Robust processing advantage for binomial phrases with variant conjunctions

Suphasiree Chantavarin[1], Emily Morgan[2], & Fernanda Ferreira[3]

[1] Faculty of Psychology, Chulalongkorn University, Bangkok, Thailand

[2] Department of Linguistics, University of California, Davis, CA, USA

[3] Department of Psychology, University of California, Davis, CA, USA

Address correspondence to:

Suphasiree Chantavarin

Faculty of Psychology, Chulalongkorn University

Rama 1 Road, Wangmai, Pathumwan

Bangkok 10330, Thailand

suphasiree.c@chula.ac.th

© Wiley Online Library, 2022. This paper is not the copy of record and may not exactly replicate the authoritative document published in the Wiley Online Library journal. The article is available online at: https://doi.org/10.1111/cogs.13187

Abstract

Prior research has shown that various types of conventional multi-word chunks are processed faster than matched novel strings, but it is unclear whether this processing advantage extends to variant multi-word chunks that are less formulaic. To determine whether the processing advantage of multi-word chunks accommodates variations in the canonical phrasal template, we examined the robustness of the processing advantage (i.e., predictability) of binomial phrases with noncanonical conjunctions (e.g., *salt and also pepper; salt as well as pepper*). Results from the cloze study (Experiment 1) showed that there was a high tendency of producing the canonical conjunct (*pepper*), even in the binomials that contained non-formulaic conjunctions. Consistent with these findings, results from two eye tracking studies (Experiment 2a and 2b) showed that canonical conjuncts were read faster than novel conjuncts that were matched on word length (e.g., *paprika*), even in the binomials with variant conjunctions. This robust online processing advantage was replicated in a self-paced reading study that compared all three conjunction types (Experiment 3). Taken together, these findings show that binomials with variant function words also receive facilitated processing relative to matched novel strings, even though both types of strings are neither conventional nor relatively frequent. Exploratory analyses revealed that this processing speed advantage was driven by the lexical-semantic association between the canonical conjuncts (*salt–pepper*), rather than lexical and phrasal frequency. Overall, these results highlight flexibility in the processing of multi-word chunks that current models of multi-word storage and processing must take into account.

Keywords: multi-word chunks; binomials; predictability; eye tracking; self-paced reading; sentence processing

1. Introduction

Corpus analyses estimate that a large portion of the language produced by native speakers consists of recurrent patterns (Erman & Warren, 2000; Jackendoff, 1995). These formulaic word combinations such as idioms (*break the ice*), conversational formulae (*have a nice day*), collocations (*spend time*), and binomial phrases (*salt and pepper*) are so familiar and frequently used that they may be stored in our memory and can be automatically retrieved during language processing. At the same time, a hallmark of language is that it permits infinite novel linguistic expressions conforming to grammatical rules, and language can be used in creative, idiosyncratic ways that deviate from the conventional forms. A speaker might say, for example, "Pass me the salt as well as the pepper". When a listener hears this sentence, is the non-formulaic phrase *salt as well as the pepper* treated as an unfamiliar novel phrase, or is does it still benefit from speeded activation based on her knowledge of the familiar binomial phrase *salt and pepper*? The current study investigates whether variant binomial phrases—those that have a non-formulaic conjunction but still retain a similar meaning to conventional binomial phrases—also receive facilitated processing relative to novel phrases, despite having a noncanonical surface form.

Under the words-and-rules approach to language (Chomsky, 1965; Pinker, 1999), words are posited as the building blocks of language learning and use, under the assumption that the mental lexicon stores representations of words and morphemes separate from combinatory grammatical rules. Larger linguistic units such as phrases are thought to be computed on the fly, rather than stored in memory. In contrast, other approaches to language such as usage-based (Bybee, 1998, 2006; Tomasello, 2009; Ellis, 2002), construction-based (Goldberg, 2006; Fillmore, Kay, & O'Connor, 1988; Dabrowska, 2014), and cognitive grammar (Langacker, 1986, 1987; Taylor, 2002) propose that linguistic units of various sizes are represented and processed within a single system (Jackendoff & Audring, 2019). Under the usage-based views, the degree of prior exposure probabilistically determines the linguistic units that are represented in long-term memory (Arnon & Snider, 2010; Snider & Arnon, 2012). The main line of evidence for this usage-based approach comes from the literature on multi-word chunks, which has shown mounting evidence that the language system is sensitive to the frequency of not only words but also phrases. Numerous studies have shown that various types of conventional multi-word chunks, including idioms, binomials, and collocations, exhibit a processing speed advantage over comparable novel phrases in language comprehension (Swinney & Cutler, 1979; Siyanova-Chanturia et al., 2011; Tremblay et al., 2011; Carrol & Conklin, 2019; Arnon & Snider, 2010; Sosa & MacFarlane, 2002) as well as production (Bybee & Scheibman, 1999; Bannard & Matthews, 2008; Arnon & Cohen Priva, 2013; Janssen & Barber, 2012; Siyanova-Chanturia & Janssen, 2018). These findings support the view that the building blocks of language acquisition and usage may constitute multi-word chunks (Arnon, McCauley, & Christiansen, 2017; Christiansen & Arnon, 2017), and also suggests that phrases may have equal footing with words in the lexicon. Yet, many questions remain regarding how these larger linguistic units are represented in the mind, and how they are accessed during sentence processing. If multi-word chunks are potential building blocks of language, it is important to determine the cognitive mechanisms that map the language input received in real-time processing onto the multi-word chunks stored from past experience.

Three main theoretical approaches have been put forward to explain how recurrent multi-word patterns are stored and retrieved from semantic memory. The first one posits that multi-word chunks may be stored and directly retrieved as fixed lexical units, without access to their constituent parts (Sinclair, 1991; Bobrow & Bell, 1973; Swinney & Cutler, 1979). However, several studies have shown that the internal syntactic and semantic components of idiomatic expressions are accessible (Cutting & Bock, 1997; Konopka & Bock, 2009; Sprenger, Levelt, & Kempen, 2006; Peterson et al., 2001), challenging the view of multi-word chunks as fixed lexicalized units. Conversely, a second approach proposes that multi-word chunks are initially processed compositionally, in which comprehenders incrementally assign idiomatic meanings to the constituent words (Gibbs, Nayak, & Cutting, 1989). Finally, a third approach views multi-word chunks as neither fully lexicalized chunks nor fully compositional strings. These hybrid approaches posit that some phrase-level representation of multi-word chunks may be stored—perhaps as constructions (i.e., highly formulaic pairings of linguistic form and meaning; Goldberg, 2003, 2006; Bybee, 2006; Bod, 2009) or as something akin to a mental template (Siyanova-Chanturia et al., 2017; Vespignani et al., 2010)—and that the processing of multi-word chunks involves recognizing these stored linguistic patterns. An example of this view is the Configuration Hypothesis of idiom processing, which proposes that an idiom's figurative meaning is activated when encountering a portion of the string that serves as the "key" or recognition point that the string being processed is idiomatic (Cacciari & Tabossi, 1988). Entrenched multi-word chunks may therefore be retrieved from semantic memory when language users receive enough perceptual information to recognize the familiar phrasal configuration and to map the input onto a frequency-based, co-activated network of form-meaning representations (Arnon & Snider, 2010; Snider & Arnon, 2012; Arnon & Cohen Priva, 2014; Dabrowska, 2014). A similar but slightly different view is that multi-word chunks are processed by matching the encountered input to a pre-activated mental template that precisely matches the unfolding configuration, and crucially, this template matching mechanism is not observed for comparable strings that are non-conventional and unfamiliar, suggesting that this mechanism is unique to the processing of frequent, familiar strings (Vespignani et al., 2010; Siyanova-Chanturia et al., 2017; Molinaro & Carreiras, 2010).

An outstanding issue is that there is a blurry distinction between the re-use of conventional, pre-stored multi-word chunks and the online computation of novel, unfamiliar strings (Buerki, 2018; Langlotz, 2006). To make matters more complicated, multi-word chunks are not necessarily syntactically and lexically frozen. Variant forms of formulaic sequences have been extensively documented in the language, especially idiom variations (e.g., *kick the proverbial bucket;* Fraser, 1970; Gibbs, 1990; Moon, 1998; Schröder, 2015; Wulff, 2008). Since these variant multi-word chunks are neither highly conventional nor completely novel strings, how can they be accounted for under the existing models of multi-word representation and processing? For instance, under the hybrid views, would variant multi-word chunks, which are similar but not identical to the canonical, frequent form, still benefit from pre-activation and pattern recognition? An approach that can help resolve this issue is to examine the kinds of modifications that reduce the conventionality of multi-word chunks without compromising their meaning, and assessing how these alterations impact their processing speed advantage. The next section describes how examining the processing advantage in variant multi-word chunks can provide novel insights regarding the mechanisms underlying the mapping and representation of recurrent multi-word chunks that characterize fluent and native-like language use (Pawley & Syder, 1983).

1.1 The processing of variant multi-word chunks

Despite the documentation of recurrent, conventional multi-word chunks, these strings also occur in variant forms that are neither highly conventional nor frequent, yet still retain the core elements (e.g., content words) of the formulaic configuration. Variations have been extensively documented for idioms (Fraser, 1970; Schröder, 2015; Moon, 1998; Langlotz, 2006; Wulff, 2008), some of which permit a certain degree of variability in the syntax and the component words (Cacciari & Glucksberg, 1991; McGlone, Glucksberg, & Cacciari, 1994; Kyriacou, Conklin, & Thompson, 2019; Mancuso et al., 2020). Thus, even though multi-word chunks may be stored in the lexicon with an internal syntactic structure, they are by no means frozen. Grammatical constraints do govern these stored multi-word chunks, just as they apply to structures constructed on the fly (Culicover, Jackendoff, & Audring, 2017; Tabossi, Wolf, & Koterle, 2009).

An open empirical question is whether variant forms of multi-chunks also show a processing advantage. Existing models of multi-word chunk representation and processing all predict that the conventional forms of multi-word chunks will be processed faster than novel strings, but these models do not directly account for the processing of multi-word chunks that contain grammatical or lexical variations. Nevertheless, most of the theoretical approaches outlined in the previous section would probably not predict such a processing advantage for variant multi-word chunks, for the following reasons. According to approaches which assume holistic storage and retrieval (Sinclair, 1991; Bobrow & Bell, 1973; Swinney & Cutler, 1979), variant multi-word chunks would not be processed faster than comparable novel phrases because the variant forms do not match the stored lexical representation. Approaches which assume that exact frequency of exposure is a key determinant of the processing speed of multi-word chunks (Arnon & Snider, 2010; Arnon & Cohen Priva, 2014) would also not predict a processing advantage because variant multi-word chunks have low phrasal frequency, so in that aspect they are comparable to novel strings. Finally, hybrid approaches which assume that the mapping of multi-word chunks requires a unique match to the stored mental template (Vespignani et al., 2010; Molinaro & Carreiras, 2010; Siyanova-Chanturia et al., 2017) would not predict a processing advantage for variant multi-word chunks because their surface form does not precisely match the canonical mental template.

This latter prediction based on the template matching view has received empirical support from studies that used electrophysiological recordings to examine the processing of multi-word chunks. One study on Spanish collocations compared the processing of literal and figurative collocations whose final word was either the canonical completion or a synonym (*an honor to the truth/reality; against wind and tide/cold*; Molinaro & Carreiras, 2010). Beyond the N400 effect that was sensitive to the semantic/literality manipulation, they also found a modulation in the P300 component that was sensitive to the contextual expectancy of the final word (canonical completion or synonym). This was interpreted as reflecting the retrieval and pre-activation of a mental template which occurs only when there is an exact match between the string being processed and the familiar phrasal configuration (for a similar study but on idioms, see Vespignani et al., 2010). In line with these findings, another ERP study of binomial phrases found that the second conjunct (*fork*) in conventional, frequent binomials (*knife and fork*) elicited smaller N400 amplitudes and larger P300 amplitudes compared to infrequent, unfamiliar phrases whose first conjunct was replaced with another word that had comparable semantic association with the second conjunct (*spoon and fork*; Siyanova-Chanturia et al., 2017). Crucially, in Experiment 1b they also found that when the canonical conjuncts were presented without the conjunction (*knife, fork*), there were no differences in the amplitudes of the P300 and the N400 between the two conditions, suggesting that the template matching mechanism only applies to multi-word chunks in their canonical surface form.

Despite these findings, there are reasons to expect that variant multi-word chunks may confer a processing advantage over novel strings, despite their non-idiomaticity and low phrasal frequency. First, the predictability of the constituents in multi-word chunks—typically operationalized as the cloze probability of the phrase-final word—is correlated with phrase-specific variables such as lexical-semantic association (LSA) in binomials, and potentially other measures like mutual information in collocations and phrase familiarity in idioms, which jointly contribute to their formulaic processing advantage (Carrol & Conklin, 2019). In general, these measures are hard to distinguish from each other because they all reflect the degree of experience with each phrase, based on the frequency of exposure and subsequent entrenchment of that phrase in memory (Tremblay & Tucker, 2011; Siyanova-Chanturia et al., 2011). As long as some of these phrase-specific features are retained, variant multi-word chunks may be processed faster than matched novel strings, even if their surface form is neither conventional nor frequent. For example, *salt and also pepper* has a low exact phrasal frequency but might still show a formulaic processing advantage due to the semantic association between *salt* and *pepper* and/or other factors that contribute to predictability.

Second, a few studies have specifically found a processing advantage for some types of variant multi-word chunks. In an eye tracking study that presented idioms (e.g., *kick the bucket*) in sentence contexts that were biased towards their figurative meaning, passivized idioms (*the bucket was kicked)* were read faster than literal, control phrases (*the apple was kicked*), suggesting that idioms with grammatical variations still retained their figurative meaning (Kyriacou, Conklin, & Thompson, 2019; see also Mancuso et al., 2020). Similarly, another eye tracking study found that collocations with intervening words (*provide some of the information*) were read faster than matched novel strings (*compare some of the information*), indicating that the statistical co-occurrence of collocates (*provide*, *information*) impacts their processing speed even when they are nonadjacent (Vilkaitė, 2016). This processing advantage for nonadjacent collocations has also been observed in an ERP study which found that in complex function words with an inserted adjective (*in the capable hands of)*, the noun *(hands)* elicited a smaller N400 in this variant form than in their canonical form without the adjective (*in the hands of*), suggesting that variant collocations received facilitated processing from the preceding adjective, rather than incurring a processing cost due to a mismatch with the canonical mental template (Molinaro et al., 2013).

In sum, it remains unclear when variant forms of multi-word chunks confer a processing advantage in terms of pre-activation and pattern recognition, despite their non-idiomaticity. Some existing approaches to multi-word storage and processing would not predict such a processing advantage, whereas other studies have suggested that multi-word chunks with grammatical variation and lexical insertion do exhibit a processing speed advantage (Kyriacou, Conklin, & Thompson, 2019; Mancuso et al., 2020; Vilkaitė, 2016). Moreover, results from ERP studies on variant multi-word chunks point to different conclusions regarding the degree of flexibility in the template matching mechanism presumed for multi-word chunks. Some studies indicate that the processing advantage of multi-word chunks is observed only when there is an exact, categorical match with the canonical mental template stored in memory (Molinaro &

Carreiras, 2010; Vespignani et al., 2010; Siyanova-Chanturia et al., 2017), whereas other studies suggest that a precise match may not necessarily be required (Molinaro et al., 2013). Besides the template matching view, other hybrid approaches such as the Configuration Hypothesis would predict that variant forms should activate the idiomatic meaning as long as the content words are present (Cacciari & Tabossi, 1988), yet it remains unclear what kinds of variant forms would retain a processing advantage over novel forms.

1.2 Current study

To better understand the representation and processing of variant multi-word chunks, the present study examined how manipulating the phrasal template affects the processing advantage in a type of multi-word chunk that has a highly conventional, well-defined template: binomial expressions. Binomial phrases are formulaic sequences of the form *A and B* that often occur in a rigid surface form (e.g., *salt and pepper, bride and groom*). Binomials vary on the degree of reversibility, with some occurring exclusively in one order (e.g., *ladies and gentlemen*) whereas others occur more equally in both orders (*cats and dogs*; Malkiel, 1959). Processing speed advantages in terms of faster reading times have been found for binomials in their canonical surface form (*bride and groom*) compared to their reversed order (*groom and bride*; Siyanova-Chanturia, Conklin, & van Heuven, 2011). The degree of lexical-semantic association between the conjuncts and the contextual predictability of the phrase have also been implicated in the processing speed advantage for attested binomials (*king and queen*), compared to novel control phrases in which the first or second conjunct was changed (*prince and queen, king and prince;* Carrol & Conklin, 2019). Because of their relatively fixed surface form, binomials are an ideal test case for the degree of flexibility in the mapping of multi-word chunks. Compared to other types of literal multi-word chunks such as collocations and lexical bundles, binomials typically occur in a rigid phrasal configuration (*A and B*), in a particular morphological inflection (#*brides and grooms*), and without the presence of determiners (#*the bride and the groom;* Lambrecht, 1984), suggesting that they may be less open to lexical variations than other types of compositional multi-word chunks.

The current study investigated the processing of binomials that contain nonformulaic conjunctions (*and also* and *as well as*), which perform the same function as *and* in linking two constituents. In three experiments, we tested whether comprehenders still show a processing advantage for variant binomials (*salt and also pepper; salt as well as pepper*) relative to novel control phrases (*salt and also paprika; salt as well as paprika*). In Experiment 1, we used the cloze task (Taylor, 1953) to investigate whether people's tendency to produce the binomial's canonical second conjunct (*pepper*) is robust even when the conjunction is non-formulaic. We then conducted two eye tracking studies (Experiment 2a and 2b) and one self-paced reading study (Experiment 3) to investigate whether the processing speed advantage for binomials, here measured via reading times on the second conjunct, is still observed even in the variant binomials.

Given that the conjunction *and* is a defining feature in the phrasal configuration of binomials, and based on the findings from Experiment 1b from Siyanova-Chanturia et al.'s (2017) study which suggested that binomials confer a processing advantage only in their canonical surface form (*A and B*), one possible outcome is that binomials with variant conjunctions would not exhibit a processing advantage over novel, control phrases. Specifically, for variant binomials (*salt and also pepper, salt as well as pepper*), there would be a low probability of producing the binomial's second conjunct (*pepper)* in a cloze task (Experiment 1), and reading times on this conjunct would be comparable to, or even slower than, a novel conjunct matched in word length (*paprika*; Experiments 2 and 3). On the other hand, if variant binomials and their component words did show a processing advantage, this would provide evidence that the processing of multi-word chunks accommodates a certain degree of variability in their surface form configuration.

2. Experiment 1: Cloze study

In Experiment 1, we conducted a cloze completion study to test the robustness of the processing advantage for variant binomials. This study assessed the likelihood of producing the canonical second conjunct (e.g., *pepper*) in binomial phrases with non-canonical conjunctions (*and also*, *as well as*). We chose the cloze task because it directly measures people's predictions of the second conjunct.

2.1 Participants

Sixty-six undergraduates at the University of California, Davis participated in the experiment for course credit. Participants gave informed consent to take part in the study. All participants were native speakers of English and did not acquire any other language before the age of 5.

2.2 Materials

The experimental stimuli consisted of 36 sentences, each sentence containing a binomial (e.g., *salt and pepper*) that varied in conjunction type (<u>*and*</u> , <u>*and also*</u> , <u>*as well as*</u>). To create the stimuli, we first selected attested binomials to serve as the baseline condition. Because we will measure the ease of semantic processing on the second conjunct relative to a novel conjunct in Experiments 2 and 3, we specifically selected binomials conjoining two concrete nouns to ensure that the binomials had a transparent, decomposable meaning. We then constructed a sentence context for each binomial. Variant binomials were created by replacing the conjunction <u>*and*</u> with <u>*and also*</u> (e.g., *salt and also pepper*) and <u>*as well as*</u> (e.g., *salt as well as pepper*). For some of the binomial items (19/36 items), we added a determiner to ensure that the sentence was grammatical (e.g., *bride and also the groom*), and in the cloze task these items were presented with the appropriate determiner (e.g., *…the bride and also the* ____). The 36 experimental items described above were counterbalanced across three lists so that participants saw all 36 items but they saw only one version of each item. In each list, the experimental items were presented with 64 filler sentences. Participants were randomly assigned to one of the three lists.

2.3 Procedure

Participants were presented with sentence fragments up to, but not including, the second conjunct (e.g., *The meat was bland, so the chef reached for the salt and also* ___), and they were instructed to type one word that best completed the sentence. Participants' responses were not timed, and they were free to proceed through the items at their own pace.

2.4 Results

Mean cloze probabilities of the canonical conjuncts (e.g., *pepper*) for each condition were similar across the different conjunction conditions (see Fig. 1). To compare the cloze probabilities between the <u>And</u> condition (*M* = 86%, *SD* = 35%, range = 27-100%), <u>Also</u> condition (*M* = 80%, *SD* = 40%, range = 23-100%), and the <u>As Well As</u> condition (*M* = 76%, *SD* = 43%, range = 18-100%), we constructed Bayesian mixed-effects logistic models using the *brms* package (version 2.16.1; Bürkner, 2018) in R (version 4.1.1; R Core Team, 2021). Conjunction Type was entered as a fixed effect, and the model included random intercepts and random slopes by subjects and by items. Conjunction Type was coded using treatment contrasts, with And as the baseline (Intercept): the first contrast compared Also to And, and the second contrast compared As Well As to And. We computed all possible pairwise comparisons between conditions using the *hypothesis()* function, with *p* values adjusted using the Bonferroni correction. The analyses indicated evidence that the mean cloze probability of the canonical conjunct in the And condition was higher than the Also condition (Estimate = -2.44, 95% CI = [-4.55, -2.24]) and the As Well As condition (Estimate = -3.44, 95% CI = [-4.81, -2.26]). There was no difference between the Also condition and the As Well As condition (Estimate = -0.16, 95% CI = [-0.78, 0.52]).[1] The data, materials, and analysis scripts for this experiment can be found on OSF: https://osf.io/6y8xb/

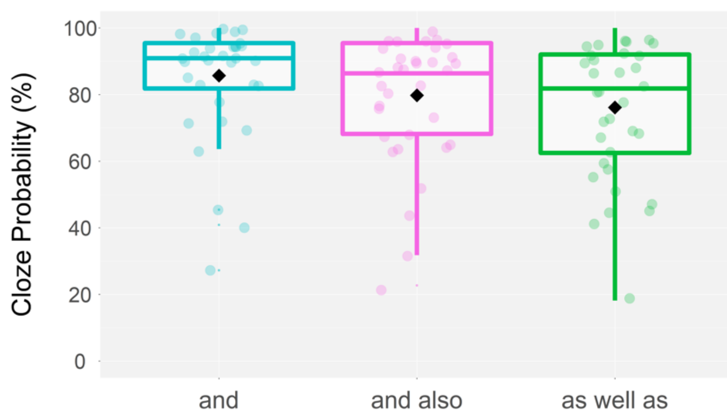

Figure 1. Boxplots of the mean cloze probabilities of the canonical conjunct (e.g., *pepper*) for each conjunction type in the cloze study (Experiment 1). Each boxplot denotes the interquartile range with the median line for the cloze responses (averaged across subjects and items) for that condition. The jittered dots overlaid on each boxplot show the mean cloze probabilities (averaged across subjects) for each of the 36 items within each condition, and the black diamond shows the mean cloze probability for each condition (And: *M* =86%, *SD* = 35%; And Also: *M* = 80%, *SD* = 40%; As Well As: *M* = 76%, *SD* = 43%).

2.5 Discussion

The descriptive statistics of the results indicated that the probability of producing the canonical conjunct (e.g., *pepper)* was high (>75%) across all conjunction types. The high cloze probability of the canonical conjunct in the And condition (86%) confirmed that the selected binomials were highly familiar and frequent strings. Crucially, the high cloze probability of the canonical conjunct in the And Also condition (80%) and As Well As condition (76%) suggests that overall, people were highly likely to produce the canonical conjunct despite encountering the variant conjunction. Overall, these findings suggest that the processing advantage for binomial phrases is relatively robust even when the function word (i.e., conjunction) in the canonical binomial template has been changed, supporting the account that the mapping of multi-word chunks is a relatively flexible process that can tolerate a certain degree of surface form variation.

The analyses also revealed that people were slightly less likely to produce the canonical conjunct (e.g., *pepper*) in variant binomial fragments conjoined with *and also* and *as well as*, compared to those conjoined with the formulaic conjunction *and*. The two variant conjunctions led to comparable rates of producing the canonical conjunct, which is attributable to their both being variant, non-*and* conjunctions. Although our research question pertains the processing advantage of variant binomials over novel strings rather than comparing the magnitude of the processing advantage in variant binomials compared to the canonical form, it is interesting to note that the small but reliable difference in cloze probability between the canonical form and the variant forms (6-10% drop) suggests that people are sensitive to the specific conjunction that is used in binomial phrases. In the non-*and* conjunctions, people were slightly more likely to producing other lexical candidates instead of the canonical conjunct.

Another interesting aspect of the results is there was greater variability and dispersion in cloze probabilities (across items) for the variant conjunctions (*and also*, *as well as*) compared to the responses for *and* (see Figure 1). The unexpected large variance in cloze probability may be due to the specific context associated with each sentence. Although a small percentage (7.6%) of the participants did tend to produce novel, noncanonical conjuncts in all three conditions, most participants responded with the canonical conjunct across items and conditions (see Appendix S2). Additionally, a few of the experimental items had a consistently high completion rate of the canonical conjunct for all the Conjunction conditions (e.g., *washer and / and also / as well as dryer*), whereas a few other items had a consistently low completion rate for all conditions (e.g., *peas and / and also / as well as carrots*). Thus, although the binomial stimuli we selected were familiar and frequent—as indicated by their relatively high phrasal frequency in COCA corpus

(Davies, 2008; see Table 1 for the distributional properties of the stimuli)—the binomial items varied in predictability, regardless of the conjunction.

In addition to assessing the mapping of binomial phrases in a production task, we were also interested in examining the flexibility of this mapping process during online comprehension of the binomials. Measuring the binomial processing advantage in terms of reading times provides a window into the real-time cognitive processes that are engaged in the processing of variant multi-word chunks. This also allows us to investigate specific factors that may contribute to the lexical predictability of variant binomials (e.g., word frequency of the second conjunct, lexical-semantic association between the conjuncts, and typicality of the phrase), beyond the cloze probability of the canonical conjunct. In the next two eye tracking studies, we therefore tested whether the variant conjunction affected reading times on the second conjunct.

3. Experiment 2a: Eye tracking study (comparing ***and*** to ***and also***)

In Experiment 2a, we varied the surface form of binomials (*salt and pepper*) by adding an intervening word (*also*) before the second conjunct, and we used eye tracking to investigate whether variant binomials confer a processing speed advantage over matched novel phrases. In order to compare the processing of binomials to the processing of novel phrases, we substituted the final word in the binomial phrase (*pepper*) with a semantically related word (*paprika*) that was congruous with the meaning of the sentence but was not the most highly predictable completion of the binomial string. These manipulations resulted in a 2 (Conjunction Type: *and* vs. *and also*) x 2 (Predictability of the string's final word: *canonical* vs. *control*/*novel*) factorial design with four experimental conditions (see Example 1 below). Based on the terminology used in past research on the processing of variant multi-word chunks (e.g., Molinaro & Carreiras, 2010), we chose the term "predictability" to refer to the contextual predictability of the final word in the multi-word chunk, which distinguishes between a highly expected, highly semantically-related conjunct (*pepper*) that is part of a familiar and relatively frequent multi-word chunk, and an unexpected, less semantically-related conjunct (*paprika*) that constitutes a novel, unfamiliar sequence. In addition to our main conditions of interest, i.e., the variant binomial (1c) and its corresponding control condition (1d), we also included the canonical binomial form (1a) and its corresponding control condition (1b) for completeness (see Example 1 below).

1a. The meat was bland, so the chef reached for the salt and pepper to add some flavor. (And–Canonical)

1b. The meat was bland, so the chef reached for the salt and paprika to add some flavor. (And–Control)

1c. The meat was bland, so the chef reached for the salt and also pepper to add some flavor. (Also–Canonical)

1d. The meat was bland, so the chef reached for the salt and also paprika to add some flavor. (Also–Control)

If the online processing advantage for binomials is disrupted in binomials with a variant surface form, the processing speed advantage should be absent in the Also conditions (1c and 1d): The canonical conjuncts (*pepper* in 1c) should be read equally slow to, or even slower than, the novel conjuncts (*paprika* in 1d). Conversely, if the online processing advantage for binomials is robust to surface form modifications, as suggested by the findings from Experiment 1, we should observe a robust processing advantage in the variant binomials. In the Also conditions, canonical conjuncts preceded by a variant conjunction (*pepper* in 1c) should be read faster than novel conjuncts in a control phrase (*paprika* in 1d).

3.1 Participants

Fifty-two undergraduates recruited from the same pool and with the same restrictions as Experiment 1 participated in this experiment for course credit. Participants reported normal or corrected-to-normal vision, and gave informed consent to take part in the study. None had participated in Experiment 1.

3.2 Materials

The experimental stimuli were the 36 sentences from Experiment 1. The full list of stimuli is included as supplementary materials (Appendix S1). The properties of the binomial items are summarized in Table 1. Each experimental item contained a binomial phrase (e.g., *salt and pepper*) that varied in conjunction type (*and* vs. *and also*) and in the predictability of the final noun (*canonical* vs. *control/novel*). These two manipulations are described below.

Table 1

Item characteristics for the 36 binomials

	And–Canonical *salt and pepper*	And also–Canonical *salt and also pepper*	As well as–Canonical *salt as well as pepper*	And–Control *salt and paprika*	And also–Control *salt and also paprika*	As well as–Control *salt as well as paprika*
Phrase length (characters)	15 (2.5)	22.06 (2.72)	24.06 (2.72)	15.31 (1.85)	22.25 (2.53)	24.25 (2.53)
Phrasal frequency (COCA corpus)	1.38 (2.23)	0	0	0.01 (0.03)	0	0
Word frequency of the second conjunct	61.90 (88.24)	-	-	31.15 (37.04)	-	-
Lexical semantic association (LSA)	0.53 (0.22)	-	-	0.27 (0.18)	-	-
Cloze score of the second conjunct (from Experiment 1)	0.86 (0.35)	0.80 (0.40)	0.76 (0.43)	0.01 (0.04)	0.01 (0.04)	0.01 (0.02)
Typicality ratings (1-7)	5.43 (0.71)	3.51 (0.37)*	3.74 (0.56)	5.10 (0.74)	3.39 (0.41)*	3.86 (0.51)

Note. Standard deviations are presented in parentheses. All calculations included the determiner preceding the second conjunct, if present for that item. We report frequencies as the occurrence per million words rather than the log frequency because the frequency for some of the unattested binomials were zero (log of zero is undefined). Because the second conjunct in the Also and As Well As conditions were the same as in the And condition, a dash is shown in place of redundant values. For cloze scores and typicality ratings, binomials were presented in their sentence context. *These mean ratings were from Experiment 2a.

3.2.1 Manipulation of conjunction type

As in Experiment 1, we created variant binomials by inserting the particle *also* before the second conjunct (e.g., *salt and also pepper*). As a result of our conjunction manipulation, the conjunction region in the Also condition was always longer than the conjunction region in the And condition (which only contained the word *and*). Since past research has shown that the probability of skipping over a parafoveal word is affected by the word's distance from the previous fixation or launch site (Brysbaert, Drieghe, & Vitu, 2012; Drieghe, Brysbaert, Desmet,

& De Baecke, 2004), we did not select word skipping as one of our a priori dependent measures in this experiment as the skipping measure may be affected by the differences in numbers of words in our stimuli.

For some items in the Also condition (19 items), we also had to add a determiner (*the*, *a*, *his*, *her*) before the target word so that the sentences were grammatically correct, e.g., *the bride and also the groom*. For these items, we included the determiner as a part of the conjunction region (e.g., *and also the*) in our analyses of reading times. Inevitably, this meant that some of the experimental items (19/36) contained a determiner before the target word whereas other items (17/36) did not. We return to this point in the Results section of this experiment, where we conducted additional analyses to test whether the presence or absence of the determiner influenced the results.

3.2.2 Manipulation of target word predictability

To create the control phrases that served as the baseline for quantifying the processing advantage of variant binomials over novel, unfamiliar strings, we replaced the final word in the binomial (*pepper*) with a semantically related word (*paprika*) that was still plausible given the context (see typicality norms in the next section). Analyses from the cloze study (Experiment 1) showed that the words that we selected as the novel conjuncts had mean cloze probability that was near-zero in both the And condition (M = 1%, SD = 9%) and the Also condition (M = 1%, SD = 9%), confirming that the novel conjuncts were not expected completions of the binomial string.

We also controlled for the length of the canonical conjuncts and novel conjuncts. For the final 36 stimuli used in this experiment, the mean length of the canonical conjuncts (M = 5.33, SD = 1.24) was not significantly different from the mean length of the novel conjuncts (M = 5.42, SD = 1.27), p = .780. However, word frequency was not equivalent across the two conditions, according to the COCA corpus (Davies, 2008; see Table 1). There was a significant difference in mean word frequency (per million) of the canonical conjuncts (M = 61.90, SD = 88.24) and the novel conjuncts (M = 31.15, SD = 37.04) across the 36 target word pairs (p < .05). We return to this point in the Results section of this experiment, where we conducted additional analyses to control for frequency differences between the binomial and the novel conjuncts.

We also measured the semantic association between the two conjuncts in each binomial item based on their contextual co-occurrence by obtaining their LSA values (Landauer & Dumais, 1997). Across the items, the mean LSA value was significantly higher between the binomial's first noun and its canonical conjunct (*salt–pepper;* $M = 0.53$, $SD = 0.22$), compared to the LSA value between the first noun and a novel conjunct (*salt–paprika;* $M = 0.27$, $SD = 0.18$), $t(35) = 6.12$, $p < .001$). Because these results showed that on average, the first noun had higher semantic association with the canonical conjunct than the novel conjunct, we included these LSA values as a predictor in the post-hoc analyses of reading times on the second conjunct.

3.2.3 Typicality and plausibility norming

Since the variant binomials (*salt and also pepper*) and the control phrases *(salt and paprika*, *salt and also paprika*) were unattested in corpora (their phrase frequencies were near-zero; see Table 1), we collected typicality ratings for the experimental stimuli to ensure that these phrases were not highly atypical or unusual. Sixty-four UC Davis undergraduates participated in this norming study. Participants were asked to rate the sentences on how typical they seemed, from 1 (*Very Atypical*) to 7 (*Very Typical*).[2] The mean ratings in each condition were as follows: 5.43 (And–Binomial), 5.10 (And–Control), 3.51 (Also–Binomial), and 3.39 (Also–Control). The mean rating for the Also condition was significantly lower than for the And condition ($F(1, 140) = 347.64$, $p < .001$), indicating that participants found the conjunction *and also* more unusual than *and*, regardless of predictability of the phrase-final word. There was also an effect of Predictability ($F(1, 140) = 5.20$, $p < .05$), indicating that the canonical conjuncts were perceived to be slightly more typical than the novel conjuncts in the sentences.

3.3 Procedure

The 36 experimental items were presented with 72 filler sentences, so that each participant read a total of 108 sentences. These items were counterbalanced across four lists in a Latin Square design. At the beginning of the experiment, each participant was randomly assigned to one of these four lists. Simple comprehension questions were presented on 27 of the filler sentences (25% of all trials) to ensure that participants were paying attention while reading.

Eye movements were recorded with an EyeLink 1000 System (SR Research Ltd.) with a sampling rate of 1,000 Hz. Head movements were minimized using a chin and forehead rest. Viewing was binocular but eye movements were recorded only for the right eye, following conventional practice in eye tracking studies. The sentences were displayed on a 1,024 x 768 display monitor. Eye tracking was calibrated at the beginning of each session and throughout the session if necessary. After calibration, participants read three filler sentences in a practice block. Afterwards, the 108 sentences were presented in random order.

Participants were instructed to read each sentence silently at a normal pace. Each trial began with a fixation dot on the left side of the monitor. Once gaze was stable, the experimenter presented the sentence. After reading the sentence, participants pressed a button on a handheld button box, which replaced the sentence with the fixation dot for the next trial. On 27 of the trials, the sentence (always a filler) was replaced with a comprehension question, with two possible answers directly below it. Participants responded to the question using the button box. Each experimental session lasted about 30 minutes.

3.4 Analysis

3.4.1 Analysis regions and data preparation

Each experimental sentence was divided into interest areas, and reading times were computed on the target word region (i.e., the second conjunct in the binomial). Analyses were conducted on three reading time measures that were determined a priori: first fixation duration, first pass reading, and second pass reading. These measures were selected because they have been shown in previous studies to be sensitive to our types of manipulations (Frisson, Rayner, & Pickering, 2005; Kliegl, Grabner, Rolfs, & Engbert, 2004; Rayner, Slattery, Drieghe, & Liversedge, 2011), and because we wished to avoid measures that might be problematic given the differences in numbers of words across the And versus the Also conditions of our stimuli (e.g., word skipping). First fixation duration is the duration of the first fixation made in the region. First pass reading (also known as gaze duration) captures the duration of all fixations made within the region during first-time reading, before exiting the region to the left or right. Second pass reading is the duration of all fixations made in the region during second-time reading, before leaving the region to the left or right. We did not have specific predictions about whether the processing advantage would emerge in early processing (first fixation and first pass) or later processing (second pass), although past literature on predictability effects would suggest the early processing measures would likely be more diagnostic. In addition to these three a priori measures, we also report data for four other standard reading measures in our main analyses:

word skipping probability, regression-out probability, regression path duration, and total reading time, although following Von der Malsburg and Angele (2017), our focus is on the a priori measures.

Standard data filters in the Data Viewer software (SR Research Ltd.) were used to automatically combine nearby fixations that were shorter than 80 ms and within 1 character from each other. None of the fixations exceeded 1,500 ms. Using Data Viewer, research assistants naïve to the research question performed data cleaning on the raw data to ensure that congruous fixations were included in the interest area of the sentence, and that suspected blinks and highly incongruous fixations were excluded from the interest area. For trials on which the target word was skipped during first pass reading, any recorded fixation durations on those trials were removed. This allowed us to distinguish trials on which the target word was fixated during early processing (i.e., first fixation duration was recorded) from trials on which the word was not fixated (or skipped) during early processing. We also deleted values of zero for total reading time on trials where the target word was skipped, because those zero values misrepresented the data.

3.4.2 Statistical analyses

After cleaning the data, we analyzed each measure separately by fitting Bayesian mixed-effects regression models using the Stan modeling language (Stan Development Team, 2018) accessed with the *brms* package (Bürkner, 2018) in R (R Core Team, 2021). We used the Bayesian framework for all the analyses because unlike the frequentist approach, the Bayesian models circumvent issues of nonconvergence and also ensure that the design-driven random effects were always included in the models. The results we report are based on whether the 95% credible interval of each parameter's posterior distribution includes zero. If the interval does not include zero, then there is substantial evidence (95% probability) that the coefficient's true value is not zero (Nicenboim & Vasishth, 2016; McElreath, 2020).

The following procedure of model specification was followed for all the reading time analyses (both a priori and post-hoc), in this experiment and in all subsequent experiments. The models included Conjunction Type and Predictability as main effects (both deviation-coded as -0.5 and 0.5), along with their interaction. Following the recommended procedure by Barr and colleagues (2013), for each model we included random intercepts and random slopes by subjects and by items, setting the random correlations between the random slopes and random intercepts to zero, as follows: (1|subject), (0+Conditions|subject), (1|item), (0+Conditions|item), such that Condition denotes the full set of main effects and their interaction. For the binary measures (i.e., word skipping probability and regression-out probability), we fitted Bayesian logistic models with the Bernoulli distribution and logit link function. For all models, we used the default priors specified by *brms*. See OSF for the full set of data, materials, and analysis scripts: https://osf.io/6y8xb/

3.5 Results: a priori analyses

3.5.1 Comprehension question accuracy

Mean accuracy for the comprehension questions was high (94%), confirming that participants were paying attention during the online reading task.

3.5.2 Target word region

Mean reading times on the second conjunct (e.g., *pepper* or *paprika*) for each measure are presented in Table 2 and Fig. 2. There was a main effect of Predictability for first fixation duration (β = 21.16, 95% CI = [12.09, 30.23]) and first pass duration (β = 33.44, 95% CI = [20.63, 45.15]), indicating that the canonical conjuncts (*pepper*) were read faster than novel conjuncts (*paprika)* across all conjunction types. For second pass duration, none of the effects received substantial evidence.

For completeness, we also report the results for four additional eye movement measures. Mixed-effects analyses indicated a main effect of Predictability for regression path duration (β = 32.10, 95% CI = [9.08, 64.00]) and total reading time (β = 44.42, 95% CI = [29.71, 59.39]). There was also a main effect of Conjunction Type for regression-out probability (β = 1.07, 95% CI = [0.45, 1.69]) and regression path duration (β = 27.22, 95% CI = [1.37, 52.25]), indicating that during first-pass reading, readers made more leftward regressions from the second conjunct and spent more time re-reading in the Also condition than in the And condition. For word skipping probability, none of the fixed effects received substantial evidence.

3.5.3 Spillover region

Following conventional practice in eye tracking reading studies, we also report the results for the spillover region (N+2), defined as the two words immediately after the target word region. There was a main effect of Conjunction Type, indicating longer reading times on the spillover word in the Also condition than the And condition in these measures: total reading time

(β = 22.74, 95% CI = [0.13, 45.81]) and second pass duration (β = 33.39, 95% CI = [6.13, 61.20]). For the other eye movement measures, none of the effects received substantial evidence.

3.6 Results: post-hoc analyses

3.6.1 Controlling for the effects of word frequency, LSA, and typicality

Beyond the effects of contextual predictability, the variance in reading times may also be explained by the frequency of the second conjunct, the semantic association between the conjuncts (LSA), and the sentence typicality ratings collected in the norming study. In particular, as previously mentioned in the Materials section, the mean word frequency of the canonical conjuncts was significantly higher than for the novel conjuncts, and thus word frequency may account for the effects of Predictability in the analyses of reading times. To assess the effects of these three variables (word frequency, LSA, and typicality), we included all three as continuous predictors in Bayesian mixed-effects analyses of reading times on the target word that also included the main predictors (Conjunction Type, Predictability, and their interaction). All of the predictors in the model were centered, and Word Frequency was log-transformed to reduce skewness in the frequency distribution. Some of these predictors were correlated with each other (Predictability with Word Frequency and LSA; Conjunction Type with Typicality), but the variance inflation factor and tolerance values showed that the degree of collinearity between the coefficients was low (all VIF values were below 3.60, and all tolerance values were above .28). One possible approach for dealing with multicollinearity is to residualize the predictors, but this method has been shown to be an ineffective remedy for this issue (Wurm & Fisicaro, 2014), and thus we report the results using raw, nonresidualized predictors.

Consistent with the original analyses of the a priori measures, there was a main effect of Predictability for these measures: first fixation duration (β = 17.65, 95% CI = [6.03, 28.90]), first pass duration (β = 28.14, 95% CI = [12.35, 43.52]), and total reading time (β = 33.49, 95% CI = [13.14, 54.61]. For second pass duration, none of the effects received substantial evidence, consistent with the original analyses. Unlike the original analyses, the effect of Conjunction Type in regressions-out and regression path duration no longer received substantial evidence when the covariates were included. These results are reported in Table 3.

3.6.2 Analyses using cloze probability instead of predictability

To further explore the factors that may drive this Predictability effect, we conducted another version of this exploratory analysis that included the cloze probability of the canonical conjunct (collected in Experiment 1) as a predictor instead of the dichotomous Predictability predictor. These analyses revealed a main effect of LSA in various measures: first fixation duration (β = -31.52, 95% CI = [-48.83, -14.62]), first pass duration (β = -49.70, 95% CI = [-73.59, -25.20]), regression-out probability (β = -1.15, 95% CI = [-2.08, -0.22]), regression path duration (β = -102.30, 95% CI = [-153.22, -52.32,]), and total reading time (β = -91.20, 95% CI = [-128.07, -53.83]). Crucially, the effect of word frequency did not receive substantial evidence in any of the analyses, suggesting that the Predictability effect in the original analyses was not reducible to word frequency effects. Additionally, there was a main effect of Typicality for total reading time (β = -16.69, 95% CI = [-33.04, -0.50]), and there was a main effect of Conjunction Type for first pass duration (β = -27.20, 95% CI = [-52.70, -3.12]). None of the other effects received substantial evidence.

3.6.3 Effect of determiner in the Also condition

As previously mentioned in the Materials section, some of the experimental items (19/36) required a determiner (*the, a, his, her*) in the Also condition to make the sentence grammatical (e.g., *bride and also the groom/priest*), whereas some items (17/36) did not (e.g., *salt and also pepper/paprika*). This raises the possibility that the predictability of the target word was systematically different between the items that had a determiner versus those that did not. To address this issue, we conducted exploratory analyses on the five eye movement measures that revealed a main effect of Predictability in the original analyses on the target word region: first fixation duration, first pass duration, regression path duration, total reading time, and second pass duration. First, we excluded items in the And condition because the determiner never appeared in the And condition (e.g., *salt and pepper/paprika; bride and groom/priest*). Thus, analyzing only the items in the Also condition, we created a categorical predictor called Determiner, deviation-coding the presence or absence of a determiner for each item (no determiner = -0.5, with determiner = 0.5). We entered Predictability, Determiner, and their interaction as fixed effects in each model.

Results from these models revealed a main effect of Predictability for first fixation duration (β = 19.95, 95% CI = [7.66, 32.17]), first pass duration (β = 33.33, 95% CI = [15.94, 50.43]), and total reading time (β = 47.47, 95% CI = [24.41, 70.06]). These results indicated that the effect of Predictability in the Also condition was robust despite the presence of a determiner before the target word.

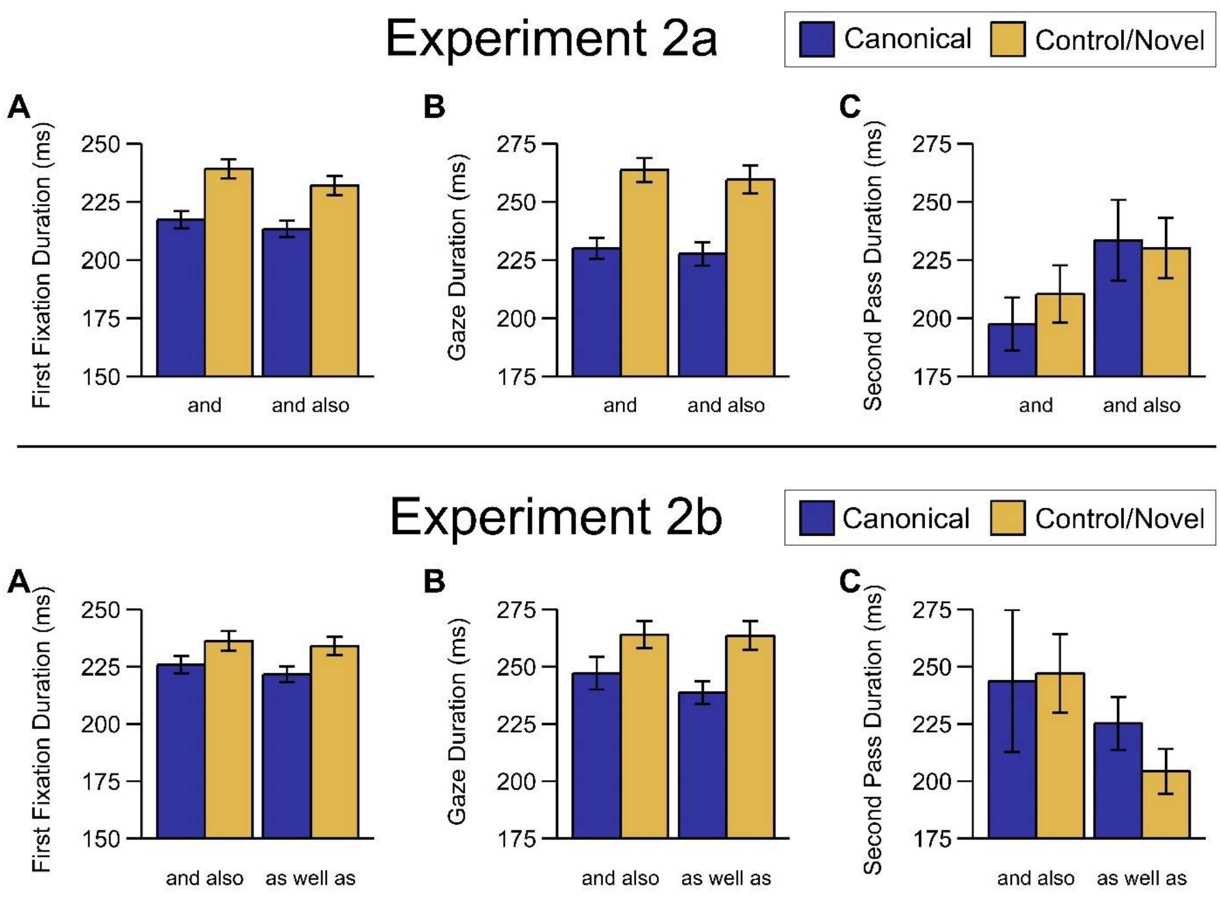

Figure 2. Mean reading times in milliseconds on the second conjunct (canonical or control/novel) for the three a priori measures in the eye tracking studies (Experiment 2a in top panel, Experiment 2b in bottom panel): (A) first fixation duration, (B) first pass (gaze) duration, and (C) second pass duration. Error bars denote the standard error of the mean.

Table 2

Mean reading times in milliseconds (standard deviations in parentheses) on the second conjunct for all eye movement measures analyzed in Experiment 2a and 2b

Measure	Experiment 2a				Experiment 2b			
	And–Binomial	And–Control	And also–Binomial	And also–Control	And also–Binomial	And also–Control	As well as–Binomial	As well as–Control
First fixation duration	217 (69)	239 (79)	213 (65)	232 (76)	226 (69)	236 (78)	222 (61)	234 (76)
First pass (gaze) duration	230 (83)	264 (98)	228 (91)	260 (110)	247 (128)	264 (110)	239 (89)	264 (115)
Second pass duration	198 (66)	311 (90)	233 (112)	230 (100)	365 (464)	362 (346)	343 (292)	336 (273)
Regression-out probability	.06 (.23)	.07 (.26)	.12 (.33)	.16 (.37)	.11 (.31)	.13 (.33)	.14 (.34)	.09 (.29)
Regression path duration	247 (146)	291 (153)	285 (270)	313 (188)	278 (160)	309 (174)	288 (184)	300 (211)
Total time	252 (110)	297 (130)	265 (151)	312 (177)	287 (191)	300 (216)	282 (137)	307 (156)
Skipping probability	.26 (.44)	.26 (.44)	.28 (.45)	.28 (.45)	.32 (.47)	.28 (.45)	.29 (.45)	.28 (.45)

Note. Target word = the second conjunct in the binomial, e.g., *pepper* (canonical conjunct) or *paprika* (novel/control conjunct). Means for word skipping probability and regression-out probability are in percentages.

Table 3

Results of Bayesian mixed-effects analyses for the three a priori reading measures on the second conjunct in Experiment 2a and 2b

	Experiment 2a			Experiment 2b		
Model parameters for each measure	β	*SE*	95% CI	β	*SE*	95% CI
First fixation duration						
Intercept	224.58	3.80	[217.09, 231.83]	228.05	3.75	[220.74, 235.29]
Word Frequency	-0.95	1.36	[-3.64, 1.71]	-1.54	1.32	[-4.20, 1.06]
LSA	-9.79	11.15	[-31.84, 11.66]	**-30.82**	**12.04**	**[-55.18, -8.06]**
Typicality	-0.49	3.96	[-8.47, 7.27]	-4.51	4.81	[-13.93, 4.60]
Conjunction Type	-6.40	8.84	[-24.04, 10.48]	-2.06	4.65	[-11.07, 6.88]
Predictability	**17.65**	**5.75**	**[6.03, 28.90]**	1.91	5.87	[-9.84, 13.25]
Conjunction Type*Predictability	-2.49	8.50	[-19.26, 13.66]	3.38	8.91	[-14.18, 21.05]
First pass duration						
Intercept	242.73	6.09	[230.77, 254.91]	250.41	6.02	[238.32, 262.22]
Word Frequency	-0.75	2.19	[-5.12, 3.40]	0.00	2.27	[-4.57, 4.34]
LSA	-13.12	17.47	[-47.68, 19.66]	-30.40	20.22	[-71.63, 8.64]
Typicality	-4.87	5.94	[-16.92, 6.49]	0.45	7.37	[-13.95, 14.59]
Conjunction Type	-13.92	12.89	[-39.85, 11.73]	-4.42	6.76	[-17.46, 8.97]
Predictability	**28.14**	**8.02**	**[12.35, 43.52]**	12.37	9.59	[-6.04, 31.55]
Conjunction Type*Predictability	-0.06	10.26	[-20.06, 19.94]	9.76	13.89	[-17.39, 37.48]
Second pass duration						
Intercept	216.46	9.94	[196.97, 236.14]	219.55	12.88	[193.41, 244.39]
Word Frequency	1.96	4.69	[-7.31, 11.02]	3.45	5.53	[-7.52, 14.35]
LSA	-21.08	38.74	[-97.42, 56.66]	-89.97	49.49	[-188.54, 8.47]
Typicality	-9.85	14.07	[-37.63, 17.00]	-33.38	22.47	[-77.21, 11.29]
Conjunction Type	2.97	29.89	[-55.80, 62.04]	-23.18	20.36	[-63.00, 16.50]
Predictability	-2.05	19.08	[-40.13, 35.24]	-28.25	24.14	[-75.07, 20.63]
Conjunction Type*Predictability	-3.09	31.90	[-67.27, 56.52]	-16.29	35.94	[-87.80, 54.10]

Note. SE = standard error; CI = credible interval. Main effects that received substantial evidence are bolded.

3.7 Discussion

Experiment 2a showed that across several eye movement measures, canonical conjuncts (e.g., *pepper*) were read faster than the novel conjuncts (e.g., *paprika*), even in the variant binomials (*salt and also pepper*) whose surface form did not match the canonical binomial template (*salt and pepper*). Post-hoc analyses showed that the predictability effect (for first fixation duration, first pass, and total reading time) received substantial evidence even after accounting for word frequency effects, semantic association effects, and sentence typicality (see Table 3). This predictability effect was also robust to the presence of a determiner (*the, a, his, her*) in some of the items. The robust predictability effect in the early measure of processing (first fixation duration and first pass) supports the account that the processing speed advantage for binomials is robust to conjunction substitutions in the phrasal template, consistent with the second set of predictions outlined in the Introduction. These results from the eye tracking study mirror the findings from the cloze study (Experiment 1) that revealed high rates of producing the canonical conjunct even for the nonformulaic conjunctions. In the post-hoc analysis that used the cloze probability of the canonical conjunct as a predictor instead of the dichotomous Predictability variable (Section 3.6.2), we also found an effect of LSA across various eye movement measures: first fixation duration, first pass duration, regression path duration, and total reading time. This effect of LSA is consistent with previous work on binomial processing, which showed that the predictability of binomials is driven by the lexical-semantic association between the conjuncts (Carrol & Conklin, 2019; for further discussion see Experiment 2b and the General Discussion).

Interestingly, we also found an effect of conjunction type in measures of later processing, which we did not expect. Readers were more likely to regress from the second conjunct (target word) and had longer regression path durations in the Also condition than the And condition. This effect of conjunction type was also evident in the spillover region for second pass duration and total reading time. The direction of this conjunction effect in the spillover region was consistent with conjunction effect in the target word region: Reading times on the spillover region were longer in the Also condition compared to the And condition. What could explain the effect of words following *and also* being re-read and re-analyzed more than words following *and*, regardless of the predictability of the target word? This effect of conjunction type seems to pattern with the results from the sentence typicality norms that we conducted on the stimuli, and so here we see concordance between the offline and online data. Sentences containing *and also* were rated as significantly less typical than sentences containing just *and*, regardless of target word predictability, which indicated that readers found the conjunction *and also* somewhat unusual in conjoined noun phrases.

To follow up on these results which showed the unusualness of *also,* we searched for the pattern of usage for the word *also* and for the bigram *and also* in the COCA corpus (Davies, 2008). The corpus showed that *also* is not commonly used in conjoined noun phrases. The particle *also* is more frequently used at the beginning of a sentence (e.g., *Also, he went to the store*) than in the middle of one, functioning as a discourse marker. These results are consistent with the literature on the usage of *also*, which showed that *also* is more commonly used in other syntactic positions in the sentence (Sopher, 1978) rather than in conjoined noun phrases. The unusualness of *also* in conjoined phrases may therefore explain readers' greater likelihood of regressing from the target word following *and also*, and readers' longer re-reading of the words following *and also*, including the spillover region after those words. The effect of conjunction type was observed in reading measures thought to reflect later stages of processing (regressions-out, regression path duration, total reading time, and second pass duration). In contrast, the earlier measures of lexical processing (first fixation duration and first pass) on the target word indicated a robust predictability effect (i.e., the contextual expectancy of canonical conjuncts over novel conjuncts), which is consistent with past literature showing that prediction is reflected in very early measures of lexical processing (Frisson et al., 2005; Rayner et al., 2011; Staub & Goddard, 2018).

Given our main finding from Experiment 1 that the tendency to produce the canonical conjunct was high even in the noncanonical conjunctions (Experiment 1), and the finding from this experiment that the online lexical processing of the canonical conjunct was still facilitated even for binomials with the variant conjunction *and also* (Experiment 2a), we conducted a second eye tracking study to assess the processing advantage of two types of variant binomials (i.e., binomials substituted with *and also* and with *as well as*) over matched novel phrases.

4. Experiment 2b: Eye tracking study (comparing *and also* to *as well as*)

In Experiment 2b we compared reading times on the second conjunct in variant binomials conjoined with *and also* vs. *as well as*. We chose to compare *as well as* with *and also* (excluding the canonical form with *and*) because this experimental design eliminated possible effects of the Determiner (previous described under post-hoc analyses in Experiment 2a), since the presence or absence of a determiner was consistent across the Also condition and the As Well As condition. This experiment had a 2 (Conjunction Type: *and also* vs. *as well as*) x 2 (Predictability of the second conjunct: *binomial* vs. *control/novel*) factorial design with four experimental conditions (see Example 2):

2a. The meat was bland, so the chef reached for the salt and also pepper to add some flavor. (Also–Binomial)

2b. The meat was bland, so the chef reached for the salt and also paprika to add some flavor. (Also–Control)

2c. The meat was bland, so the chef reached for the salt as well as pepper to add some flavor. (As well as–Binomial)

2d. The meat was bland, so the chef reached for the salt as well as paprika to add some flavor. (As well as–Control)

Under the prediction that the online processing advantage for binomials would be disrupted in binomials with an altered surface form, the canonical conjuncts should be read faster than novel conjuncts in the Also condition, replicating the results from Experiment 2a, whereas this predictability effect should be absent in the As Well As condition (e.g., perhaps due to the greater dissimilarity in orthographic form and string length between *as well as* and the formulaic conjunction *and*, relative to the dissimilarity between *and also* and *and*). On the other hand, if the processing advantage of binomial phrases is robust to surface form modifications, there should be a robust effect of predictability in these two variant binomial forms: canonical conjuncts would be read faster than novel conjuncts in both the Also condition and the As Well As condition.

4.1 Participants

Fifty-two undergraduates from the same pool and selected with the same restrictions participated in this experiment. None had participated in any of the other experiments reported here.

4.2 Materials

The experimental stimuli were the same as in Experiment 2a, except for the manipulation of conjunction type (see Table 1 for a summary of the binomials' properties). Note that for the 19 items that required a determiner before the target word, the determiner was present in both the Also condition (*bride and also the groom/priest)* and the As Well As condition (*bride as well as the groom/priest)*. Thus, unlike in Experiment 2a, in Experiment 2b the presence or absence of a determiner in each item was consistent across all four experimental conditions.

4.2.1 Manipulation of target word predictability

We used the same binomial and novel conjuncts as Experiment 2a as our target words.

4.2.2 Manipulation of conjunction type

The items in the Also conditions in this experiment were identical to the items in the Also conditions in Experiment 2a. For the As Well As conditions, we replaced the conjunction *and* with *as well as* (e.g., *salt as well as pepper*).

4.2.3 Typicality and plausibility norming

As in Experiment 2a, we collected typicality ratings for the experimental stimuli to validate that the conjunctions we selected (*and also* and *as well as*) were not highly unusual. Fifty-two UC Davis undergraduates participated in this norming study. All 36 items were counterbalanced across four lists with the same filler sentences used in the typicality norming in Experiment 2a. The mean ratings in each condition were as follows: 3.61 (Also–Binomial), 3.57 (Also–Control), 3.74 (AWA–Binomial), and 3.86 (AWA–Control). The small mean difference between the Conjunction conditions was significant, $F(1, 140) = 6.51$, $p < .05$, indicating that participants found the conjunction *and also* slightly more unusual than *as well as*, regardless of target word predictability. The mean difference between the Predictability conditions was not significant ($p = .527$), indicating that the binomial and novel conjuncts were equally typical in the sentences. This finding is in contrast to Experiment 2a, in which we found that canonical conjuncts were significantly more typical than the novel conjuncts. It is unclear why there is a discrepancy given that the same target words and sentence frames were used in both experiments, but we note that the differences between the mean typicality ratings for the binomial and novel target words in Experiment 2a were modest (4.47 and 4.25, respectively), considering the full range of the scale (1 to 7).[3]

4.3 Procedure

The experimental procedure was identical to the procedure described in Experiment 2a. The 36 experimental sentences were presented with the same 72 filler sentences used in Experiment 2a, and each participant read a total of 108 sentences. These items were counterbalanced across four lists, and at the beginning of the experiment each participant was randomly assigned to one of these four lists. As in Experiment 2a, there were simple comprehension questions for 27 of the same filler sentences (25% of all trials).

4.4 Results: a priori analyses

The analysis regions, dependent measures, data cleaning procedures, and statistical analyses for this experiment were identical to Experiment 2a (described under Section 3.4). The data, materials, and analysis scripts for this experiment are available on OSF: https://osf.io/6y8xb/

4.4.1 Comprehension question accuracy

Mean accuracy for the comprehension questions was high (96%), confirming that participants were paying attention during the reading task.

4.4.2 Target word region

Mean reading times on the target word (e.g., *pepper* or *paprika*) for each measure are presented in Table 2 and Fig. 2. There was a main effect of Predictability for first fixation duration (β = 11.28, 95% CI = [0.94, 21.73]) and first pass reading (β = 20.60, 95% CI = [4.79, 36.55]), indicating that for both types of variant binomials (*and also*, *as well as*), the canonical conjuncts were read faster than novel conjuncts. For second pass duration, none of the effects received substantial evidence.

For completeness, we also report the results for five additional eye movement measures. Mixed-effects analyses indicated a main effect of Predictability for total reading time (β = 33.38, 95% CI = [6.28, 60.42]). For the remaining eye movement measures, none of the effects received substantial evidence.

4.4.3 Spillover region

For all measures, none of the effects received substantial evidence.

4.5 Results: post-hoc analyses

4.5.1 Controlling for the effects of word frequency, LSA, and typicality

As in Experiment 2a, we conducted additional analyses that included word frequency, LSA, and typicality ratings as continuous predictors in the analyses of reading times. All of the predictors in the model were centered, and Word Frequency was log-transformed. The final model for each reading time measure included Word Frequency, LSA, Typicality Ratings, Conjunction Type, and Predictability (raw, nonresidualized predictors) as fixed effects. The degree of collinearity between these coefficients was low (all VIF values were below 1.80, and all tolerance values were above .56).

For first fixation duration and first pass duration, the effect of Predictability did not receive substantial evidence, unlike the a priori analyses. Instead, there was a main effect of LSA for first fixation duration (β = -30.82, 95% CI = [-55.18, -8.06]) and total reading time (β = -63.40, 95% CI = [-128.21, -1.46]) . For the other measures, none of the effects received substantial evidence, consistent with the a priori analyses. These results are summarized in Table 3.

4.5.2 Analyses using cloze probability instead of predictability

As with Experiment 2a, we conducted exploratory analyses using the cloze probability of the canonical conjunct as a predictor instead of the dichotomous Predictability variable. Results indicated a main effect of LSA for first fixation duration (β = -36.89, 95% CI = [-54.89, -18.94]), first pass duration (β = -54.67, 95% CI = [-82.96, -26.65]), regression path duration (β = -63.07, 95% CI = [-110.48, -16.60]), and total reading time (β = -96.48, 95% CI = [-141.51, -53.35]). None of the other effects received substantial evidence in any of the measures.

4.6 Discussion

Consistent with the results from Experiment 2a, results from Experiment 2b indicated robust effects of Predictability, indicating that canonical conjuncts (e.g., *pepper*) were read faster than novel conjuncts (e.g., *paprika*), across both types of variant conjunctions. In the post-hoc analyses that included word frequency, typicality, and LSA (Section 4.5.1), there was an effect of LSA for first fixation duration and total reading time. This effect of LSA was also found in the post-hoc analyses that used the cloze probability of the canonical conjunct as a predictor instead of Predictability (Section 4.5.2), across numerous eye movement measures: first fixation duration, first pass duration, regression path duration, and total reading time.

It is worth noting that, in the exploratory analyses that included both LSA and Predictability as predictors (Section 4.5.1 and Table 3), the analysis of first fixation duration indicated an effect of LSA but no effect of Predictability. This does not necessarily mean that LSA is a confound in the Predictability manipulation. Rather, it suggests that LSA is one possible factor contributing to the predictability of the binomials. Since Predictability and LSA are mildly correlated with each other, the degree of evidence for each effect may strongly depend on whether these two variables are included in the same model. When both predictors are included in the model (see Table 3), sometimes the effect of Predictability receives substantial evidence only (Experiment 2a), and other times the effect of LSA receives substantial evidence only (Experiment 2b). This discrepancy may reflect the fact that neither Predictability nor LSA are given credit for their shared variance, because this variance is not uniquely assigned to either predictor (Wurm & Fisicaro, 2014). On the other hand, when only one of these predictors is included, such as in the exploratory analyses that included LSA but excluded Predictability (see Section 4.5.2 and Section 3.6.2), LSA explains the variance in the same eye movement measures in both Experiment 2a and 2b. As discussed in Experiment 2a, we interpret this robust LSA effect as indicating that semantic association partly explains the Predictability effect that was observed in the a priori analysis. This conclusion is consistent with previous research which showed that semantic association underlies the online processing advantage of canonical binomial phrases conjoined with *and* (Carrol & Conklin, 2019). We elaborate on this point in the General Discussion.

A potential concern regarding Experiment 2a and 2b is that the eye tracking task may not fully capture the influence of the variant conjunction on the processing advantage of binomials, because reading processes may unfold too quickly to allow people enough time to evaluate the information conveyed by the conjunctions (and the determiners). To ensure that comprehenders are using the conjunction information in mapping the multi-word chunks onto the lexicon, we conducted a third reading experiment using a different reading paradigm and included all three conjunction types in the same reading study.

5. Experiment 3: Self-paced reading study

In Experiment 3, we aimed to replicate the results from Experiment 2 by comparing all three conjunctions in the same reading experiment using a self-paced reading task. We chose self-paced reading rather than eye tracking for the following reasons. The self-paced reading paradigm constrains readers to fixate on each word in the sentence, including the function words that are often skipped during unconstrained reading. More importantly, the self-paced reading paradigm allows us to circumvent issues surrounding parafoveal preview and the effects that preview of the second conjunct might have had on both processing times for the conjunction and the second conjunct, because only one word is visible at a time in self-paced reading. The other issue we wished to address is the potential confound associated with conjunction length differences and their potential effects on the skipping rate and fixation durations on the target word. The self-paced reading task addresses these concerns because in self-paced reading, subjects are required to read each word in the sentence one at a time with no possibility of regressing to earlier parts of the sentence. Finally, the simplicity of the task allowed us to test the large number of participants required to ensure sufficient power to evaluate our hypotheses given this expanded experimental design.

This study had a 3 (Conjunction Type: *and* vs. *and also* vs. *as well as*) x 2 (Predictability of target word: *canonical* vs. *control/novel*) factorial design with six conditions (see Example 3). The stimuli were identical to the ones used in Experiment 1 and 2.

3a. The meat was bland, so the chef reached for the salt and pepper to add some flavor. (And–Canonical)

3b. The meat was bland, so the chef reached for the salt and paprika to add some flavor. (And–Control)

3c. The meat was bland, so the chef reached for the salt as well as pepper to add some flavor. (Also–Canonical)

3d. The meat was bland, so the chef reached for the salt as well as paprika to add some flavor. (Also–Control)

3e. The meat was bland, so the chef reached for the salt as well as pepper to add some flavor. (As well as– Canonical)

3f. The meat was bland, so the chef reached for the salt as well as paprika to add some flavor. (As well as–Control)

If the online processing advantage for binomials is disrupted in the presence of nonformulaic conjunctions (*and also, as well as*), the processing advantage for canonical conjuncts over novel conjuncts should be absent in the two variant conjunction conditions. In contrast, if the online processing advantage for binomials is robust despite the presence of a nonformulaic conjunction, canonical conjuncts should be read faster than novel conjuncts in all three Conjunction conditions. This pattern of findings would replicate the main results from Experiment 2a and 2b.

5.1 Participants

One hundred eighty-two undergraduates from the same subject pool and selected according to the same criteria participated in this experiment. None had participated in the other experiments reported above. Data from eight participants were excluded as they did not meet the a priori criterion we had established for data inclusion (see Results section), and thus the analyses included data from 174 participants.

5.2 Materials

The 36 stimuli for this experiment consisted of the same experimental items used in Experiment 2a and 2b. Each sentence contained a binomial (*salt and pepper*) that varied in conjunction type (*and* vs. *and also* vs. *as well as*) and target word predictability (*canonical* vs. *control/novel*). The 36 experimental sentences were presented with the same 72 filler sentences used in Experiment 2, and each participant read a total of 108 sentences. These items were counterbalanced across six lists, and at the beginning of the experiment each participant was randomly assigned to one of these six lists.

5.3 Procedure

The sentences were presented in a standard self-paced, noncumulative, moving window paradigm (Just, Carpenter, & Woolley, 1982) on a 1,024 x 768 monitor, using the Ibex Farm platform (http://spellout.net/ibexfarm/). Participants were tested on individual computers in a small room which accommodated up to 8 participants. An experimenter remained in the room with the participants for the duration of the experiment. Before the experiment, participants read three filler sentences (identical to the ones used in Experiment 2) in a practice block to become familiarized with the self-paced reading paradigm. Afterwards, the 108 sentences were presented in random order.

The first screen instructed participants to read each sentence silently at a normal pace. Each trial displayed a sentence outline, with underlines in place of each letter in the sentence. Spaces and punctuation marks appeared unchanged throughout the trial. Participants read the sentences one word at a time by pressing the spacebar. Each spacebar press revealed a word while the previous word changed back into its underlined form. On 25% of the trials, the sentence was replaced with a comprehension question centered on the screen with two possible answers directly below, each labeled with (1) or (2). These questions were identical to and followed the same items as in Experiment 2a and 2b. Participants responded to the question by either pressing the corresponding number on the keyboard or clicking on the answer directly. Between two trials, participants saw a transition screen instructing them to press the spacebar to see the next sentence. Reading times were recorded as the time between button presses. Each experimental session lasted about 30 minutes.

5.4 Analysis

Following conventional practice for self-paced reading data, we treated outliers by excluding data on trials whose reading times were 3 standard deviations above or below the mean for each subject (e.g., Brothers, Swaab, & Traxler, 2017). This trim affected approximately 1.5% of the reading time data for the experimental items, and the percentage of trials removed was similar across conditions (1.4% - 1.7%). Reading times on the second conjunct (i.e., target word) were then analyzed using Bayesian mixed-effects models, with subject- and item-variability as crossed random effects, and with Conjunction Type, Predictability, and their interaction as fixed effects. Conjunction Type was coded using dummy contrasts, with the contrast matrix specified as follows: the first contrast compared Also to the And (contrast coefficients: And = 0, Also = 1, As Well As = 0), and the second contrast compared As Well As to And (contrast coefficients: And = 0, Also = 0, As Well As = 1). The two levels of Predictability were deviation coded as -0.5 and 0.5. The procedure for these statistical analyses was the same as Experiment 2 (described in Section 3.4.2 under Experiment 2a). See OSF for the data, materials, and analysis scripts: https://osf.io/6y8xb/

5.5 Results: a priori analyses

Seven subjects performed below 85% on the comprehension questions, and based on this a priori criterion for inclusion, we removed those subjects from the data analyses. To ensure that the data were fully counterbalanced across the six lists, we also excluded data from one subject and analyzed the reading time data for the remaining 174 subjects only.

5.5.1 Comprehension question accuracy

Mean accuracy for the comprehension questions was high (95%) and comparable to what we had observed in Experiment 2a and 2b.

5.5.2 Target word region

Mean reading times on the target word (e.g., *pepper* or *paprika*) are reported in Table 4 and Fig. 3. Mixed-effects analyses revealed that the credible interval for the effect of Predictability included 0, providing insufficient evidence for the effect of Predictability (β = 5.26, 95% CI = [-4.54, 15.21]). The results from these analyses are reported in Table 5.

5.5.3 Spillover region

Since past research has shown that in self-paced reading studies the reading time effects may spill over into the subsequent region (Ferreira & Henderson, 1990; Smith & Levy, 2013), we also report the results for the spillover region, defined as the word immediately after the target word (N+1).[4] The credible interval for the effect of Predictability on the spillover word did not include 0, indicating substantial evidence for the effect of Predictability (β = 12.76, 95% CI = [3.18, 22.53]).

5.6 Results: post-hoc analyses

As in Experiment 2a and 2b, we conducted exploratory analyses that included log word frequency, LSA, and typicality ratings as predictors. The degree of collinearity between these coefficients was low (all VIF values were below 3.16, and all tolerance values were above .32). Analyses on the target word revealed that the credible interval for Predictability included zero (β = 2.45, 95% CI = [-9.03, 13.91]), indicating insufficient evidence for the effect of Predictability. For the spillover word (N+1), however, the credible interval for Predictability did not include 0 (β = 11.34, 95% CI = [0.55, 21.86]), indicating an effect of Predictability. None of the other effects in the model received sufficient evidence. These results are presented in Table 6.

5.7 Discussion

Experiment 3 compared all three conjunction types in a self-paced reading experiment. The a priori analyses showed that canonical conjuncts were read faster than novel conjuncts, even in binomials with variant conjunctions (for both *and also* and *as well as*). The fact that this predictability effect was observed in the spillover region but not in the target word region is unsurprising because the effects in self-paced reading often show up in later regions rather than on the critical region. This processing speed advantage also received substantial evidence in the exploratory analyses that included word frequency, semantic association, and typicality ratings as covariates. Word frequency did not account for the reading times in the exploratory analyses, indicating that the processing speed advantage was not reducible to lexical frequency, consistent with the findings from Experiment 2a and 2b.

Overall, these results from self-paced reading replicated the main findings from the eye tracking experiments (Experiment 2a and 2b), providing clear evidence for facilitated processing of the canonical conjuncts even in variant binomials that contain non-*and* conjunctions. The self-paced reading data also confirmed that the lexical processing advantage in variant binomials that we found in the eye tracking studies were not due to different lengths of the various conjunctions nor their potential effects on parafoveal preview of the second conjunct. Even when participants read the sentences in a self-paced reading paradigm that required them to spend time directly fixating the conjunction, the canonical conjunct was still read faster than novel conjuncts. Given that two eye tracking studies and one self-paced reading study yielded the same pattern of results, we are more confident that variant binomials (e.g., *bride and also the groom*) do confer an online processing speed advantage over matched novel phrases (*bride and also the priest*).

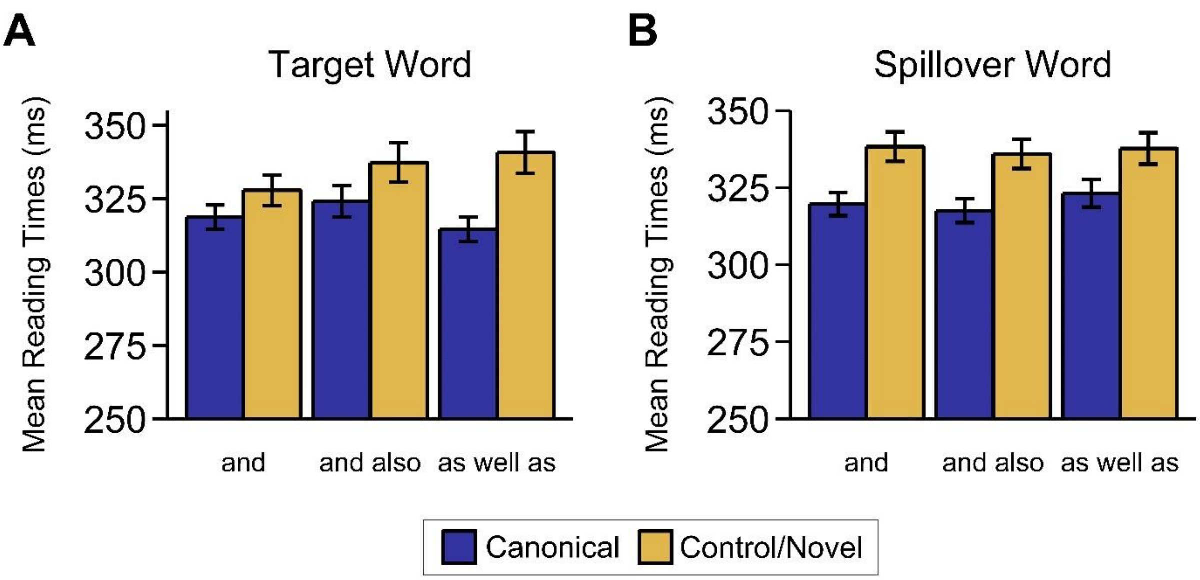

Figure 3. Mean reading times in milliseconds in the self-paced reading study (Experiment 3) on (A) the second conjunct, and (B) the spillover word (N+1). Error bars denote the standard error of the mean.

Table 4

Mean reading times in milliseconds (standard deviations in parentheses) on the second conjunct and the spillover word (N+1) in Experiment 3

Analysis Region	And–Binomial *salt and pepper*	And–Control *salt and paprika*	And also–Binomial *salt and also pepper*	And also–Control *salt and also paprika*	As Well As–Binomial *salt as well as pepper*	As Well As–Control *salt as well as paprika*
Second conjunct	314 (118)	320 (135)	313 (119)	318 (134)	310 (123)	321 (145)
Spillover word	317 (112)	329 (128)	315 (112)	327 (126)	317 (122)	328 (123)

Table 5

Results of the a priori mixed-effects analyses of reading times on the second conjunct and the spillover word (N+1) in Experiment 3

	Second Conjunct		
Model parameters	β	*SE*	95% CI
Intercept	316.81	6.90	[303.22, 330.44]
Conjunction Type (also)	-0.22	4.20	[-8.42, 8.08]
Conjunction Type (as well as)	-0.43	3.87	[-8.00, 7.25]
Predictability	5.26	5.05	[-4.54, 15.21]
Conjunction Type (also)*Predictability	0.85	7.07	[-13.02, 14.67]
Conjunction Type (as well as)* Predictability	7.23	7.90	[-8.57, 22.72]
	Spillover Region (N+1)*		
Model parameters	β	*SE*	95% CI
Intercept	320.82	6.25	[308.79, 333.27]
Conjunction Type (also)	2.01	3.48	[-4.69, 8.90]
Conjunction Type (as well as)	0.95	3.42	[-5.76, 7.71]
Predictability	**12.76**	**4.88**	**[3.18, 22.53]**
Conjunction Type (also)*Predictability	0.79	6.64	[-12.33, 13.96]
Conjunction Type (as well as)* Predictability	-1.55	6.65	[-14.86, 11.31]

Note. *SE* = standard error; CI = credible interval. *N+1 = the 1st word after the second conjunct. Main effects that received substantial evidence are bolded.

Table 6

Results of exploratory mixed-effects analyses of reading times on the second conjunct and the spillover word (N+1) in Experiment 3

	Second Conjunct		
Model parameters	β	*SE*	95% CI
Intercept	315.57	7.65	[300.46, 330.50]
Word Frequency	-1.38	0.97	[-3.33, 0.51]
LSA	-7.15	8.97	[-24.64, 10.47]
Typicality	1.22	2.93	[-4.58, 6.94]
Conjunction Type (also)	1.90	6.75	[-11.49, 15.06]
Conjunction Type (as well as)	1.28	5.79	[-10.27, 12.50]
Predictability	2.45	5.84	[-9.03, 13.91]
Conjunction Type (also)*Predictability	0.55	7.26	[-13.63, 14.88]
Conjunction Type (as well as)* Predictability	6.65	7.97	[-9.08, 22.37]
	Spillover Region (N+1)*		
Model parameters	β	*SE*	95% CI
Intercept	321.24	6.52	[308.37, 333.98]
Word Frequency	-0.51	0.87	[-2.21, 1.19]
LSA	-4.01	7.61	[-18.99, 11.23]
Typicality	0.30	2.63	[-4.90, 5.50]
Conjunction Type (also)	1.44	5.88	[-10.12, 12.87]
Conjunction Type (as well as)	0.80	3.55	[-6.20, 7.68]
Predictability	**11.34**	**5.44**	**[0.55, 21.86]**
Conjunction Type (also)*Predictability	0.73	6.71	[-12.44, 13.90]
Conjunction Type (as well as)* Predictability	-1.74	6.74	[-15.02, 11.51]

Note. SE = standard error; CI = credible interval. *N+1 = the 1st word after the second conjunct. Main effects that received substantial evidence are bolded.

6. General discussion

Prior research has shown that various types of multi-word chunks are processed faster than matched novel strings (Siyanova-Chanturia et al., 2011; Tremblay et al., 2011; Carrol & Conklin, 2019; Arnon & Snider, 2010; Sosa & MacFarlane, 2002; Bannard & Matthews, 2008; Arnon & Cohen Priva, 2013; Janssen & Barber, 2012; Siyanova-Chanturia & Janssen, 2018), but it is unclear whether this processing advantage extends to variant multi-word chunks that are neither highly formulaic nor recurrent. In the current study, we investigated the processing of variant binomial expressions (*salt and also pepper, salt as well as pepper)* in a cloze completion study (Experiment 1), two eye tracking studies (Experiment 2a and 2b), and a self-paced reading study (Experiment 3). Results from Experiment 1 indicated that the cloze probability of the canonical conjunct (e.g., *pepper* in *salt and pepper*) was high even for binomials with variant conjunctions (80% for *and also* and 76% for *as well as*), suggesting that the canonical conjuncts remain robustly predictable despite a variant conjunction. Consistent with these findings, Experiment 2 and Experiment 3 showed that canonical conjuncts were read faster than novel/control conjuncts even in binomials with variant conjunctions, and even when some of the binomials required the insertion of a determiner (*the, a, his, her*) for grammatical correctness (e.g., *bride and also the groom*). Considering all three sets of experiments together, these findings indicate that variant binomials do show a processing advantage over matched control phrases. This was observed even though the variant binomials and control phrases both have low conventionality and phrasal frequency, and even though the conjunction *and* is a defining feature of the canonical configuration of binomials. Overall, these findings support the view that the entrenchment of recurrent multi-word patterns in long-term memory can facilitate language processing in a way that goes beyond the surface statistical regularities in the language (e.g., whole-phrase frequency and lexical frequency), and the findings also suggest that the processing of multi-word chunks may not necessarily require an exact match with the canonical phrasal template.

We now turn to the discussion of the processing advantage of variant binomials in more detail, which we define as the high likelihood of producing the canonical conjunct over other lexical candidates for binomials with noncanonical conjunctions (e.g., *salt and also…*), and the facilitated reading times on the second conjunct in variant binomials (e.g., *salt and also pepper)* than in matched control phrases (*salt and also paprika*), despite the fact that both types of strings are unattested in corpora. This was an intriguing finding because the processing advantage of variant binomials is not attributable to overall lexical frequency (word frequency did not account for the facilitated reading times in any of the exploratory analyses) nor phrasal frequency (for both variants, phrasal frequency is 0; see Table 1), which previous studies have found to be implicated in the processing advantage for canonical binomials (Siyanova et al., 2011; Carrol & Conklin, 2019).

What cognitive mechanisms can explain a formulaic processing advantage independent of phrasal frequency? Based on the idea of analogy in cognitive processing (Bod, 2009; Gentner & Markman, 1997) and the idea of exemplar-based learning in usage-based views of grammar (Abbot-Smith & Tomasello, 2006; Bod, 2006; Goldberg, 2003), one potential explanation of our results is that variant binomials may be processed by direct comparison to the canonical binomial template *(A and B;* see Schmitt, 2006 for a related discussion). During online processing, the language processing system may try to find the closest match to this phrasal template via fuzzy logic or exemplar categorization, and importantly, this match to memory may not have to be exact. An unattested phrase like *salt as well as pepper* may not be precisely the same as the stored exemplar *salt and pepper,* but they are similar enough that readers can recognize the lexical-semantic similarity between the two. Comprehenders may process an unconventional, unfamiliar phrase like *salt as well as pepper* by direct analogy to *salt and pepper,* leading to the speeded activation and retrieval of the canonical conjunct (*pepper*) despite the variant conjunction (for a similar view in the processing of idioms, see McGlone, Glucksberg, & Cacciari, 1994). Under this proposed account, language users can generalize beyond the frequency-based statistical regularities in the input and exhibit a processing advantage even for nonformulaic, unattested binomials. Though this is one of many plausible explanations of our findings, this account precisely explains the cloze data obtained in Experiment 1, which showed that variant binomials are still overwhelmingly predictable despite having a non-formulaic conjunction. This account can also explain the small yet reliable decrease in the cloze probability of the canonical conjunct in the variant conjunctions (<u>*and also*</u>, <u>*as well as*</u>) compared to the canonical conjunction, since the alternative forms are functionally similar but not semantically identical.

The robust processing advantage for variant binomials may be specifically driven by the strong lexical-semantic association between the two conjuncts. Evidence for this comes from the exploratory analyses in Experiment 2a and 2b, which showed that the degree of semantic association predicted reading times on the second conjunct across several eye movement measures. Although the effect of LSA did not receive substantial evidence in some of the exploratory models in Experiment 2a (Section 3.6.1) and Experiment 3 (Section 5.6), this does not undermine these conclusions because LSA may have been accounted for by the predictability of the conjunct, which emerged as the better predictor of reading times in those models. In other words, we do not see our Predictability manipulation and our LSA measure as being competing theoretical accounts of our findings—rather, semantic associations as measured by LSA are one possible large contributor to the predictability effect. These findings are consistent with prior research which showed that the processing advantage for canonical binomials is driven by the joint effects of lexical predictability and lexical-semantic priming (Carrol & Conklin, 2019), beyond their phrasal frequency and word frequency. Our study extends this previous research by showing that even if the binomial's surface form contains variant function words, the tight semantic association between the conjuncts can still lead to a robust expectation for the canonical conjunct. It is possible that this processing advantage includes something above and beyond lexical predictability effects (DeLong et al., 2014; Kuperberg & Jaeger, 2016; Vespignani et al., 2010; Staub, 2015) and semantic priming effects (Hoey, 2005; Jones & Estes, 2012), for example, activation of a mental template (Siyanova et al., 2018) and/or a more complex, cohesive representation of multi-word units within the lexicon (Carrol & Conklin, 2017), but further research is needed to specifically test the contribution of these phrase-level factors in the processing of variant multi-word chunks. Alternatively, the effect of LSA can also be explained by the idea of shallow or heuristic processing strategy in the sentence processing literature (e.g., Sanford & Graesser, 2006; Ferreira, 2003; Ferreira & Patson, 2007; Dwivedi, 2013), which has found that the language processor may generate an initial interpretation by recognizing the lexical-semantic association between words during real-time processing, and a "deeper", more detailed syntactic and semantic computation may later take place only if the task explicitly probes for an interpretation of the sentence (Dwivedi, 2013; Dwivedi, Goertz, & Selvanayagam, 2018; Dwivedi & Gibson, 2017).

Overall, our findings add to the literature which has shown a processing advantage for multi-word chunks over novel strings (Siyanova-Chanturia et al., 2011; Tremblay et al., 2011; Carrol & Conklin, 2019; Siyanova-Chanturia et al., 2017; Vespignani et al., 2020; Molinaro & Carreiras, 2010), and crucially, the current study contributes the novel finding that this processing advantage is observed despite variations in function words, and even if the function word is a defining feature of the multi-word chunk's highly conventional surface form. On the one hand, our findings dovetail with the small set of studies which have found that variant idioms (Kyriacou, Conklin, & Thompson, 2019; Mancuso et al., 2020) and collocations (Vilkaite, 2016) do show a processing speed advantage over matched novel strings. In particular, our findings are consistent with the Vilkaite's (2016) study on collocations, which found a processing advantage for collocates despite intervening words (e.g., *provide some of the information)*. On the other hand, how can our findings be reconciled with the study by Siyanova-Chanturia and colleagues (2017), which found that when the binomials' conjuncts are presented without any conjunctions (e.g., *knife, fork*), the paired conjuncts no longer activate a binomial template nor show facilitated integration? As we suggested earlier, one possible interpretation is that multi-word chunks with variant function words may be matched via analogy or fuzzy matching to the canonical form, rather than a precise template matching mechanism. In the case of variant binomials, the first conjunct (*knife*) and the variant conjunction (*and also the)* may be sufficient to activate the highly associated conjunct entrenched in memory (*fork*) and facilitate its processing, even if the function words in the string do not precisely match the familiar binomial template. As long as a similar conjunction is present, this may be sufficient to facilitate retrieval of the canonical conjunct. Another possible factor which may explain the discrepancy is that in the current study, the binomials were presented in a sentence context rather than in isolation like in Siyanova-Chanturia et al.'s (2017) study, which may support the contextual predictability of the canonical conjunct and explain why we observed a robust processing advantage in variant binomials.

Beyond the processing advantage of variant multi-word chunks over novel strings, another interesting finding that emerged from our experiments is the effect of the variant conjunctions relative to the canonical conjunction, which has implications on the interpretation of similar but non-identical conjunctions. Our results indicated that variant conjunctions (*and* *also*, *as well as*) caused a small but reliable decrease in the expectancy of the canonical conjunct (Experiment 1), and binomials with *and also* received lower typicality ratings than binomials with *and* (Experiment 2a). In addition, conjuncts preceded by *and also* received more leftward regressions and longer regression durations than those preceded by *and* (Experiment 2a), though this effect did not receive substantial evidence in the analyses that included word frequency, LSA, and typicality. Taken together, these results suggest that different conjunctions may convey different pragmatic cues, and that comprehenders are sensitive to these subtle semantic differences, especially when such subtleties are highly relevant to the linguistic task at hand (e.g., in the untimed cloze task that requires the selection of a specific lexical candidate). Although *and*, *and also*, and *as well as* may vary on several dimensions, in this paper we focused on the relative conventionality of the conjunctions in binomials. Prior research suggests that the variant conjunctions, especially *and also*, may convey additive meaning (De Cesare, 2015; Krifka, 1998), unlike the plain *and*. This additive nature and the relative frequency of these conjunctions may account for the differences in reading times on the second conjunct between conjunction conditions, beyond the variance captured by their relative typicality (which we included as a predictor of reading times on the second conjunct in Experiment 2 and 3). Investigating the precise meaning differences between these conjunctions (e.g., semantic/pragmatics, typicality, additive implicatures, and relative frequency between the variant conjunctions) remains a question for future investigation.

Comparing the processing advantage of the variant binomials to the canonical binomials, it is somewhat surprising that variant conjunctions caused a small but reliable reduction in the cloze probability of the canonical conjunct in Experiment 1, yet we did not find an interaction effect between conjunction type and predictability of the conjuncts in the online reading measures (Experiment 2 and 3). This is unlikely to be due to low statistical power, since Experiment 2 and 3 contained sample sizes that far exceed the typical number of participants typically reported in reading studies[5]. One possible explanation is that eye tracking and self-paced reading may not be sensitive to small differences in the cloze probability between the different conjunctions, even if they are standard paradigms for examining reading processes. Another possibility is variant conjunction may disrupt the processing advantage in binomials only when readers have sufficient time to process the linguistic information such as in the cloze task, and reading may simple be too fast a process to capture their influence. Nevertheless, even if we allow for the possibility that there is a small but undetectable interaction effect of conjunction type and predictability in the online reading measures, this does not undermine the most important finding that there is a large binomial processing advantage of the the variant binomials, relative to the control phrases. The main conclusions from this study are based on the processing advantage of the variant binomials over matched control phrases, rather than on the presence or absence of an interaction effect between the canonical and variant binomials.

The present study has important implications on the current models of multi-word representation and processing. First, our findings add to the body of evidence that the processing of multi-word chunks involves accessing their conjunct semantic and syntactic parts (Cutting & Bock, 1997; Konopka & Bock, 2009; Sprenger, Levelt, & Kempen, 2006) rather than being processed as pre-assembled, unanalyzed wholes, as we have shown that internally modifying the function words in binomial phrases does not disrupt the processing advantage for the conjunct noun. Conversely, our results are more congruent with the compositional views (Gibbs, Nayak, & Cutting, 1989) as well as the hybrid views of multi-word storage and retrieval (e.g., Arnon & Snider, 2010; Snider & Arnon, 2012; Arnon & Cohen Priva, 2014), including the Configuration Hypothesis of idiom processing (Cacciari & Tabossi, 1988; Fanari, Cacciari, & Tabossi, 2010; Cacciari & Corradini, 2015) which predicts that linguistic strings will be processed as idiomatic as long as they remain a recognizable configuration. Nevertheless, the fact that we found a robust processing advantage for variant binomials suggests that the definition of multi-word chunks and their implicated processing advantage in these current models may need to be revised to accommodate the variant forms. By definition, multi-word chunks must be recurrent, in the sense that they occur more frequently than comparable novel phrases. The fact that variant binomials still confer a processing advantage over control phrases that the language system still recognizes these strings as being similar enough to the canonical configuration in terms of surface form and overall meaning, instead of treating them like novel strings.

Our findings therefore support construction-based approaches (Goldberg, 2006; Bod, 2009; Fillmore, Kay, & O’Connor, 1988) and usage-based approaches (Tomasello, 2009; Bybee, 1998, 2006) to language, which assume a central role of meaningfulness in the processing of linguistic units. Consistent with these approaches, our findings showed that the function words in idiomatic strings may be modified without disrupting their processing advantage, as long as the modification does not change the overall meaning of the string. Variant binomials may be different from the conventional form, but they still communicate the same overall meaning from the two frequently paired conjuncts that have developed into a cohesive form-meaning pairing over time. This conclusion is also congruent with prior research which found that the overall meaningfulness of idioms and literal multi-word chunks predicts their processing speed, independent of their idiomaticity and whole-string frequency (Jolsvai, McCauley, & Christiansen, 2020). Our study is also relevant to computational approaches to language because it emphasizes the need to consider variability in the representation and mechanisms supporting language learning and use, such that these models of multi-word representation and processing must be able to account for the processing speed advantage of noncanonical variants of recurrent multi-word chunks over matched novel strings (e.g., Geeraert, Newman, & Baayen, 2017).

An important remaining issue is to what extent the processing advantage of multi-word chunks can tolerate variations. For instance, we might expect that the processing advantage may no longer be robust if the variant function words drastically change the overall meaningfulness of the string. One possible approach for addressing this question is increase the degree of modifications to the canonical form, such as examining the processing advantage in binomials with semantically-variant function words and an intervening content word (e.g., *not only the salt but also the fragrant pepper).* Another productive avenue for future research is to examine how the flexibility in the mapping of multi-word chunks varies depending on the semantic properties of the specific items. Results are likely to be different, for example, in irreversible, figurative yet somewhat decomposable binomials (e.g., *safe and sound; hustle and bustle),* whose surface form may not be as open to modifications as the literal *Noun and Noun* binomials we examined in this study. Another interesting question that remains unexplored is whether a processing speed advantage for variant multi-word chunks would also be observed in speech comprehension. This flexibility in processing may not generalize to spoken multi-word chunks because highly conventionalized forms of spoken multi-word chunks tend to undergo phonological reduction (e.g., the conjunction *and* in binomials; Bybee & Scheibman, 1999; Arnon & Cohen Priva, 2013), which may serve as an important cue for prediction. Hence, if that cue is unavailable, the upcoming conjuncts of the string may no longer be highly predictable.

6.1 Conclusions

The experiments reported here showed that there is a robust processing advantage even for variant binomials with noncanonical conjunctions (*and also*, *as well as*) and with intervening determiners (*the, a*, *his*, *her*). This facilitation was observed in both the cloze completion task (Experiment 1) and in the reading tasks (Experiment 2a, 2b, and 3). Crucially, this processing advantage is not attributable to the strings' lexical and phrasal frequency because variant binomials such as *bride as well as the groom* are unattested in corpora, and word frequency did not explain the variance in the eye movement measures. These findings show that although frequency shapes the representation and processing of linguistic units (Arnon & Snider, 2010; Snider & Arnon, 2012; Bybee, 1998, 2006; Tomasello, 2009; Ellis, 2002), the faster processing of familiar phrases seems to hold even if their highly canonical, frequent surface form contains variant function words. This study also highlights the key role of semantic association in the processing of binomials over and above conventionality and frequency of use, consistent with prior research (Carrol & Conklin, 2019; Jolsvai, McCauley, & Christiansen, 2020). Our results also add to the body of literature that has found a robust processing advantage for other types of variant multi-word chunks (Kyriacou et al., 2019; Molinaro et al., 2013; Vilkaite, 2016; Mancuso et al., 2020) and provides the novel insight that binomials do accommodate variations in their function words, despite having a relatively rigid phrasal template (i.e., the highly conventional conjunction *and*). Overall, the evidence we found for flexibility in the processing of multi-word chunks is testimony to the combinatoriality or open-endedness of language. That is, the language system can flexibly recognize and efficiently process variant forms of highly conventional strings, despite the documentation of recurrent multi-word chunks in the language.

Notes

1. Prior to conducting this experiment, we had conducted two preliminary cloze studies that compared *and* to *and also*, and *and also* to *as well as*, respectively. For brevity, we report these results in this footnote rather than in the main text. The results from these two studies are remarkably similar to the cloze data reported in Experiment 1 with respect to the absolute values of the means as well as the data patterns: The first study showed that *and also* (*M* = 80%, *SD* = 40%) led to lower rates of producing the canonical conjunct compared to *and* (*M* = 86%, *SD* = 35%), *p* < .01, and the second study showed that *and also* (*M* = 79%, *SD* = 41%) and *as well as* (*M* = 76%, *SD* = 43%) led to similar rates of producing the canonical conjunct (*p* = .722). These nearly identical results highlight the reliability of the cloze preferences.
2. In this typicality norming study we normed all 44 experimental sentences that were initially created, but here we report the data analysis on the 36 sentences that were actually used in the eye tracking experiments. In the analysis that included all 44 items, the means in each condition differed only slightly and the significance of the effects were also the same, whether 44 or 36 items were included in the analyses.
3. Since the items in the Also condition were identical to the items used in Experiment 2a, we also compared the typicality ratings for the Also items across the two norming experiments. We included Predictability (*canonical* or *novel* conjunct) and Experiment (*2a* or *2b*) as predictors of the typicality ratings. Analyses showed a significant main effect of Experiment ($F(1, 140) = 4.10$, $p < .05$), indicating that the mean typicality ratings for the Also items were significantly higher in Experiment 2b (*M* = 3.59, *SD* = 0.44) than Experiment 2a (*M* = 3.45, *SD* = 0.39). The magnitude of this difference was small, considering the full range of the scale (1 to 7). The main effect of Predictability and the interaction between Experiment and Predictability were not significant: Mean typicality ratings did not significantly differ between the predictable and unpredictable conditions overall, and the typicality ratings for the canonical vs. novel conjuncts did not significantly differ across the two experiments.
4. Since previous research on binomial processing had found that some effects in the self-paced reading paradigm may show up in even later regions of the sentence (Morgan & Levy, 2016), we also examined reading times for the second (N+2) and third (N+3) word after the target

region. These later spillover regions contained some missing data since one item had only two words after the target region, and three items had only three words following the target region. Nevertheless, reading times in these two later regions revealed an effect of Predictability for both the N+2 (β = 10.07, 95% CI = [1.17, 19.00]) and N+3 (β = 9.04, 95% CI = [0.79, 17.41]) regions. Detailed results of these analyses on the later spillover regions are reported as supplementary materials (Appendix S3). When additional predictors (word frequency, LSA, and typicality) were included in the model, the effect of Predictability was still evident in the N+2 region (β = 14.97, 95% CI = [4.70, 25.08]) but not in the N+3 region (β = 8.96, 95% CI = [-0.92, 18.90]).

5. We also conducted a sensitivity analysis to compute the statistical power required to detect the Conjunction x Predictability interaction effect in Experiment 3, using the web-based application PANGEA which can estimate power for a complex design with crossed participant and stimulus random factors (Westfall, 2016; https://jakewestfall.shinyapps.io/pangea). The sensitivity analysis indicated that 174 participants and 36 items provide 93-99% statistical power to detect an effect size of d = .3 to .4 (the typical range of effect sizes reported in psychology; Brysbaert et al., 2018), suggesting that Experiment 3 was sufficiently powered to detect the interaction effect.

References

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *Linguistic Review*, *23*(3), 275–290. https://doi.org/10.1515/TLR.2006.011

Arcara, G., Lacaita, G., Mattaloni, E., Passarini, L., Mondini, S., Benincà, P., & Semenza, C. (2012). Is “hit and run” a single word? The processing of irreversible binomials in neglect dyslexia. *Frontiers in Psychology*, *3*, 1–11. https://doi.org/10.3389/fpsyg.2012.00011

Arnon, I., & Cohen Priva, U. (2013). More than Words: The effect of multi-word frequency and constituency on phonetic duration. *Language and Speech*, *56*(3), 349–371. https://doi.org/10.1177/0023830913484891

Arnon, I., & Cohen Priva, U. (2014). Time and again: The changing effect of word and multiword frequency on phonetic duration for highly frequent sequences. *The Mental Lexicon, 9*(3), 377-400. https://doi.org/10.1075/ml.9.3.01arn

Arnon, I., McCauley, S. M., & Christiansen, M. H. (2017). Digging up the building blocks of language: Age-of-acquisition effects for multiword phrases. *Journal of Memory and Language*, *92*, 265–280. https://doi.org/10.1016/j.jml.2016.07.004

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*(1), 67–82. https://doi.org/10.1016/j.jml.2009.09.005

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, *19*(3), 241-248.

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. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48, https://doi.org/10.18637/jss.v067.i01.

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395, https://doi.org/10.32614/RJ-2018-017

Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory &*

Cognition, *1*(3), 343-346.

Bod, R. (2006). Exemplar-based syntax: How to get productivity from examples. *Linguistic Review*, *23*(3), 291–320. https://doi.org/10.1515/TLR.2006.012

Bod, R. (2009). From exemplar to grammar: A probabilistic analogy-based model of language learning. *Cognitive Science*, *33*(5), 752–793. https://doi.org/10.1111/j.1551-6709.2009.01031.x

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language, 93*, 203–216. https://doi.org/10.1016/j.jml.2016.10.002

Brysbaert, M., Drieghe, D., & Vitu, F. (2012). Word skipping: Implications for theories of eye movement control in reading. *Cognitive Processes in Eye Guidance*, 125–147. https://doi.org/10.1093/acprof:oso/9780198566816.003.0003

Brysbaert, M. & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. Journal of Cognition, 1(1): 9, 1–20.

Buerki, A. (2018). Formulaic sequences: a drop in the ocean of constructions or something more significant?. In I. MacKenzie & M. A. Kayman (Eds.), *Formulaicity and Creativity in Language and Literature* (pp. 15-36). Taylor & Francis.

Bybee, J. 1998. The emergent lexicon. *Chicago Linguistics Society, 34,* 421-435.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 711-733.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: The reduction of don't in English. *Linguistics: An Interdisciplinary Journal of the Language Sciences, 37*(4), 575-596.

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., & Glucksberg, S. (1991). Chapter 9 understanding idiomatic expressions: The contribution of word meanings. *Advances in Psychology*, 217–240.

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

Carrol, G., & Conklin, K. (2014). Eye-tracking multi-word units: some methodological questions. *Journal of Eye Movement Research*, *7*(5), 1-11.

Carrol, G., & Conklin, K. (2019). Is all formulaic language created equal? Unpacking the processing advantage for different types of formulaic sequences. *Language and Speech*. https://doi.org/10.1177/0023830918823230

Chomsky, N. (1965). Aspects of the theory of syntax. MIT Press.

Christiansen, M. H., & Arnon, I. (2017). More than words: The role of multiword sequences in language learning and use. *Topics in Cognitive Science*, *9*(3), 542–551. https://doi.org/10.1111/tops.12274

Culicover, P. W., Jackendoff, R., & Audring, J. (2017). Multiword constructions in the grammar. *Topics in Cognitive Science*, *9*(3), 552-568.

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

Dąbrowska, E. (2014). Recycling utterances: A speaker's guide to sentence processing. *Cognitive Linguistics*, *25*(4), 617-653.

Davies, M. (2008-) *The Corpus of Contemporary American English (COCA): One billion words, 1990-2019*. Available online at https://www.english-corpora.org/coca/.

DeLong, K. A., Troyer, M., & Kutas, M. (2014). Pre-processing in sentence comprehension: Sensitivity to likely upcoming meaning and structure. *Language and Linguistics Compass, 8*(12), 631-645.

De Cesare, A.-M. (2015). Defining Focusing Modifiers in a cross-linguistic perspective. A discussion based on English, German, French and Italian. In K. Pittner, D. Elsner, & F. Barteld (Eds.), *Adverbs: Functional and diachronic aspects* (pp. 47–82). https://doi.org/https://doi.org/10.1075/slcs.170.03ces

Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, C. (2004). Word skipping in reading: On the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology*, *16*(1–2), 79–103. https://doi.org/10.1080/09541440340000141

Dwivedi, V. D. (2013). Interpreting quantifier scope ambiguity: Evidence of heuristic first, algorithmic second processing. *PloS One*, *8*(11), e81461.

Dwivedi, V. D., & Gibson, R. M. (2017). An ERP investigation of quantifier scope ambiguous sentences: Evidence for number in events. *Journal of Neurolinguistics*, *42*, 63-82.

Dwivedi, V. D., Goertz, K. E., & Selvanayagam, J. (2018). Heuristics in language comprehension. *Journal of Behavioral and Brain Science*, *8*(07), 430.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*(2), 143–188. https://doi.org/10.1017/S0272263102002024

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, *20*(1), 29–62. https://doi.org/10.1515/text.1.2000.20.1.29

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.

Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 555–568.

Fillmore, C. J., Kay, P., & O'connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language*, 501-538.

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

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning Memory and Cognition*, *31*(5), 862–877. https://doi.org/10.1037/0278-7393.31.5.862

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.

Gentner, D., & Markman, A. B. (1997). Structural alignment in analogy and similarity. *American Psychologist*, *52*(52), 45–56. Retrieved from https://pdfs.semanticscholar.org/7d65/5b765d962d47fa3eeb677c7411056b38165b.pdf

Gibbs, R. W. (1990). Psycholinguistic studies on the conceptual basis of idiomaticity. *Cognitive Linguistics, 1*(4), 417–451.

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, *7*(5), 219–224. https://doi.org/10.1016/S1364-6613(03)00080-9

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. New

York: Oxford University Press.

Hoey, M. (2005). *Lexical priming: A new theory of words and language.* London, UK: Routledge.

Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E. van der Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and Psychological Perspectives* (pp. 133-166). Lawrence Erlbaum.

Jackendoff, R., & Audring, J. (2018). Relational morphology in the parallel architecture. In F. Masini & J. Audring (Eds.), *The Oxford Handbook of Morphological Theory* (pp. 390–408). https://doi.org/10.1093/oxfordhb/9780199668984.013.33

Jackendoff, R., & Audring, J. (2019). *The texture of the lexicon: relational morphology and the parallel architecture.* Oxford University Press.

Janssen, N., & Barber, H. A. (2012). Phrase frequency effects in language production. *PLoS ONE*, *7*(3). https://doi.org/10.1371/journal.pone.0033202

Jolsvai, H., McCauley, S. M., & Christiansen, M. H. (2020). Meaningfulness beats frequency in multiword chunk processing. *Cognitive Science*, *44*(10), e12885.

Jones, L. L., & Estes, Z. (2012). Lexical priming: Associative, semantic, and thematic influences on word recognition. In J. S. Adelman (Ed.), *Visual word recognition: Meaning and context, individuals and development* (pp. 44–72). Psychology Press.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*(2), 228.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1–2), 262–284. https://doi.org/10.1080/09541440340000213

Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology*, *58*(1), 68-101.

Krifka, M. (1998). Additive particles under stress. *Semantics and Linguistic Theory, 8*(3), 111. https://doi.org/10.3765/salt.v8i0.2799

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31*(1), 32-59.

Kyriacou, M., Conklin, K., & Thompson, D. (2019). Passivizability of idioms: Has the wrong tree been barked up? *Language and Speech*, *63*(2), 404-435. https://doi.org/10.1177/0023830919847691

Lambrecht, K. (1984). Formulaicity, frame semantics, and pragmatics in German binomial expressions. *Language*, 753-796.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240. https://doi.org/10.1037/0033-295X.104.2.211

Langacker, R. W. (1986). An introduction to cognitive grammar. *Cognitive Science*, *10*(1), 1-40.

Langacker, R. W. (1987). *Foundations of cognitive grammar – Volume 1: Theoretical prerequisites*. Stanford University Press.

Langlotz, A. (2006). *Idiomatic creativity: A cognitive-linguistic model of idiom-representation and idiom-variation in English* (Vol. 17). John Benjamins Publishing.

Malkiel, Y. (1959). Studies in irreversible binomials. *Lingua*, *8*, 113–160. https://doi.org/10.1016/0024-3841(59)90018-X

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*, 99-124

McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC Press.

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

Molinaro, N., & Carreiras, M. (2010). Electrophysiological evidence of interaction between contextual expectation and semantic integration during the processing of collocations. *Biological Psychology*, *83*(3), 176-190.

Molinaro, N., Canal, P., Vespignani, F., Pesciarelli, F., & Cacciari, C. (2013). Are complex function words processed as semantically empty strings? A reading time and ERP study of collocational complex prepositions. *Language and Cognitive Processes*, *28*(6), 762–788. https://doi.org/10.1080/01690965.2012.665465

Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. Oxford University Press.

Morgan, E., & Levy, R. (2016). Abstract knowledge versus direct experience in processing of binomial expressions. *Cognition*, *157*(617), 384–402. https://doi.org/10.1016/j.cognition.2016.09.011

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational Ideas—Part II. *Linguistics and Language Compass*, *10*(11), 591–613. https://doi.org/10.1111/lnc3.12207

Pawley, A., & Syder, F. H. (1983). Natural selection in syntax: Notes on adaptive variation and change in vernacular and literary grammar. *Journal of Pragmatics, 7*(5), 551-579.

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.

Pinker, S. (1999). *Words and rules: The ingredients of language*. Harper Collins.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 514–528. https://doi.org/10.1037/a0020990

Schmitt, N. (2006). Formulaic language: Fixed and varied. *Estudios de Lingüística Inglesa Aplicada*, 6, 13-39.

Schröder, D. (2015). Chapter 12 Spilling some linguistic beans: On the syntactic flexibility of idioms. In W. Malec & M. Rusinek (Eds.), *Within Language, Beyond Theories (Volume III): Discourse Analysis, Pragmatics and Corpus-based Studies* (pp. 190-205). Cambridge Scholars Publishing.

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

Siyanova-Chanturia, A., Conklin, K., Caffarra, S., Kaan, E., & van Heuven, W. J. B. (2017). Representation and processing of multi-word expressions in the brain. *Brain and Language, 175*(October), 111–122. https://doi.org/10.1016/j.bandl.2017.10.004

Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. J. B. (2011). Seeing a phrase “time and again” matters: The role of phrasal frequency in the processing of multiword sequences.

Journal of Experimental Psychology: Learning Memory and Cognition, *37*(3), 776–784. https://doi.org/10.1037/a0022531

Siyanova-Chanturia, A., & Janssen, N. (2018). Production of familiar phrases: Frequency effects in native speakers and second language learners. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(12), 2009.

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. 127–164). De Gruyter. https://doi.org/https://doi.org/10.1515/9783110274073

Sopher, H. (1978). Positional features of "also." *Linguistics*, *59*(6), 538–555. https://doi.org/10.1080/00138387808597931

Sosa, A. V., & MacFarlane, J. (2002). Evidence for frequency-based constituents in the mental lexicon: Collocations involving the word of. *Brain and Language*, *83*(2), 227–236. https://doi.org/10.1016/S0093-934X(02)00032-9

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

Stan Development Team (2018). *RStan: the R interface to Stan.* R package version 2.17.3. http://mc-stan.org

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass*, *9*(8), 311-327.

Staub, A., & Goddard, K. (2018). The role of preview validity in predictability and frequency effects on eye movements in reading. *Journal of Experimental Psychology: Learning Memory and Cognition*. https://doi.org/10.1037/xlm0000561

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, *18*(5), 523–534. https://doi.org/10.1016/S0022-5371(79)90284-6

Tabossi, P., Wolf, K., & Koterle, S. (2009). Idiom syntax: Idiosyncratic or principled?. *Journal of Memory and Language*, *61*(1), 77-96.

Taylor, J. R. (2002). *Cognitive grammar.* Oxford: Oxford University Press.

Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism*

Bulletin, *30*(4), 415–433. https://doi.org/10.1177/107769905303000401

Tomasello, M. (2009). *Constructing a language*. Harvard University Press.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning*, *61*(2), 569–613. https://doi.org/10.1111/j.1467-9922.2010.00622.x

Tremblay, A., & Tucker, B. V. (2011). The effects of N-gram probabilistic measures on the recognition and production of four-word sequences. *The Mental Lexicon, 6*(2), 302-324.

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

Vilkaite, L. (2016). Are nonadjacent collocations processed faster? *Journal of Experimental Psychology: Learning Memory and Cognition*, *42*(10), 1632–1642. https://doi.org/10.1037/xlm0000259

Von der Malsburg, T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language, 94*, 119-133.

Westfall, J. (2016). *PANGEA: Power ANalysis for GEneral Anova designs*. Retrieved from http://jakewestfall.org/publications/pangea.pdf

Wulff, S. (2008). *Rethinking idiomaticity: A usage-based approach*. Bloomsbury Publishing.

Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language*, *72*(1), 37–48. https://doi.org/10.1016/j.jml.2013.12.003W