Ambiguity Resolution in Passivized Idioms;

Is There a Shift in the Most Likely Interpretation?

Marianna Kyriacou[1], Kathy Conklin[2], Dominic Thompson[2]

[1]Center for Multilingualism in Society across the Lifespan (MultiLing), University of Oslo, Oslo, Norway

[2]School of English, University of Nottingham, Nottingham, United Kingdom

***Correspondence:**

Corresponding Author

marianna.kyriacou@iln.uio.no

Abstract

Ambiguous but canonical idioms (*kick the bucket*) are processed fast in both their figurative ('die') and literal ('boot the pail') senses, although processing costs associated with meaning integration may emerge in post-idiom regions. Modified versions (*the bucket was kicked*) are processed more slowly than canonical configurations when intended figuratively. We hypothesised that modifications delay idiom recognition and prioritise the literal meaning, yielding processing costs when the context warrants a figurative interpretation. To test this, we designed an eye-tracking study, where passivized idioms were followed by 'keywords' relating to their literal (bucket - *water*) or figurative (dead - *body*) meaning, or were incongruent (*time*). The remaining context was identical. The findings showed a facilitation for the literal meaning: keywords and passivized idioms in the literal condition were read significantly faster in go-past and total reading time respectively, compared to both the figurative and control conditions. However, both literal and figurative keywords were processed equally fast (and significantly faster than controls) in total reading time. In support of our hypothesis, the literal meaning of passivized idioms appears to be more highly activated and easier to integrate, although the figurative meaning receives *some* activation which facilitates its (full) retrieval if necessary.

Keywords: idioms, modifications, passive voice, eye-tracking, figurative meaning

Public significance statement

This study shows that altering the form of (otherwise familiar and) ambiguous idiomatic expressions (*e.g., the beans were spilled* as opposed to *spilled the beans*), can affect their initial interpretation, with the literal meaning becoming more dominant. Idiomatic meanings are still available but require extra processing effort. The research demonstrates that the way we understand recurrent linguistic units is influenced by how often we encounter them in a particular way.

Ambiguity Resolution in Passivized Idioms;

Is There a Shift in the Most Likely Interpretation?

Idioms are recurrent figurative expressions with a conventionalised form and meaning (*shoot the breeze* = 'to have a casual conversation') that become entrenched in memory following repeated exposure (Conklin & Carrol, 2021). While idioms such as *shoot the breeze* are unlikely to be understood literally (it is physically impossible to shoot something insubstantial like the wind), others have a perfectly plausible literal interpretation: e.g., *kick the bucket, spill the beans, break the ice, draw the line*. Such idioms involve meaning selection when embedded in a context that requires activation of a specific interpretation, thus resembling the processing of semantically ambiguous words such as homonyms (*bank* meaning 'side of a river' or 'financial institution'), and polysemes (*church* referring to the 'building' or 'religious organisation'). In polysemes and homonyms frequency determines which meaning is activated faster, as can context (Duffy et al., 1988, 2001). In idioms, in addition to frequency and context, familiarity (with the figurative meaning), transparency (how easy it is to guess the meaning of an idiom based on its component words), and literalness (how likely it is for an idiom to be used literally) may affect when and how fast the literal versus the figurative meaning is accessed.

These factors have been found to also modulate the processing and acceptability ratings of modified idiom forms such as *the bucket was kicked* (Geeraert et al., 2017; Haeuser et al., 2020; Kyriacou et al., 2020, 2021; Mancuso et al., 2020; Mancuso & Laudanna, 2019; McGlone et al., 1994), but we currently do not know how the figurative versus the literal meaning of a modified idiom is activated, or integrated as part of longer discourse. In the current study, we explore ambiguity resolution in modified idioms during on-line reading. Specifically, an eye-tracking experiment examines which meaning of a literally plausible

idiom receives greater activation when the phrase appears in the passive voice, and is followed by information disambiguating the intended meaning (e.g., *When **the bucket was kicked**, she got rid of the **body/water** with the help of her devoted housekeeper*).

Meaning Activation in Canonical Idioms

Most of what we know about meaning activation in idioms comes from research on canonical idiom processing. Early theories assumed that idioms are not compositional phrases but instead (large) lexical chunks, and argued for either a literal-first computation, with the figurative meaning being accessed only when the literal was deemed inappropriate in the context (Bobrow & Bell, 1973); a figurative-first computation with the literal meaning being completely by-passed (Gibbs, 1980); or a parallel computation of both meanings from the phrase onset, with the figurative meaning being faster to access due to direct retrieval from the lexicon (lexical route) as opposed to word-by-word computation (Swinney & Cutler, 1979). This helped to explain why idioms (*kick the bucket*) produced faster responses than literal control phrases (*lift the bucket*) in phrasal decision tasks (Swinney & Cutler, 1979), but it also entailed that fast idiom processing could only lead to activation of the figurative meaning; the literal meaning (i.e., *kick the bucket* meaning 'boot the pail') would not have a stored representation in the lexicon, and therefore, its activation should produce similar response times as literal control phrases (*lift the bucket*).

However, fast activation of the figurative meaning does not necessarily preclude simultaneous (or partial) activation of the literal. For instance, *kick the bucket* primes both DIE (figurative associate) and PAIL (literal associate) equally well in both production and comprehension studies (Cutting & Bock, 1997; Holsinger, 2013; Holsinger & Kaiser, 2013; Peterson et al., 2001; Snider & Arnon, 2012; Sprenger et al., 2006). It has been argued that meaning activation spreads from the literal component words (kick, the, bucket) to specific

lexico-syntactic representations or superlemmas (*kick-the-bucket*), and from there to the concept ('die')—with the process being reversed in production (*Superlemma Theory*; Sprenger et al., 2006). A superlemma level effectively acts as an intermediary step, or in other words an additional layer, linking the (literal) idiom component words to the figurative meaning. Furthermore, meaning activation is neither parallel (in the sense that both meanings are computed directly from the phrase onset), nor is it purely lexical or compositional. Figurative meanings can be retrieved before (Cacciari & Corradini, 2015), or after literal ones (Swinney, 1981) depending on how soon an idiom is recognised. Prior to recognition, idioms are most likely processed compositionally, while after recognition they are most likely processed lexically, thus benefitting from a hybrid representation.

The point of recognition acts as the 'key' to the figurative meaning (*Configuration Hypothesis*; Cacciari & Tabossi, 1988), and the earlier the key occurs in a phrase, the faster the figurative meaning becomes available. In predictable idioms (i.e., idioms that are recognised before the phrase offset: *take the bull…*[*by the horns*]), recognition occurs before encountering the whole phrase. Greater idiom frequency and familiarity, as well as increased phrase length (Fanari et al., 2010), can equally boost initial idiom recognition, speeding up the retrieval of the figurative meaning (*Multidetermined Model*; Libben & Titone, 2008; Titone & Libben, 2014). The role of transparency is less straightforward as many studies have failed to find an effect (Carrol & Conklin, 2014; Libben & Titone, 2008; Tabossi et al., 2008; Van de Voort & Vonk, 1995), while others that did, have produced contradicting findings. It remains unclear whether the figurative meaning is activated more quickly in transparent idioms, where a connection between the literal and figurative meaning is more easily established (Caillies & Butcher, 2007; Gibbs et al., 1989; Gibbs & Nayak, 1989), or conversely, in opaque idioms, where the phrases are more lexical-like and hence should benefit more from the lexical route (Carrol & Conklin, 2020; Titone & Libben, 2014). Literal

plausibility (here used interchangeably with literalness) has a facilitative effect on the processing of idioms (Cronk & Schweigert, 1992; Mueller & Gibbs, 1987), likely because when an idiom can also be used literally, the overall frequency of the phrase increases. Although a highly plausible literal meaning could in theory undermine activation of the figurative, due to increased competition between the alternative senses, evidence suggests that highly literal idioms benefit both their interpretations (at least when post-idiom context is not at odds with preceding information), while idioms with low literalness induce a cost when intended in their literal sense (Beck & Weber, 2020; Mancuso et al., 2020; Titone & Libben, 2014).

Crucially, while such factors may influence how fast the figurative meaning will be activated relative to the literal, with familiarity and predictability particularly accelerating the retrieval of the figurative meaning, research suggests that both meanings of familiar and/or predictable idioms are activated even in the absence of context, provided that the phrases are encountered in full, and are 'known' to the (native) comprehender. For example, *pull one's leg* primes both WALK (literal associate) and JOKE (figurative associate) when these target words are presented a few milliseconds after the phrase offset (Beck & Weber, 2016). When embedded in a literally- or figuratively-biasing context, activation of the respective meaning is further facilitated, especially when context precedes idioms, as this can yield strong contextual predictions leading to earlier idiom recognition or easier processing of the literal meaning (Beck & Weber, 2020; Canal et al., 2017; Colombo, 2014; Fanari et al., 2010; Holsinger & Kaiser, 2013; Ortony et al., 1978; Titone & Connine, 1999). Furthermore, relevant findings show that familiar/predictable idioms are processed equally fast in *both* senses (e.g., *at the end of the day* meaning 'in the evening' or 'eventually') when they are preceded by biasing contexts (Conklin & Schmitt, 2008; Siyanova-Chanturia et al., 2011), and even if they have a dominant figurative interpretation, or are semantically transparent

(Milburn & Warren, 2019). These findings suggest that ambiguous canonical idioms (that are nevertheless familiar/predictable) activate both meanings, and no processing differences are observed in the idiom region when one meaning is preferred over the other.

Canonical Idioms and Lexical Ambiguity

Evidence from the time course of meaning activation and selection in ambiguous words has demonstrated that the presence of a biasing context (as well as the relative strength of that context) interacts with meaning frequency and influences initial lexical access (e.g., Duffy, Kambe, & Rayner, 2001; David S Gorfein, 1989; Hogaboam & Perfetti, 1975; Martin, Vu, Kellas, & Metcalf, 1999; Onifer & Swinney, 1981; Simpson, 1981, 1984; Vu, Kellas, & Paul, 1998). Balanced homonyms have been found to slow reading in a neutral context, probably reflecting competition between the equally viable (but semantically distinct) interpretations, while unbalanced homonyms tend to favour their dominant (most frequent) meaning (Duffy et al., 1988). Activation of subordinate (less frequent) meanings in unbalanced homonyms comes at a processing cost (subordinate bias effect) even if preceding context biases its activation (Rayner et al., 1994). Thus, the dominant meaning of a biased homonym is more readily committed to, and is difficult, if not impossible, to fully suppress.

The findings reviewed in the previous section suggest that fast(er) idiom recognition, and hence fast(er) activation of the figurative meaning, does not hamper activation of the literal. That is, at the idiom (ambiguity) level, there seems to be no costly meaning competition in the form of a subordinate bias effect; unlike unbalanced homonyms, familiar/predictable idioms are *not* read more slowly in their less dominant literal sense. Furthermore, findings suggest that idioms in both their figurative and literal senses are processed faster than matched control phrases; for example, *at the end of the day* is read faster than *at the end of the war* (Siyanova-Chanturia et al., 2011), and a number of studies

corroborate this 'idiom superiority effect' over matched control phrases (e.g., Carrol, 2021; Carrol et al., 2016; Carrol & Conklin, 2014, 2017, 2020; Haeuser et al., 2020; Laurent et al., 2006; Rommers et al., 2013; Strandburg et al., 1993; Swinney & Cutler, 1979; Tabossi et al., 2009; Titone et al., 2019; Underwood et al., 2004; Vespignani et al., 2010). This further suggests, that unlike balanced homonyms, ambiguous idioms do not slow reading. In fact, this idiom processing advantage mirrors findings from studies on regular polysemes (i.e., ambiguous words whose senses are derived by rule such as *chicken* meaning 'the living animal' or 'cooked meat'). In particular, findings have demonstrated that regular polysemes are faster to process than matched, unambiguous words (Frazier & Rayner, 1990; Klepousniotou et al., 2012; Klepousniotou & Baum, 2007; Pickering & Frisson, 2001), presumably because the different senses of the word share a core underspecified meaning, that results in lack of competition and facilitates processing. This argument is not uncontested (see Foraker & Murphy, 2012; Klein & Murphy, 2002), and similar concerns have been raised about the robustness of the idiom superiority effect. Specifically, Kyriacou et al. (2021) argued that an idiom advantage might be discernible only when idioms are compared to poorly matched literal phrases, and that studies focusing on the idiom region alone might have potentially missed processing costs emerging in other regions.

Indeed, in line with the above argument, some evidence implying costly competition between the different meanings of a (canonical and familiar) idiom, comes from studies investigating how a respective meaning is integrated with post-idiom context. For example, a disambiguating context that favours a figurative interpretation, is read faster than a literal context when preceded by idioms that are highly familiar/predictable, as well as literally implausible (Beck & Weber, 2020; Cacciari & Corradini, 2015; Titone et al., 2019). Conversely, literal and more decomposable idioms may be more easily integrated with a literal context (Beck & Weber, 2020; Titone et al., 2019). Interestingly, analogous effects

have been noted for balanced irregular polysemes, whose senses only bear *some* semantic similarity to one another (e.g., *wire* in the sense of ‘electrical cable’ and ‘spying tool’). In this case too, ambiguity seems to be resolved in subsequent regions rather than in the ambiguity region itself (Brocher et al., 2018).

Taken together, despite the open questions regarding the idiom processing advantage over matched unambiguous phrases, the findings would suggest that canonical idioms are more like balanced (irregular) polysemes as they are unlikely to lead to dominance effects in the ambiguity (idiom) region, but processing costs associated with meaning selection and integration may become visible in subsequent regions. Specifically, slower processing may be observed in post-idiom, disambiguating regions, when the literal meaning of a highly familiar/predictable idiom is intended, or for the figurative meaning of a less familiar/predictable and potentially more transparent idiom respectively. Therefore, although both idiom meanings can be activated in the absence of context, idiom-specific variables may increase (or reduce) the activation of a given meaning, thus facilitating (or hindering) its integration with upcoming contextual information.

Meaning Activation in Modified Idioms

Contrary to previous assumptions claiming that (at least some) idioms are syntactically inflexible, and therefore unable to undergo complex syntactic modifications without losing their figurative meaning (e.g., Fraser, 1970; Gibbs et al., 1989; Nunberg, 1978), recent psycholinguistic evidence has demonstrated that this is not the case. For example, passivized idioms (*the bucket was kicked*) embedded in a context biasing the activation of their figurative concept (i.e., ‘death’), were processed significantly faster than anomalous control phrases (*the apple was kicked*; Kyriacou et al., 2020). This suggests that the figurative meaning was available, since a strictly literal understanding of the phrases

would have been just as contextually anomalous as the controls. Passivized idioms have also been found to prime target words related to their figurative meaning in phrasal decision tasks (*the bucket was kicked*–DIE; Mancuso et al., 2020), and idioms modified by extra adjectives (*spill the spicy beans* in a scenario involving the revelation of scandalous secrets) are understood figuratively (Kyriacou et al., 2021). In addition, although greater frequency and familiarity can aid the processing of modified idiom forms, just as in canonical versions, the figurative meaning appears to be available even for those that are less familiar and less frequent (Kyriacou et al., 2020). Combined, these studies indicate that the idiomatic meaning is accessible despite phrasal modifications, but they were not designed to measure the activation level of the literal interpretation.

The aforementioned studies report a processing cost for modified idiom forms relative to the canonical form. For instance, passivized idioms (*the beans were spilled*) induced significantly longer reading times than their canonical configurations (*spilled the beans*) (Kyriacou et al., 2020), and response times to figurative targets (SECRET) were significantly slower, at least when an overt semantic association judgement was required (Mancuso et al., 2020). It is possible that this processing delay is due to the modification itself. It is well established, for instance, that the passive voice requires more processing effort than the active (Kamide et al., 2003; Knoeferle, 2007; Knoeferle et al., 2005; Mack et al., 2013; Meyer et al., 2012). Importantly, however, this processing cost might be indicative of a shift in meaning balance. Akin to what has been noted for canonical idioms of lower familiarity/predictability, the figurative meaning of modified idioms might be *slower* to activate than the literal, making its integration with a figurative context laborious. In this case, unfamiliarity (as well as unpredictability and lower frequency) would stem from the novelty of the phrasal configuration, which could delay idiom recognition even if the canonical form is highly familiar to a speaker. Assuming that the figurative meaning of a canonical idiom is ‘known’,

it should become available and be retrievable given some extra processing time (i.e., some time after the phrase offset). Conversely, integrating the literal meaning (with a literal context) might be easier and hence faster, as the literal meaning would be fully available and active at phrase offset. In previous research, passivized idioms were always preceded by a figuratively biasing context, making it impossible to draw conclusions as to how meaning ambiguity would be resolved if the same phrases were intended literally. It is also unclear whether the figurative meaning of a modified idiom would receive *some* activation by default out of context, or if it would only be considered when a context necessitates its activation. What is more, idiom processing models like the *Configuration Hypothesis* (Cacciari & Tabossi, 1988), the *Multidetermined Model* (Libben & Titone, 2008; Titone & Libben, 2014), and the *Superlemma Theory* (Sprenger et al., 2006) do not specify if, or how, phrasal modifications affect idiom processing, nor how they influence the interplay between figurative and literal meaning activation.

The Present Study

The current eye-tracking study explores meaning activation and integration in passivized idioms (*the bucket was kicked*). It investigates whether the processing cost observed for modified idioms (Kyriacou et al., 2020; Mancuso et al., 2020) is solely due to the modification itself (i.e., active vs. passive) or can also be attributed to the increased availability of the literal meaning. To avoid initial contextual bias towards a particular idiom interpretation, the phrases were preceded by a single neutral word (e.g., <u>*When*</u> ***the bucket was kicked*** …). In line with evidence from both irregular polysemes and idioms, suggesting that ambiguity may be resolved outside the ambiguous region (Brocher et al., 2018; Cacciari & Corradini, 2015; Titone et al., 2019), we kept the remaining context constant to enable comparisons between conditions, with the exception of a 'keyword' that disambiguated the intended meaning of the phrase (… *she got rid of the* ***body/water*** *with the help of her devoted*

housekeeper). Figurative and literal keywords were matched for word length and frequency. A control condition with contextually anomalous keywords (***time***) was also introduced as a baseline, against which the activation of the figurative meaning could be measured. If the figurative meaning failed to be activated, then the figurative keyword (*body*) should be equally anomalous as the control keyword (*time*) and there should not be a processing difference between them.

We hypothesised that if the literal meaning of passivized idioms is activated first and/or receives greater activation due to the 'novel' presentation of the phrases, then the literal keywords (*water*) should be significantly faster to read than both the figurative (*body*) and control (*time*) keywords. The figurative keywords should be read faster than controls, if the figurative meaning receives some activation. Alternatively, if both idiom meanings receive (almost) equal and/or simultaneous activation, as has been demonstrated for canonical idioms (Milburn & Warren, 2019; Siyanova-Chanturia et al., 2011), then the figurative and literal keywords (*body/water*) should be equally fast to read, and both should be significantly faster than controls (*time*).

In addition to the keywords, we examined the post-keyword regions for 'spillover' effects; if the figurative meaning is more difficult to integrate with later parts of the sentence due to lesser/weaker activation, post-keyword regions (*with the help of her devoted housekeeper*) might elicit longer total reading times in the figurative condition. We also looked at the idiom region (*the bucket was kicked*), expecting to find differences in total reading time. Keywords in the figurative condition might cause longer regressive refixations back to the idiom, thus increasing the overall (re)reading time of the phrases. Therefore, we had three regions of interest (ROIs), (a) the ambiguous/idiom region (*the bucket was kicked*), (b) the disambiguating/keyword region (*water/body/time*), and (c) the post-disambiguation region (*with the help of her devoted housekeeper*).

Methodology

Participants

Forty-five native speakers of British English took part in the eye-tracking experiment (mean age = 20.02, SD = 3.28; 37 female). Participants received compensation or course credit for their participation. This study was approved by the ethics committee of the Faculty of Arts at the University of Nottingham.

Materials

Forty-five idioms were selected from the Collins Cobuilt Idioms Dictionary (Sinclair, 2011), many of which were also used in Kyriacou et al., (2020). All idioms had a plausible figurative and literal interpretation (e.g., *kick the bucket, spill the beans*) and the dominance of each meaning was assessed through the literalness questionnaire discussed in later paragraphs. All idioms consisted of a verb plus a noun phrase (V + NP) to enable us to use them in the passive voice. For a full list of items see Appendix 1. One idiom (*rock the boat*) was removed from analyses due to an error in the design of the experiment, thus leaving forty-four idioms in total. Idioms were always presented in the passive (*the bucket was kicked, the beans were spilled*).

The passivized idioms were allocated three keywords: (a) one that matched the figurative meaning of the phrase (Figurative Condition), (b) one that matched the literal meaning of the phrase (Literal Condition), and (c) one that was unrelated and contextually anomalous (Control Condition). Examples of the stimuli can be seen in Table 1. As the purpose of the keywords was to disambiguate the intended meaning of the phrase, they differed across conditions. The remaining context was identical in all three conditions. Care

was taken so that a keyword[1] never occurred at a line break, avoiding fixation contamination due to the programming and execution of saccades (Conklin et al., 2018), and were of the same grammatical class (e.g., noun, verb) for each idiom in question. The sentences were of a comparable length and keywords were placed approximately in the middle or towards the end of each sentence.

Table 1

Example Stimuli Sentences Across Conditions: (a) Figurative, (b) Literal, and (c) Control.

1.	Kick the bucket	a	When the bucket was kicked, she got rid of the body with the help of her devoted housekeeper.
		b	When the bucket was kicked, she got rid of the water with the help of her devoted housekeeper.
		c	When the bucket was kicked, she got rid of the time with the help of her devoted housekeeper.
2.	Tie the knot	a	Once the knot was tied, he felt confident that his wife would not go anywhere.
		b	Once the knot was tied, he felt confident that his boat would not go anywhere.
		c	Once the knot was tied, he felt confident that his idea would not go anywhere.
3.	Clear the air	a	When the air was cleared, they were finally able to reconcile and everything went back to normal.
		b	When the air was cleared, they were finally able to breathe and everything went back to normal.

[1] Most of the keywords (27/44) completed the clause in which they were embedded. For instance, upon encountering *body* and *water* in (1) the clause *she got rid of the x* is completed. In contrast, *wife* and *boat* in (2) are the subjects of a secondary clause whose meaning is not completed until the end of the sentence. This could potentially influence how quickly the meaning of a phrase is integrated, since the completion of a clause by a keyword may allow disambiguation of the phrase more readily. We explored this as a factor when analysing the eye-tracking data but found no significant modulation in any measure investigated.

c When the air was cleared, they were finally able to jump and everything went back to normal.

Note. The idioms and keywords are underlined in the examples above but were not demarcated in any way for the participants.

Figurative and literal keywords were matched for frequency ($p = 0.90$) and word length ($p = 0.67$), but control keywords were deliberately chosen to be more frequent and significantly shorter (all ps < 0.05) (see Table 2). This would ensure that any adverse processing observed for the control keywords in the eye-tracking experiment would be attributable to their contextual incongruency and not to other lexical properties. To this end, we also conducted a lexical decision task using the keywords along with an equal amount of pseudowords that matched the keywords in length. Twenty native speakers of British English (age = 19.45, SD = 1.02; 19 = females) took part in the experiment and were asked to decide whether each string of letters was a real English word or not by pressing a corresponding key on the keyboard. The experiment was designed using PsychoPy (Peirce, 2007) and words were presented randomly one-by-one in the middle of the screen, in white Arial font over a black background. The whole procedure lasted for approximately 5 minutes. Response times shorter than 350ms and longer than 2000ms were removed from the analysis, leading to a data loss of 3.65%, out of which 1.55% were keywords. Keywords elicited a significantly faster response time (M = 620ms, SD = 220) than pseudowords (M = 780ms, SD = 304); $t(5085) = 21.55$, $p < 0.005$), and crucially no significant effect of Condition (figurative, literal, control) was found for the response time of keywords [$F(2, 2596) = 0.212$; $p = 0.80$] (for means see Table 2).

Table 2
Summary of Keyword Characteristics

		Frequency		Word Length		Response Time	
Keyword Type	N	Mean	SD	Mean	SD	Mean	SE
Figurative	44	65.08	62.38	6.48	2.42	616.67	7.43
Literal	44	51.39	67.24	6.11	2.05	619.73	7.29
Control	44	186.20	235.10	4.30	1.41	623.57	7.77

Note. Frequency values were extracted from the BNC and are reported above as instances per million words, Word Length was calculated as number of characters, and Response Time to keywords in the lexical decision task is reported in milliseconds.

Participants (n = 64) from the same population, who did not participate in the eye-tracking experiment, rated the idioms[2] for familiarity (with an idiom's figurative meaning; n = 16), frequency (perceived frequency of occurrence in language use; n = 16), transparency (how clear the figurative meaning is based solely on the idioms' component words; n = 16), and literalness (how likely it is to encounter a given idiom in its figurative vs. its literal sense; n =16). These norming data were collected via separate questionnaires using Likert rating scales from 1 (very unfamiliar/infrequent/transparent) to 5 (very familiar/frequent/opaque), and for literalness on a scale from 1 (most likely literally) to 7 (most likely figuratively). Of note, the familiarity rating questionnaire included dictionary definitions of the idioms to ensure that the ratings reflected familiarity with the correct meaning of the phrase (e.g., *to kick the bucket* means 'to die'). Table 3 demonstrates that the idioms were rated as highly familiar, fairly frequent, more frequently occurring in their figurative sense, and transparent. Using corpus-based frequencies from the BNC (BNC Consortium, 2007) we also calculated the transitional probability of the final words of the phrases (e.g., *kicked* in *the bucket was kicked*) as a measure of phrasal predictability (McDonald & Shillcock, 2003a, 2003b). As expected due to the infrequent phrase configuration, transitional probability was very low, ranging between 0.00 to 0.11 (M = 0.02), resembling the very low predictability scores of passivized idioms reported in Kyriacou et al. (2020). Because of this, transitional probability was not included as a predictor in the analyses.

Table 3
Summary of Idiom Characteristics

	N	Mean	SD
Familiarity (/5)	44	4.30	0.63
Transparency (/5)	44	2.69	0.58

[2] All the ratings were gathered for the canonical form of the idioms and used in the analyses as predictors affecting the processing of their modified forms and their associated keywords.

Literalness (/7)	44	5.31	0.82
Frequency (/5)	44	3.40	0.69

Note. Familiarity, frequency, and transparency were assessed in norming tasks on a scale from 1 (very unfamiliar/very infrequent/ very transparent) to 5 (very familiar/ very frequent/ very opaque) and literalness on a scale from 1 (most likely encountered literally) to 7 (most likely encountered figuratively).

As part of the norming procedures we also assessed the association strength between the keywords and the idiom content words (e.g., association of *bucket* and *kick* with the keywords *body/water/time*) using the University of Florida Free Association Norms (Nelson et al., 2004). The association strengths between idiom content words and both the figurative and control keywords were mostly 0, whereas for literal keywords there were occasionally some strong associations (e.g., *air* from *clear the air* was highly associated with the literal keyword *breathe*). An ANOVA test showed that Condition was significant for both idiomatic nouns (*bucket*) [$F(2, 1602) = 77.10$, $p < .001$] and verbs (*kicked*) [$F(2, 1602) = 29.84$, $p < .001$)] and further pairwise comparisons revealed that literal keywords had significantly stronger association strength with both idiomatic nouns and verbs compared to both figurative and control keywords ($p < .001$), while no difference was found between figurative and control keywords in any case (all ps > 0.9). To be able to include association strength as a predictor in the data analyses, we recoded association strength as original association score plus 1 and then log-transformed it to reduce the skewness of the distribution.

Procedure

Three experimental lists were created using a Latin square design so that each participant encountered a target phrase in only one of the three conditions (figurative, literal, or control). To distract participants from the purpose of the study and especially the semantic incongruency of the control keywords, 45 filler sentences having a similar form were created,

with some containing semantically anomalous words (**When the soup was ruined, everyone wanted to go out for trees instead*). The same filler sentences were used across all lists.

An SR EyeLink 1000+ desktop-mount tracker was used to track and record eye-movements during reading with a minimum sampling rate of 500Hz. To minimize head movement during reading, a chin- and forehead-rest was used. Participants were seated facing a computer monitor and a 9-point calibration was performed at the beginning of the experiment and whenever necessary thereafter. Participants were asked to read each sentence as quickly as possible but for comprehension, and to press ENTER after they finished reading a sentence to proceed to the next one. Around half of the filler sentences were followed by a yes or no comprehension question to ensure participants' attention to the content of the sentences. The presentation of the stimuli was randomized per participant and each sentence was preceded by a drift correction. Sentences were triple-spaced and presented in black font (Courier New, size 14) on a white background. The whole task lasted about 20 minutes.

Results

Participants' accuracy on the comprehension questions was high (93%), indicating that they had understood the task and the sentences. We focused the analyses on three regions of interest (ROIs): (a) the ambiguous/idiom region (*the bucket was kicked*), (b) the disambiguation/keyword region (*body/water/time*), and (c) the post-disambiguation region (*with the help of her devoted housekeeper*). Of note, the term 'idiom' is used throughout in the remaining section to refer to the ambiguous passivized phrases, regardless of whether the phrase was meant figuratively or literally. When referring to a particular sense of the phrase, this is specified.

For all ROIs we examined the total reading time (the sum of all fixation durations including those of regressive refixations). For the keyword region we also considered first

pass gaze duration (an early measure which consists of the duration of all fixations and refixations in the ROI, from when the ROI is first fixated and before the eye moves to the right) and go-past time (which includes the first pass gaze duration plus any time spent revisiting previous parts to the left of the ROI). The mean reading times for the different ROIs and measures are presented in Table 4 and detailed model outputs are reported in Appendix 2.

Table 4
Mean Reading Time Across Measures and Regions of Interest

	First Pass Gaze Duration		Go-Past Reading Time	
	Mean	SE	Mean	SE
Keyword Region				
Literal	210	4.63	236	7.14
Figurative	215	4.76	250	7.60
Control	212	4.73	255	7.80

	Total Reading Time	
	Mean	SE
Keyword Region		
Literal	240	8.49
Figurative	250	8.89
Control	279	9.95
Idiom Region		
Literal	713	33.50
Figurative	749	35.20
Control	752	35.1
Post-Disambiguation Region		
Literal	769	39.50
Figurative	798	41.00
Control	918	47.10

Note. All fixation durations above are reported in milliseconds.

The data analyses were carried out using mixed-effects models and the *lme4* package, version 1.1-28 (Bates et al., 2014) in R, version 4.1.3 (R Core Team, 2021). All fixation durations were log-transformed to reduce the positive skewness of the distribution and were analysed using linear mixed-effects models. Condition, a fixed effect, was operationalised as

a three-level factor (literal, figurative, and control), with the literal condition set as the baseline. Further pairwise comparisons between the three conditions (e.g., figurative vs. control) were carried out using the *emmeans* package (Lenth, 2018). By-item and by-participant intercepts and slopes were included as random effects initially (Barr et al., 2013; Bates et al., 2015), but due to model convergence issues this was simplified to by-item and by-participant intercepts only for the phrase region, and a by-participant intercept only for the keyword and post-disambiguation regions. Familiarity, frequency, transparency, and literalness were added stepwise as predictors in the models of idiom and keyword region and were modelled as interactions with Condition, to explore whether they have a bearing on the processing of the idiomatic phrases and/or on the degree of activation of the figurative versus the literal meaning. Association strength was also included as predictors in keyword analyses.[3]

To reduce collinearity (k = 35) between familiarity and transparency (r = -0.50), familiarity and frequency (r = 0.60), literalness and transparency (r = 0.47), and between frequency and transparency (r = -0.67), after centering these variables, familiarity was residualised over both transparency and frequency, frequency over transparency, and literalness over both transparency and familiarity (k = 1). The new residualised factors created for familiarity, frequency, and literalness all highly correlated with their corresponding original variables (all rs ≥ 0.74), but not with one another (all rs ≤ 0.01).

Idiom Region Analysis

In the idiom region, outlier removal processes resulted in very little data loss (0.35%). For total reading time, Condition was significant; passivized idioms used literally were read

[3] For all models we also included Trial Sequence Number as a predictor, to make sure that main effects were not driven by increased exposure and practice with the stimuli. Trial Sequence Number was indeed significant in many models, but as the main effects were unaffected by this, we do not report further details in the Results or Discussion.

significantly faster than those used figuratively (β_{FIG} = 0.05, SE = 0.02, *t = 2.34*; *p* = 0.02) and the control condition (β_{CON} = 0.05, SE = 0.02, *t* = 2.52; *p* =0.01). Further pairwise comparisons showed that the reading time of the idioms in the figurative condition did not differ significantly from the control condition (*p* = 0.98). There was a significant effect of frequency (β = -0.10, SE = 0.03, *t* = -3.41; *p* = 0.001), whereby the more frequent the canonical idiom (based on ratings), the faster its passivized form was read, but there was no significant interaction between frequency and Condition ($p_s \geq 0.22$)

Keyword Region Analyses

In the keyword region, outlier removal processes again resulted in little data loss (0.33%). However, 18.68% of keywords were skipped during first pass reading and were not later fixated. Importantly, skipping affected all conditions equally (115 = literal, 122 = figurative, and 133 = control). The skipping rate may be due to the short length of the keywords (M = 5.63 characters), which may have led to them being processed while in parafoveal view (Conklin et al., 2018; Hyönä, 2011; Rayner, 2009), thus eliminating the need for a direct fixation.

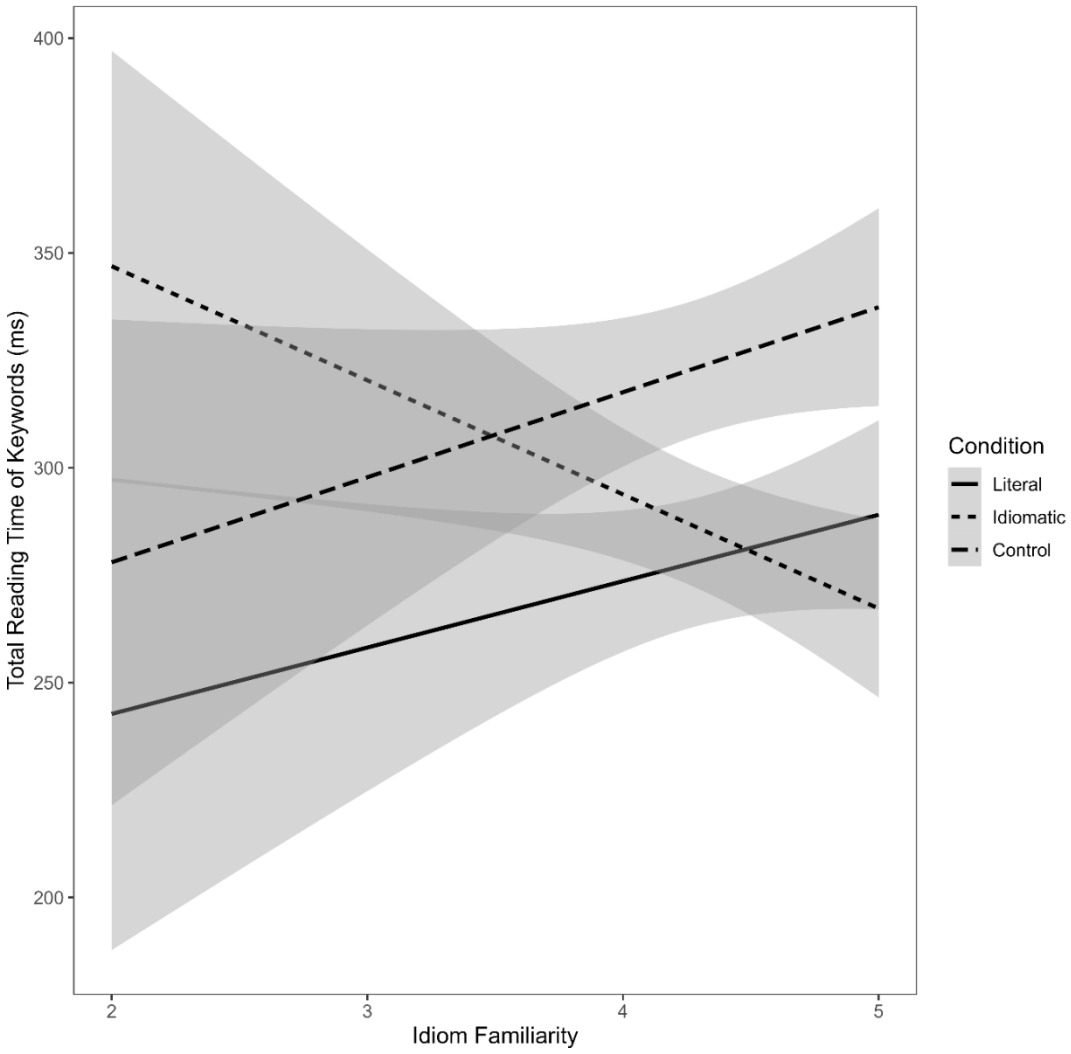
First Pass Gaze Duration. There was no effect of Condition on the first pass gaze duration of the keywords (β_{FIG} = 0.02, SE = 0.02, *t = 1.03*; *p* = 0.30 and β_{CON} = 0.01, SE = 0.02, *t* = 0.52; *p* =0.60), or of any other factor.

Go-Past Reading Time. Condition was significant in the go-past reading, whereby the figurative keywords (β_{FIG} = 0.06, SE = 0.03, *t = 2.01*; *p* = 0.04) and control keywords (β_{CON} = 0.08, SE = 0.03, *t* = 2.69; *p* = 0.007) were both read significantly more slowly than the literal keywords. Further pairwise comparisons showed that the reading time of the figurative and control keywords did not differ significantly (*p* = 0.77). Greater frequency

significantly accelerated the go-past reading time of keywords (β = -0.07, SE = 0.03, t = -2.34; p =0.01), but did not significantly interact with Condition ($p_s \geq 0.36$)

Total Reading Time. Condition was also significant in the total reading time. In this case, however, the literal keywords were *not* significantly faster to read than the figurative keywords (β_{FIG} = 0.04, SE = 0.03, *t = 1.54*; p = 0.12), but they were still significantly faster than the control keywords (β_{CON} = 0.15, SE = 0.03, t = 5.42; $p < 0.001$). Further pairwise comparisons showed that the figurative keywords were *also* read significantly faster than controls (p = 0.004). Greater frequency made total reading time of keywords significantly faster (β = -0.07, SE = 0.03, t = -2.54; p =0.01), but did not significantly interact with Condition ($p_s \geq 0.49$). Interestingly, familiarity did; the significant interaction with Condition revealed that greater familiarity (with the figurative meaning of an idiom) led to slower reading time of literal keywords (β = 0.08, SE = 0.03, t = -3.00; p =0.003), but to faster reading time of figurative keywords (β = -0.14, SE = 0.04, t = -3.73; $p < 0.001$). We ran the same model using the Figurative condition as the baseline of Condition to further examine the interaction between familiarity and Condition. The relevelled model indicated that control keywords were also read significantly more slowly than figurative keywords as idiom familiarity increased (β = 0.09, SE = 0.04, t = 2.56; p =0.01; see Figure 1).

Figure 1
The Effects of Idiom Familiarity Ratings on the Total Reading Time of Keywords Respectively



Post-Disambiguation Region Analysis

No outliers were removed from the post-disambiguation dataset, but 0.64% of data was lost due to track loss and/or premature pressing of the ENTER key by the participants. The analysis of the total reading time revealed that this region was read significantly more slowly in the control than in the literal condition (β_{CON} = 0.18, SE = 0.03, t = 5,91; p < 0.001), but there was no difference between the literal and figurative condition (β_{FIG} = 0.04,

SE = 0.03, t = 1.23; p = 0.21). This region was read significantly more quickly in the figurative than in the control condition ($p < 0.001$).

Discussion

This study aimed to investigate the activation of the literal and figurative meanings of idioms encountered in the passive voice. We hypothesised that the deviation from the established phrasal configuration would slow idiom recognition, delaying the retrieval of its figurative meaning while prioritizing the literal. This could explain why modified, and specifically passivized idioms elicit a processing cost when intended in their figurative sense relative to their canonical forms (Kyriacou et al., 2020, 2021; Mancuso et al., 2020). To investigate whether differences in activation of the literal and figurative meanings might account for this observed pattern, and to explore whether passivized idioms are recognised in the absence of preceding contextual bias, we designed an eye-tracking study using passivized idioms embedded in sentences, where the intended meaning became clear in a post-idiom disambiguating (keyword) region. We observed eye-movements to three ROIs: the idiom region (*the bucket was kicked*), the keyword region (*body/water/time*), and the post-disambiguation region (*with the help of her devoted housekeeper*). The different ROIs enabled us to pinpoint where and when ambiguity arises and is resolved. Overall, the findings suggest that although both meanings were activated, the degree of their activation differed significantly, making the literal meaning easier to integrate. The findings are discussed in more detail below.

We start by considering the findings of the keyword analyses, as this region was the locus of ambiguity resolution. In first pass gaze duration, we did not find any differences in the reading of either the literal, figurative, or control keywords, which suggests that this measure reflects initial word recognition operations. In go-past reading time, however, the

literal keywords (*water*) were read significantly faster relative to both the figurative (*body*) and control (*time*) keywords. There was no difference between the figurative and control keywords. This indicates that initially the literal meaning of the passivized idioms had greater activation than the figurative, making the literal keyword easier and faster to process. Notably, however, in total reading time (a late measure), the difference between literal and figurative keywords disappeared, and both were read significantly faster than controls, implying that although the literal meaning was activated more quickly and/or to a greater degree, the figurative meaning was accessible and relatively easy to retrieve.

The findings from total reading time, therefore, suggest that *some* activation of the figurative meaning had occurred *prior* to encountering the disambiguating region (i.e., before a figurative interpretation was triggered by the keyword). Recall that keywords appeared a few words after the idioms and, hence, there was time to potentially activate both interpretations (Beck & Weber, 2016). Had the figurative meaning not been activated at all, we would have expected the figurative keywords to be slower than the literal ones in total reading time (although faster than controls). We argue, therefore, that partial activation of the figurative meaning better accounts for the present data, and lines up with previous findings demonstrating that passivized idioms prime figurative associates in neutral contexts (Mancuso et al., 2020). Further evidence in support of this, comes from the analysis of the post-disambiguation region (*with the help of her devoted housekeeper*), where no spillover effect was found as a function of idiom sense, while in the control condition reading time was significantly slower. The figurative interpretation was not treated as anomalous, and any delays associated with its activation and/or integration were quickly and fully resolved in the keyword region. It appears, therefore, that the figurative meaning of passivized idioms is activated, albeit to a lesser degree, even without preceding contextual bias. When subsequent

information is felicitous with a figurative interpretation, some extra processing time is required to fully activate it and integrate it in the context.

More effortful integration of the figurative meaning was also observed in the idiom region. Passivized idioms in the figurative condition elicited significantly longer total reading times than the (same) phrases in the literal condition. Since the passivized idioms were preceded by minimal, neutral context, and the participants could not have known beforehand what the 'correct' interpretation would be, the inflated total reading times in the figurative condition are the result of longer regressive refixations, and are therefore suggestive of reanalysis, once the intended meaning of the phrase became evident.[4] Thus, unlike findings from studies on canonical idioms showing a lack of costly meaning competition in the idiom region, especially when idioms are preceded by biasing context (Conklin & Schmitt, 2008; Milburn & Warren, 2019; Siyanova-Chanturia et al., 2011), in this study, the slower processing of the passivized idioms in the figurative condition resembles the dominance bias effect typically observed for the subordinate meaning of a biased homonym (e.g., Dholakia et al., 2016; Duffy et al., 1988; Gottlob et al., 1999; Gunter et al., 2003), and is indicative of greater competition (from the literal meaning of the phrase).

Here, we also observed meaning resolution taking place at the point of disambiguation (keyword level), where a similar conclusion can be drawn based on the data from go-past reading time. That is, the literal meaning appears to be the dominant one, although suppressing it is fairly easy, as literal and figurative keywords were read with comparable ease in total reading time. In the Introduction, we mentioned that a similar

[4] For the idiom region we also examined regression likelihood (i.e., the probability of a regression back into the idiom region from subsequent areas of the sentence. As we found no effect of Condition ($\beta_{FIG} = 0.16$, SE = 0.13, $z = 1.20$; $p = 0.22$) and ($\beta_{CON} = 0.20$, SE = 0.13, $z = 1.49$; $p = 0.13$), we do not report these data. This, however, suggests that while idioms were not more likely to elicit a regression in the figurative than in the literal condition, regressive fixation durations in the former condition were significantly longer, as evidenced by the significantly longer total reading times.

facilitation for the literal meaning is sometimes observed in disambiguating regions that follow less familiar and/or less predictable canonical idioms (Cacciari & Corradini, 2015; Titone et al., 2019). The idioms we used in the present study were short and hence unpredictable (when in their canonical forms), but they were rated to be highly familiar. Because the idioms were passivized, they had an atypical and unfamiliar syntactic configuration, which lends support to our hypothesis that modifications render idioms less familiar, at least on the surface, delaying both idiom recognition and activation of the figurative meaning.

On the other hand, integrating the figurative meaning with a subsequent figurative context has been found to be easier, when canonical idioms are highly familiar or predictable (Cacciari & Corradini, 2015; Titone et al., 2019), suggesting that increased familiarity (and predictability) can increase the level of activation of the figurative meaning. Interestingly, despite the unfamiliarity generated by the passive construction in our study (which led to overall faster processing of the literal meaning), we too found that figurative keywords were read significantly faster, and literal keywords significantly more slowly, as familiarity increased. Given that this construct reflected the degree of familiarity with the figurative meaning of an idiom, the results demonstrate that knowing an idiom well leads to greater activation of its idiomatic meaning, even if the phrase is encountered in a novel way (see also Kyriacou et al., 2020, 2021). Greater phrasal frequency also had a facilitative effect on the go-past reading time of keywords and on the total reading time of both keywords and idioms, but it did not interact with Condition. This is unsurprising, since frequency merely reflects how often a phrase occurs in language, without differentiating between figurative and literal uses. We did not find any further effects of literalness or transparency, likely because our idioms were fairly balanced in both measures with little variability among the items. Further, there was no effect of semantic association strength on the processing of the keywords.

While current models of idiom processing do not address whether or how phrasal modifications influence meaning activation, we argue that the present data support a hybrid representation of idioms. First, and in line with the *Superlemma Hypothesis* (Sprenger et al., 2006), we found that activation of the figurative meaning can spread from the component words of an idiom (*bucket, kicked*). Crucially, however, in the current study the component words do not appear in their expected position in the phrase. Second, since activation of the figurative meaning would require an additional, intermediate step (the superlemma level) which does not apply to literal meanings, this might partly explain why the literal meaning was faster to activate in the present study. Third, the figurative meaning of more familiar idioms—which were arguably recognised sooner than less familiar ones—was available significantly more quickly, as evidenced by the total reading time of the keywords. This lends support to the *Configuration Hypothesis* (Cacciari & Tabossi, 1988), which posits that faster idiom recognition leads to faster activation of the figurative meaning. Finally, the influence exerted by idiom variables, and specifically by familiarity and frequency, line up with the *Multidetermined Model* (Libben & Titone, 2008), which predicts that such factors affect (canonical) idiom processing, potentially at different stages and timeframes during comprehension (Morid et al., 2021; Titone & Libben, 2014). It is worth highlighting, that the present data extend this model by including evidence from the processing of modified idioms, and further suggest that phrasal properties measured on canonical idiom forms (as was done in this study) are good predictors of processing ease (or difficulty) for modified versions. The direction of the effect seems to be the same: a factor that boosts idiom processing in a canonical form will likely benefit processing of a modified version. This is probably due to these factors tapping into lexical retrieval, making the figurative meaning more readily available even if modifications slow idiom recognition.

It is important to note that the current study investigated passivized idioms, and therefore, one should be cautious about generalising to other types of modifications. The passive voice is a complex grammatical structure, and therefore, it might influence processing differently to other modifications (e.g., adjective insertion). Specifically, as more attentional resources might be required for the processing of the passive, fewer might be available to attend to the processing of potentially nonliteral strings. It remains to be explored whether other types of manipulation influence meaning activation in the same way.

Conclusion

The current study demonstrates that the processing of passivized idioms is akin to that of unbalanced homonyms, as there was a clear bias for one interpretation, namely, the literal. However, competition and meaning resolution was evident not only in the ambiguous (idiom) region, but also in the disambiguating (keyword) region that followed. The literal meaning was available faster and was integrated more easily in the context; keywords associated to the literal meaning enjoyed a processing advantage in the earlier stages of processing (go-past reading time), and passivized idioms elicited significantly shorter total reading times when intended literally. The figurative meaning, on the other hand, produced similar fixation durations as anomalous control keywords in go-past time, but they were read as fast as the literal ones in total reading time. This suggests that the figurative meaning received some activation which facilitated its retrieval, although initial activation was smaller relative to the literal meaning, and longer processing time was required for its integration. We conclude that the processing costs associated with syntactic modifications in idioms, and in particular with the passive voice (Kyriacou et al., 2020; Mancuso et al., 2020), may indeed be partially accounted for, by greater activation of the literal meaning.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R Package Version*, *1*(7), 1–23.

Beck, S. D., & Weber, A. (2016). 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.

Beck, S. D., & Weber, A. (2020). Context and Literality in Idiom Processing: Evidence from Self-Paced Reading. *Journal of Psycholinguistic Research*, 1–27.

BNC Consortium. (2007). The British national corpus, version 3 (BNC XML Edition). *Distributed by Oxford University Computing Services on Behalf of the BNC Consortium*, *5*(65), 6.

Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition*, *1*(3), 343–346.

Brocher, A., Koenig, J.-P., Mauner, G., & Foraker, S. (2018). About sharing and commitment: the retrieval of biased and balanced irregular polysemes. *Language, Cognition and Neuroscience*, *33*(4), 443–466.

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.

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language, 27*(6), 668–683.

Caillies, S., & Butcher, K. (2007). Processing of idiomatic expressions: Evidence for a new hybrid view. *Metaphor and Symbol*, *22*(1), 79–108.

Canal, P., Pesciarelli, F., Vespignani, F., Molinaro, N., & Cacciari, C. (2017). Basic Composition and Enriched Integration in Idiom Processing: An EEG Study. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *43*(6), 928–943.

Carrol, G. (2021). Psycholinguistic approaches to figuration. In *Figurative Language–Intersubjectivity and Usage* (pp. 307–338). John Benjamins.

Carrol, G., & Conklin, K. (2014). Getting your wires crossed: Evidence for fast processing of L1 idioms in an L2. *Bilingualism: Language and Cognition*, *17*(04), 784–797.

Carrol, G., & Conklin, K. (2017). Cross language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea 'has legs.' *Bilingualism: Language and Cognition*, *20*(2), 299–317.

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. https://doi.org/10.1177/0023830918823230

Carrol, G., Conklin, K., & Gyllstad, H. (2016). Found in translation: the influence of the L1 on the reading of idioms in a L2. *Studies in Second Language Acquisition*, *38*(3), 403–443.

Colombo, L. (2014). The comprehension of ambiguous idioms in context. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 163–200). Psychology Press.

Conklin, K., & Carrol, G. (2021). Words go together like ‘bread and butter’: The rapid, automatic acquisition of lexical patterns. *Applied Linguistics*, *42*(3), 492–513.

Conklin, K., Pellicer-Sánchez, A., & Carrol, G. (2018). *Eye-Tracking: A Guide for Applied Linguistics Research*. Cambridge University Press.

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

Cronk, B. C., & Schweigert, W. A. (1992). The comprehension of idioms: The effects of familiarity, literalness, and usage. *Applied Psycholinguistics*, *13*(02), 131–146.

Cutting, J. C., & Bock, K. (1997). That’s the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition*, *25*(1), 57–71.

Dholakia, A., Meade, G., & Coch, D. (2016). The N400 elicited by homonyms in puns: Two primes are not better than one. *Psychophysiology*, *53*(12), 1799–1810.

Duffy, S. A., Kambe, G., & Rayner, K. (2001). The effect of prior disambiguating context on the comprehension of ambiguous words: Evidence from eye movements. In D. S. Gorfein (Ed.), *On the consequences of meaning selection: perspectives on resolving lexical ambiguity* (pp. 22–47). American Psychological Association.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, *27*(4), 429–446.

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.

Foraker, S., & Murphy, G. L. (2012). Polysemy in sentence comprehension: Effects of

meaning dominance. *Journal of Memory and Language*, *67*(4), 407–425.

Fraser, B. (1970). Idioms within a transformational grammar. *Foundations of Language*, *6*(1), 22–42.

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, *29*(2), 181–200.

Geeraert, K., Newman, J., & Baayen, R. H. (2017). Idiom variation: Experimental data and a blueprint of a computational model. *Topics in Cognitive Science*, *9*(3), 653–669.

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

Gibbs, R. W., & Nayak, N. P. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology*, *21*(1), 100–138.

Gibbs, R. W., Nayak, N. P., & Cutting, C. (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language*, *28*(5), 576–593.

Gorfein, D. S. (1989). Resolving semantic ambiguity. *Resolution of Semantic Ambiguity, Apr, 1988, Adelphi U, Garden City, NY, US*.

Gottlob, L. R., Goldinger, S. D., Stone, G. O., & Van Orden, G. C. (1999). Reading homographs: Orthographic, phonologic, and semantic dynamics. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(2), 561.

Gunter, T. C., Wagner, S., & Friederici, A. D. (2003). Working memory and lexical ambiguity resolution as revealed by ERPs: A difficult case for activation theories. *Journal of Cognitive Neuroscience*, *15*(5), 643–657.

Haeuser, K. I., Baum, S., & Titone, D. (2020). Effects of aging and noncanonical form presentation on idiom processing: Evidence from eye tracking. *Applied Psycholinguistics*, 1–27.

Hogaboam, T. W., & Perfetti, C. A. (1975). Lexical ambiguity and sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, *14*(3), 265–274.

Holsinger, E. (2013). Representing idioms: Syntactic and contextual effects on idiom processing. *Language and Speech*, *56*(3), 373–394.

Holsinger, E., & Kaiser, E. (2013). Processing (non) compositional expressions: Mistakes and recovery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 866.

Hyönä, J. (2011). *Foveal and parafoveal processing during reading.*

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, *32*(1), 37–55.

Klein, D. E., & Murphy, G. L. (2002). Paper has been my ruin: Conceptual relations of polysemous senses. *Journal of Memory and Language*, *47*(4), 548–570.

Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics*, *20*(1), 1–24.

Klepousniotou, E., Pike, G. B., Steinhauer, K., & Gracco, V. (2012). Not all ambiguous words are created equal: An EEG investigation of homonymy and polysemy. *Brain and Language*, *123*(1), 11–21.

Knoeferle, P. (2007). Comparing the time course of processing initially ambiguous and

unambiguous German SVO/OVS sentences in depicted events. In R. P. G. van Gompel, W. S. Murray, M. H. Fischer, & R. L. Hill (Eds.), *Eye Movements: a window on mind and brain* (pp. 517–533). Elsevier Science.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, *95*(1), 95–127.

Kyriacou, M., Conklin, K., & Thompson, D. (2020). Passivizability of Idioms: Has the Wrong Tree Been Barked Up? *Language and Speech*, *63*(2), 404–435. https://doi.org/10.1177/0023830919847691

Kyriacou, M., Conklin, K., & Thompson, D. (2021). When the Idiom Advantage Comes Up Short: Eye-Tracking Canonical and Modified Idioms. *Frontiers in Psychology*, *12*(1), 1–16. https://www.frontiersin.org/article/10.3389/fpsyg.2021.675046

Laurent, J.-P., Denhières, G., Passerieux, C., Iakimova, G., & Hardy-Baylé, M.-C. (2006). On understanding idiomatic language: The salience hypothesis assessed by ERPs. *Brain Research*, *1068*(1), 151–160.

Lenth, R. (2018). Emmeans: Estimated marginal means, aka least-square means. *R Package Version 1.2*.

Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition*, *36*(6), 1103–1121.

Mack, J. E., Meltzer-Asscher, A., Barbieri, E., & Thompson, C. K. (2013). Neural correlates of processing passive sentences. *Brain Sciences*, *3*(3), 1198–1214.

Mancuso, A., Elia, A., Laudanna, A., & Vietri, S. (2020). The Role of Syntactic Variability and Literal Interpretation Plausibility in Idiom Comprehension. *Journal of*

Psycholinguistic Research, *49*(1), 99–124. https://doi.org/10.1007/s10936-019-09673-8

Mancuso, A., & Laudanna, A. (2019). Objective frequency values of canonical and syntactically modified idioms: Preliminary normative data. *CEUR Workshop Proceedings*, *2481*.

Martin, C., Vu, H., Kellas, G., & Metcalf, K. (1999). Strength of discourse context as a determinant of the subordinate bias effect. *The Quarterly Journal of Experimental Psychology Section A*, *52*(4), 813–839.

McDonald, S. A., & Shillcock, R. C. (2003a). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, *14*(6), 648–652.

McDonald, S. A., & Shillcock, R. C. (2003b). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research*, *43*(16), 1735–1751.

McGlone, M. S., Glucksberg, S., & Cacciari, C. (1994). Semantic productivity and idiom comprehension. *Discourse Processes*, *17*(2), 167–190.

Meyer, A. M., Mack, J. E., & Thompson, C. K. (2012). Tracking passive sentence comprehension in agrammatic aphasia. *Journal of Neurolinguistics*, *25*(1), 31–43.

Milburn, E., & Warren, T. (2019). Idioms show effects of meaning relatedness and dominance similar to those seen for ambiguous words. *Psychonomic Bulletin & Review*, *26*(2), 591–598.

Morid, M., Bachar, N., & Sabourin, L. (2021). Capturing the multi-determined nature of idiom processing using ERPs. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *75*(2), 155.

Mueller, R. A. G., & Gibbs, R. W. (1987). Processing idioms with multiple meanings.

Journal of Psycholinguistic Research, *16*(1), 63–81.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407.

Nunberg, G. (1978). *The pragmatics of reference*. Indiana University Linguistics Club.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition*, *9*(3), 225–236.

Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. J. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior*, *17*(4), 465–477.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13.

Peterson, R. R., Burgess, C., Dell, G. S., & Eberhard, K. M. (2001). Dissociation between syntactic and semantic processing during idiom comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1223.

Pickering, M. J., & Frisson, S. (2001). Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(2), 556.

R Core Team. (2021). *R: A language and environment for statistical computing*. http://www.r-project.org/

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506.

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory and Language*, *33*(4), 527–544.

Rommers, J., Dijkstra, T., & Bastiaansen, M. (2013). Context-dependent semantic processing in the human brain: Evidence from idiom comprehension. *Journal of Cognitive Neuroscience*, *25*(5), 762–776.

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior*, *20*(1), 120–136.

Simpson, G. B. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin*, *96*(2), 316.

Sinclair, J. (2011). *Collins COBUILD idioms dictionary*. Harper Collins.

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*, *27*(2), 251–272.

Snider, N., & Arnon, I. (2012). A unified lexicon and grammar? Compositional and non-compositional phrases in the lexicon. In D. Divjak & S. T. Gries (Eds.), *Frequency effects in language representation* (pp. 127–163). Walter de Gruyter.

Sprenger, S. A., Levelt, W. J. M., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, *54*(2), 161–184.

Strandburg, R. J., Marsh, J. T., Brown, W. S., Asarnow, R. F., Guthrie, D., & Higa, J. (1993). Event-related potentials in high-functioning adult autistics: Linguistic and nonlinguistic visual information processing tasks. *Neuropsychologia*, *31*(5), 413–434.

Swinney, D. A. (1981). Lexical processing during sentence comprehension: Effects of higher

order constraints and implications for representation. In *Advances in Psychology* (Vol. 7, pp. 201–209). Elsevier.

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, *18*(5), 523–534.

Tabossi, P., Fanari, R., & Wolf, K. (2008). Processing idiomatic expressions: Effects of semantic compositionality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 313.

Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory & Cognition*, *37*(4), 529–540.

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

Titone, D. A., & Libben, M. R. (2014). Time-dependent effects of decomposability, familiarity and literal plausibility on idiom priming: A cross-modal priming investigation. *The Mental Lexicon*, *9*(3), 473–496.

Titone, D. A., Lovseth, K., Kasparian, K., & Tiv, M. (2019). Are Figurative Interpretations of Idioms Directly Retrieved, Compositionally Built, or Both? Evidence from Eye Movement Measures of Reading. *In Print*.

Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 153–172). John Benjamins Publishing Company.

Van de Voort, M. E. C., & Vonk, W. (1995). You Don’t Die lmmecliately When You Kick an Empty Bucket: A Processing View on Semantic and Syntactic Characteristics of ldioms. In M. Everaert, E.-J. Van der Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms:*

Structural and Psychological Perspectives (pp. 283–300). Psychology Press.

Vespignani, F., Canal, P., Molinaro, N., Fonda, S., & Cacciari, C. (2010). Predictive mechanisms in idiom comprehension. *Journal of Cognitive Neuroscience*, *22*(8), 1682–1700.

Vu, H., Kellas, G., & Paul, S. T. (1998). Sources of sentence constraint on lexical ambiguity resolution. *Memory & Cognition*, *26*(5), 979–1001.

Appendix 1

Full List of Idioms (in their Canonical Form) Along with their Keywords in the Figurative, Literal, and Control Conditions.

No.	Idiom	Keyword		
		Figurative	*Literal*	*Control*
1	Blow a fuse	temper	power	life
2	Break the ice	joke	cocktail	wall
3	Chew the fat	tale	bite	ball
4	Clear the air	reconcile	breathe	jump
5	Cook your goose	fraud	recipe	radio
6	Count the cost	consequences	debts	teeth
7	Crack the whip	pupils	horses	eyes
8	Cross the line	fight	alarm	sun
9	Deliver the goods	retailer	courier	layer
10	Drag your feet	study	walk	give
11	Draw the line	husband	property	thought
12	Drop the ball	project	team	door
13	Fan the flames	war	fire	day
14	Flex your muscles	influence	strength	life
15	Grease the wheels	plans	trolley	music
16	Hit the bottle	drunk	slashed	rich
17	Hit the roof	anger	ceiling	desk
18	Hit the wall	athlete	driver	paper
19	Kick the bucket	body	water	time
20	Know the score	issue	game	cut
21	Lick your wounds	office	vet	age
22	Line your pockets	apartment	wallet	glass
23	Lose your marbles	care	game	area
24	Miss the boat	opportunistic	punctual	vivid
25	Miss the mark	debate	shot	key
26	Move the goalposts	deal	match	year
27	Play the game	board	team	room
28	Prepare the ground	ideas	bulbs	love
29	Pull one's leg	prank	pain	book
30	Pull the strings	firm	puppet	week
31	Pull your weight	colleagues	trainer	pen
32	Rock the boat	transaction	sailing	cream
33	Scratch the surface	secrets	treasures	clouds
34	Set the pace	sales	speed	hair
35	Smell a rat	felony	rodent	bud
36	Spill the beans	lawyer	mop	sky
37	Stem the tide	drinking	flooding	chair
38	Tie the knot	wife	boat	idea
39	Tighten your belt	expenses	calories	trees
40	Touch a nerve	topic	needle	hat
41	Turn the page	adventure	illustration	dish
42	Turn the tables	authority	space	world
43	Twist one's arm	club	hospital	air

44	Twist the knife	dispute	bleeding	wind
45	Waste your breath	convince	revive	cook

Appendix 2

Outputs Of All Final Models Across All Eye-Tracking Measures Investigated.

Idiom Region

	Total Reading Time			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	6.65	0.05	135.86	**<0.001**
CONDITION [Idiomatic]	0.05	0.02	2.34	**0.02**
CONDITION [Control]	0.05	0.02	2.52	**0.012**
Frequency	-0.1	0.03	-3.41	**0.001**
Trial Sequence Number	0	0	-5.11	**<0.001**
CONDITION [Idiomatic] * Frequency	0.01	0.03	0.24	0.809
CONDITION [Control] * Frequency	0.03	0.03	1.21	0.228
Random Effects				
σ^2	0.15			
$\tau_{00\ PPT}$	0.08			
$\tau_{00\ ITEM}$	0.01			
ICC	0.38			
N_{PPT}	45			
N_{ITEM}	44			
Observations	1966			

Keyword Region

	First Pass Gaze Duration			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	5.35	0.02	244.14	**<0.001**
CONDITION [Idiomatic]	0.02	0.02	1.03	0.305
CONDITION [Control]	0.01	0.02	0.52	0.604
Random Effects				
σ^2	0.14			
$\tau_{00\ PPT}$	0.01			
ICC	0.07			
N_{PPT}	45			
Observations	1604			

	Go Past Reading Time			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	5.46	0.03	181.78	**<0.001**
CONDITION [Idiomatic]	0.06	0.03	2.01	**0.045**
CONDITION [Control]	0.08	0.03	2.69	**0.007**
Frequency	-0.07	0.03	-2.34	**0.019**
CONDITION [Idiomatic] * Frequency	0.04	0.04	0.91	0.363
CONDITION [Control] * Frequency	0.03	0.04	0.73	0.467
Random Effects				
σ^2	0.23			
$\tau_{00\ PPT}$	0.02			
ICC	0.08			
N_{PPT}	45			
Observations	1604			

	Total Reading Time			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	5.54	0.04	136.15	**<0.001**
CONDITION [Idiomatic]	0.04	0.03	1.54	0.124
CONDITION [Control]	0.15	0.03	5.42	**<0.001**
Familiarity	0.08	0.03	3	**0.003**
Frequency	-0.07	0.03	-2.54	**0.011**
Trial Sequence Number	0	0	-2.83	**0.005**
CONDITION [Idiomatic] * Familiarity	-0.14	0.04	-3.73	**<0.001**
CONDITION [Control] * Familiarity	-0.04	0.04	-1.18	0.236
CONDITION [Idiomatic] * Frequency	0	0.04	0.04	0.964
CONDITION [Control] * Frequency	0.03	0.04	0.69	0.491
Random Effects				
σ^2	0.21			
$\tau_{00\ PPT}$	0.04			
ICC	0.15			
N_{PPT}	45			
Observations	1604			

	Total Reading Time (Figurative Condition as Baseline)			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	5.58	0.04	135.79	**<0.001**
CONDITION [Literal]	-0.04	0.03	-1.54	0.124
CONDITION [Control]	0.11	0.03	3.87	**<0.001**
Familiarity	-0.06	0.03	-2.27	**0.023**
Frequency	-0.07	0.03	-2.4	**0.017**
Trial Sequence Number	0	0	-2.83	**0.005**
CONDITION [Literal] * Familiarity	0.14	0.04	3.73	**<0.001**
CONDITION [Control] * Familiarity	0.09	0.04	2.56	**0.011**
CONDITION [Literal] * Frequency	0	0.04	-0.04	0.964
CONDITION [Control] * Frequency	0.03	0.04	0.63	0.526
Random Effects				
σ^2	0.21			
$\tau_{00\ PPT}$	0.04			
ICC	0.15			
N_{PPT}	45			
Observations	1604			

Post-Disambiguating Region

	Total Reading Time			
Predictors	*β*	*SE*	*t*	*p*
(Intercept)	6.7	0.06	121.68	**<0.001**
CONDITION [Idiomatic]	0.04	0.03	1.23	0.218
CONDITION [Control]	0.18	0.03	5.91	**<0.001**
Trial Sequence Number	0	0	-2.74	**0.006**
Random Effects				
σ^2	0.29			
$\tau_{00\ PPT}$	0.1			
ICC	0.25			
N_{PPT}	45			
Observations	1967			