibility with language (i.e. the more they find non-standard or erroneous forms acceptable). Shorter fixations are observed on the idioms, both the canonical form and variants, if speakers are more flexible with language. It is interesting to note that this finding is not restricted to variants. `Gender` also shows a significant main effect – males tend to fixate less long on the idiom than females, although we are not quite sure why. Finally, a main effect of `Trial` is also significant; participants fixate less long on the idiom the further into the experiment they get. But the degree to which each participant is affected by the order of presentation varies, as evidenced by significant by-Subject random slopes for `Trial`. By-Item random slopes for `Condition` with correlation parameters are also significant in this model. These slopes indicate that participants’ fixation durations vary depending on which idiom occurred in which condition – participants found certain idioms easier or more difficult to understand depending on the condition in which they occurred.⁴

3.2 Altered Word as Area of Interest

We also investigated the fixation duration on the Altered Word (i.e. the word in the idiom that was manipulated). The fixed effects for this model are shown in Table 2. Since there is no altered word in the literal condition, this section focuses on the four idiom variants (i.e. lexical variation, partial forms, idiom blends, and integrated concepts) and how they compare to the canonical form.

The interaction between `Condition` and `PortionAltered` is seen in the third panel of Figure 1. The overall pattern is that longer fixation durations occur at the end of the idiom, which is also true for the canonical form. Since the idiom occurs at the end of a sentence, these longer fixations on the canonical form and variants may reflect a sentence wrap-up effect (Rayner et al., 2000; Hirotani et al., 2006). Nevertheless, the altered word for most variants shows significantly longer fixations than the canonical form. This is not true of lexical variation, which is the only variant that does not have significantly longer fixations

than the canonical form ($t = 1.54$). In other words, a lexically altered variant is just as easy to process as the canonical form. Partial forms however appear considerably different from the canonical form. Longer fixations are observed on the altered word when the beginning has been altered, such as *use the grapevine*. But when the ending is altered (e.g. *spilled it*), fixations on the altered word are not significantly different from the canonical form ($t = -1.44$). Since altering the verb does not always result in significantly longer fixations (cf. the non-significantly different lexical variant when the beginning is altered), this finding suggests that altering the verb to a semantically vague verb, in order to make the sentence grammatical, significantly inhibits processing.

The second interaction, shown in the last panel of Figure 1, is between knowledge of the idiom (i.e. `KnowIdiom`) and the participant’s flexibility with language (i.e. `PC2.LQ`). Flexibility with language only appears to be facilitative for those who do not know the idiom, illustrated by the non-significant slope for those who know the expression ($t = -1.29$). Other strategies are apparently relied upon to interpret the alternation when knowledge of the expression is not available.

Additional main effects are also observed on the altered word. Fixation durations are longer on the altered word when the co-occurrence frequencies of the idiom are higher. Thus, altering part of a more frequent sequence causes greater processing costs. In addition, participants have shorter fixation durations when the variant is rated as more acceptable (i.e. `meanVariationRating`). The more the variation strategy is preferred for a particular idiom, the easier it is to interpret. Finally, the further the participants get into the experiment (i.e. `Trial`), the shorter their fixation durations on the altered words.

	Estimate	Std. Error	t-value
Intercept	5.70	0.06	98.48
Condition=Concept	0.47	0.06	8.28
Condition=Blend	0.15	0.06	2.67
Condition=Lexical	0.09	0.06	1.54
Condition=Partial	0.30	0.07	4.61
PortionAltered=Ending	0.27	0.06	4.49
KnowIdiom=Yes	-0.04	0.03	-1.29
PC2.LQ	-0.10	0.03	-3.12
PC1.logFrequency	0.03	0.01	4.70
meanVariationRating	-0.07	0.02	-4.27
Trial	-0.04	0.01	-2.79
I(PortionAltered=Ending Condition=Concept)	-0.12	0.08	-1.46
I(PortionAltered=Ending Condition=Blend)	-0.09	0.08	-1.17
I(PortionAltered=Ending Condition=Lexical)	-0.02	0.08	-0.26
I(PortionAltered=Ending Condition=Partial)	-0.40	0.09	-4.42
I(PC2.LQ KnowIdiom=Yes)	0.06	0.02	2.27

Table 2: Fixed Effects for the Altered Word as AOI

⁴Both models have the same random effects structure.

3.3 Spillover from the Altered Word

As the idiom occurred in sentence-final position, spillover effects from an altered noun (i.e. the end of the idiom) are not able to be determined. However, for variants in which the beginning portion of the idiom was altered (the verb), it may appear to the participant reading the text as though the ending was manipulated (e.g. as if the ‘blending idiom’ was the intended idiom in *call the strings*, or part of an idiom was inserted into an otherwise non-idiomatic text, such as *use the grapevine*); therefore, we examined the fixation duration on the first content word after the verb when the verb was manipulated (i.e. the alternation occurred at the beginning of the idiom).

Only main effects are observed in the model, shown in Table 3. Spillover effects are observed for all variant types (i.e. *Condition*), but the longest durations are for integrated concepts and partial forms. Incorporating an additional word into an idiom results in a processing cost likely due to the surprisal of this extra word. Integrating this additional information into the idiom and idiomatic context requires extra time. The largest spillover effect is with partial forms. It appears that the semantically vague words used in these sentences (to make them grammatical) make these partial forms more difficult to comprehend and cause considerable spillover effects. It remains to be determined whether partial forms from more naturalistic language produce this same effect.

The last two effects are *PC1.Frequency* and *KnowIdiom*. The higher the co-occurrences frequencies of the idiom, the longer the fixation duration on the first content word after the alternation. Modifying a frequent multi-word sequence inhibits processing. However, these spillover effects are reduced if the idiom is familiar (i.e. *KnowIdiom*).

	Estimate	Std. Error	t-value
Intercept	5.95	0.08	73.41
Condition=Concept	0.27	0.07	3.76
Condition=Blend	0.17	0.06	2.75
Condition=Lexical	0.14	0.05	2.92
Condition=Partial	0.30	0.06	4.62
PC1.logFrequency	0.04	0.01	3.54
KnowIdiom=Yes	-0.11	0.05	-2.32

Table 3: Fixed Effects for the First Content Word After the Verb

4 Discussion

This study further confirms that idioms are not nearly as fixed or frozen as previously assumed,

but can actually be modified in a variety of ways while still retaining their idiomatic meaning. Furthermore, this modification does not always end in a processing disadvantage, answering our first research question. Some variants, in fact, do not show any ‘variant’ processing costs. Lexical variation, formal idiom blends, and a literal meaning of the idiom are not processed significantly longer than the canonical form. Longer fixations are observed on the altered word (at least for idiom blends) and some spillover effects are observed if the verb was altered, but this does not result in longer processing times for the idiom as a whole. These results are partly in line with our predictions. Only formal idiom blends were predicted to be processed slower than the canonical form, due to the potential surprisal at this ‘erroneous’ form. But that is not what is observed. Intentional or not, altering a word within an idiom to a synonymous or non-synonymous word does not result in a processing cost.

Some variants, on the other hand, are processed differently from the canonical form. The variant showing the greatest difference from the canonical form is the partial form of the idiom (e.g. *use the grapevine*). This idiom variant is fixated on less than the canonical form, as predicted, largely due to the omission of a word (or words) from the expression. Yet despite this overall shorter fixation, participants fixated significantly longer on the ‘replacement’ verbs (i.e. the semantically vague verbs used to connect the idiom to the sentence) and significant spillover effects were observed on the first content word after these verbs. A similar inhibitory effect was not observed if the ending of the expression was modified (e.g. *spilled it*). These results are likely due to the design of the experiment. Using tightly controlled stimuli made these partial forms unnatural and difficult to interpret. A study investigating partial forms in naturally occurring language may shed more light on the degree of difficulty for processing this variant.

Idioms with additional integrated concepts are also processed significantly different from the canonical form, but this longer fixation time observed on the whole idiom is largely attributable to the extra word in the expression. This extra word makes the reading time longer, as expected. This longer duration on the whole idiom is very similar to the Altered Word AOI, suggesting that

this variant experiences very little processing costs over and above having to read an extra word.

Meanwhile, not all variant types are processed differently from each other, answering our second research question. Lexical variation and formal idiom blends are actually processed quite similarly, showing comparable durations to each other in addition to the canonical form. These variants maintain the same length as the original expression, and perhaps better maintain the idiom's original metaphorical meaning, leading to comparable fixation durations between these two variants. However, other variants are processed quite differently. Adding an additional element (integrated concept) or omitting part of the expression (partial form) results in processing differences – requiring longer or shorter reading times, respectively.

These findings of course do not imply that all idioms can be altered using all variation strategies. Variability with the different strategies is also evident in the results. The random effects structure in both models had significant by-Item random slopes with correlation parameters for *Condition*. This indicates that specific idioms can be easier or more difficult to process depending with which condition (i.e. variation strategy) they occurred. Furthermore, idiom variants which are preferred (i.e. rated as more acceptable) show shorter reading times, or are easier to process. These results reveal that the way in which each idiom is modified can greatly affect how easy it is to understand.

This study also incorporated additional, and sometimes novel, predictor variables to shed new light on idioms and idiomatic variation. An objective measure of compositionality (i.e. LSA scores), was used in this study, and interestingly, these scores are only predictive for the idiom as a whole, and not at the word level, suggestive of the analytical nature of idiom interpretation and not necessarily reflective of a bottom-up (i.e. decomposable) process (Gibbs et al., 1989). Meanwhile, length is surprisingly seldom investigated in the literature (Fanari et al., 2010), yet appears to play a role in idiomatic comprehension. The same can be said for speaker-specific variables. Every speaker's independent knowledge of each idiom (not just an average measure of familiarity), as well as their general flexibility with nonstandard or erroneous usage, proves facilitative in understanding idioms and idiom variants.

In sum, this study found that some variant types are processed similarly to (i.e. not significantly different from) the canonical form. Not all alternations to the canonical form resulted in a processing disadvantage. These findings suggest then that idioms are not processed differently from literal language, as some scholars have claimed (Swinney and Cutler, 1979; Sprenger et al., 2006). Proposing that idioms are stored as 'large words' and understood, say upon activation of an 'idiom key' (Cacciari and Tabossi, 1988), runs into difficulty when the idiom is varied but does not take any longer to process than the canonical form (or a literal meaning of this form). All variant forms would therefore also have to be stored, burdening the Mental Lexicon with a plethora of (infrequent) forms.

More recent approaches to language challenge the traditional view of the Mental Lexicon (i.e. as a list of dictionary entries) and instead suggest that words themselves do not possess meaning but rather are cues to meaning, modulated by experience and context (Elman, 2004; Ramscar and Baayen, 2013). Under this view, idioms would not need to be regarded any differently, but would simply be a sequence of words which are cues to the intended meaning. Geeraert et al. (2017) investigated this approach with idioms using the Naive Discriminative Learner (NDL), which utilizes wide learning networks to approximate error implicit learning. They found that the idiomatic meaning receives initial support upon encountering the first word, and continues to receive support for the duration of the idiom. Alternations to the idiom affect the activation of the idiomatic meaning. If a word is changed or omitted, there is an abrupt decline in activation. However, the idiomatic meaning can also be repaired after such a decline, as with integrated concepts. Those findings are in line with the results from this study. We can manipulate idioms in various ways and still understand them, and in some instances, without any processing costs.

Acknowledgments

This study was supported by an Izaak Walton Killam memorial scholarship awarded to the first author. In addition, we would like to thank Lauren Rudat for her suggestions on improving the stimuli, and to the anonymous reviewers for their suggestions on improving the paper.

References

- John Ayto, editor. 2009. *From the horse's mouth: Oxford dictionary of English idioms*. Oxford University Press, Oxford.
- David A. Balota, Melvin J. Yap, Michael J. Cortese, Keith A. Hutchison, Brett Kessler, Bjorn Loftis, James H. Neely, Douglas L. Nelson, Greg B. Simpson, and Rebecca Treiman. 2007. English lexicon project. *Behavior Research Methods*, 39(3):445–459.
- Michael Barlow. 2000. Usage, blends and grammar. In Michael Barlow and Suzanne Kemmer, editors, *Usage-based models of language*, pages 315–345. CSLI Publications, Stanford, CA.
- Douglas Bates, Martin Maechler, Ben Bolker, and Steven Walker. 2014. lme4: Linear mixed-effects models using eigen and S4.
- Samuel A. Bobrow and Susan M. Bell. 1973. On catching on to idiomatic expressions. *Memory and Cognition*, 1:343–346.
- Cristina Cacciari and Patrizia Tabossi. 1988. The comprehension of idioms. *Journal of Memory and Language*, 27:668–683.
- J. Cooper Cutting and Kathryn Bock. 1997. That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition*, 25(1):57–71.
- Patrick J. Duffley. 2013. How creativity strains conventionality in the use of idiomatic expressions. In Mike Borkent, Barbara Dancygier, and Jennifer Hinneill, editors, *Language and the creative mind*, pages 49–61. CSLI Publications, Stanford, CA.
- Jeffrey L. Elman. 2004. An alternative view of the mental lexicon. *Trends in Cognitive Sciences*, 8(7):301–306.
- Rachele Fanari, Cristina Cacciari, and Patrizia Tabossi. 2010. The role of idiom length and context in spoken idiom comprehension. *European Journal of Cognitive Psychology*, 22(3):321–334.
- David Fay. 1982. Substitutions and splices: A study of sentence blends. In Anne Cutler, editor, *Slips of the tongue and language production*, pages 163–195. Mouton de Gruyter, Amsterdam.
- Max Freyd. 1923. The graphic rating scale. *The Journal of Educational Psychology*, 14:83–102.
- Frederik Funke and Ulf-Dietrich Reips. 2012. Why semantic differentials in web-based research should be made from visual analogue scales and not from 5-point scales. *Field Methods*, 24(3):310–327.
- Kristina Geeraert, John Newman, and R. Harald Baayen. 2017. Idiom variation: Experimental data and a blueprint of a computational model. In Morten Christiansen and Inbal Arnon, editors, *More than words: The role of multiword sequences in language learning and use*, Topics in Cognitive Science. Wiley, doi:10.1111/tops.12263.
- Kristina Geeraert. 2016. *Climbing on the bandwagon of idiomatic variation: A multi-methodological approach*. Ph.D. thesis, University of Alberta.
- Raymond W. Gibbs and Nandini P. Nayak. 1989. Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology*, 21:100–138.
- Raymond W. Gibbs, Nandini P. Nayak, John L. Bolton, and Melissa E. Keppel. 1989. Speakers' assumptions about the lexical flexibility of idioms. *Memory & Cognition*, 17(1):58–68.
- Raymond W. Gibbs. 1980. Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, 8(2):149–156.
- M. H. S. Hayes and D. G. Patterson. 1921. Experimental development of the graphic rating scale. *Psychology Bulletin*, 18:98–99.
- Masako Hirotsu, Lyn Frazier, and Keith Rayner. 2006. Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language*, 54:425–443.
- Thomas K. Landauer, Peter W. Foltz, and Darrell Laham. 1998. Introduction to latent semantic analysis. *Discourse Processes*, 25:259–184.
- Andreas Langlotz. 2006. *Idiomatic creativity: A cognitive-linguistic model of idiom-representation and idiom-variation in English*. John Benjamins.
- Matthew S. McGlone, Sam Glucksberg, and Cristina Cacciari. 1994. Semantic productivity and idiom comprehension. *Discourse Processes*, 17:167–190.
- George A. Miller and Walter G. Charles. 1991. Contextual correlates of semantic similarity. *Language and Cognitive Processes*, 6(1):1–28.
- Rosamund Moon. 1998. *Fixed expressions and idioms in English*. Oxford University Press, Oxford.
- R Core Team. 2014. R: A language and environment for statistical computing.
- Michael Ramscar and R. Harald Baayen. 2013. Production, comprehension, and synthesis: A communicative perspective on language. *Frontiers in Psychology*, May 02.
- Keith Rayner, Gretchen Kambe, and Susan A. Duffy. 2000. The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology*, 53A(4):1061–1080.
- Daniela Schröder. 2013. *The syntactic flexibility of idioms: A corpus-based approach*. AVM, Munich.
- John Sinclair, editor. 2011. *Collins COBUILD idioms dictionary*. Harper Collins.

- Simone A. Sprenger, Willem J. M. Levelt, and Gerard Kempen. 2006. Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, 54:161–184.
- David A. Swinney and Anne Cutler. 1979. The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behaviour*, 18:523–534.
- Patrizia Tabossi, Rachele Fanari, and Kinou Wolf. 2008. Processing idiomatic expressions: Effects of semantic compositionality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(2):313–327.
- Debra A. Titone and Cynthia M. Connine. 1994a. Comprehension of idiomatic expressions: Effects of predictability and literalness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5):1126–1138.
- Debra A. Titone and Cynthia M. Connine. 1994b. Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literalness. *Metaphor and Symbolic Activity*, 9(4):247–270.
- Debra A. Titone and Cynthia M. Connine. 1999. On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics*, 31:1655–1674.