Comprehension and written production of English modifier-noun phrases: effects of meaning predictability
Christina L. Gagné, Thomas L. Spalding, & Juana R. Park, University of Alberta

Abstract
We apply the concept of meaning predictability developed by Štekauer (2005, 2006, Štekauer, Diaz-Negrillo & Valera 2011) to the cognitive processing of modifier-noun phrases (e.g., wool basket) and determine whether meaning predictability is a viable psycholinguistic construct. We review and describe three sets of experiments that examined the role of meaning predictability. In the first set of experiments, we investigated how each of the two components of meaning predictability (i.e., prevalence and goodness) influenced the reaction times in a series of tasks that involved the interpretation of modifier-noun phrases. The results of these experiments showed that highly predictable meanings elicit faster responses in a sense/nonsense judgment task and in a meaning verification task. Furthermore, the results indicated that goodness and prevalence played different roles in the two tasks, which suggests that these two components of meaning predictability are best evaluated separately rather than as a single predictability rate measure. In the second set of experiments, we investigated the influence of sentential context on response times in a comprehension task and found that the influence of context depended on meaning dominance. Finally, in the third set of experiments we investigated whether the influence of meaning dominance extends to written production. We found that meaning dominance influences inter-key typing latencies but that this influence differs across the various positions in the phrase.

Keywords: novel compounds, modifier-noun phrases, meaning predictability, meaning dominance, psycholinguistics
1 Introduction

The notion of meaning predictability (Štekauer 2005, 2006, Štekauer, Diaz-Negrillo & Valera 2011) is highly relevant for psycholinguistic theories of compound processing because many compounds allow multiple interpretations. For example, *snowman* is most commonly interpreted as a man that is made of snow (such as Frosty the Snowman), but can also refer to a man that shovels snow (i.e., it is a type of occupation akin to the compound *garbageman*). In this chapter, we discuss three streams of psycholinguistic research that each examine a particular aspect of meaning predictability and meaning dominance. Our aim is to determine whether the ease with which a compound can be processed is affected by meaning predictability and dominance. For example, are compounds with high predictability more easily processed than ones with lower predictability, and does this effect depend on other factors?

Meaning predictability was defined as the probability that one particular meaning of a word or combination of words will be preferred over all its other possible interpretations when it is heard or read for the first time without any context clues. Štekauer defines predictability in two ways. Predictability Rate refers to the strength of a particular interpretation (or “reading”). Objectified Predictability Rate refers to an interpretation’s ranking relative to other interpretations. We will focus on Predictability Rate, as it is more applicable to evaluating competing readings for a given word or combination of words, while the Objectified Predictability Rate is more applicable to evaluating readings across different words or combinations of words. In his research, Štekauer calculated Predictability Rate for each naming unit by multiplying the proportion of participants providing that interpretation (i.e., prevalence) by the goodness of that interpretation. He predicted that the interpretation with the highest predictability would be the one that is most likely to be selected when the naming unit is encountered out of context. His data confirmed this prediction.

Štekauer’s (2006) theory of meaning predictability posits that meaning predictability is based on linguistic characteristics, as well as on real-world knowledge (i.e., knowledge about the actual objects). Central to this theory is the notion that potential meanings are in competition with each other. The notion of competition fits
well with the Competition Among Relations in Nominals (CARIN) theory of conceptual combination (Gagné & Shoben 1997) and, the more recent version, the Relational Interpretation Competitive Evaluation (RICE) theory (Spalding, Gagne, Mullaly, & Ji 2010) because these theories also rely heavily on the notion of competition among potential interpretations. CARIN and RICE posit that the meaning of a compound word or novel modifier-noun phrase is mutually determined by the ease with which the constituents can be assigned to a particular role within a relational structure and by the availability of the appropriate relational structure. Various relations compete for selection, and increased competition is associated with increased processing difficulty due the extra time required to settle on a particular relational interpretation.

This prediction has been confirmed in that several studies have shown that increased competition is associated with increased processing difficulty. The ease of constructing a meaning depends on the relative strengths of different possible relational interpretations, largely because relational interpretations compete to be selected (Gagné & Shoben 1997; Spalding & Gagné 2008). Relational competition is affected by knowledge about the constituents, such as the modifier's general usage with various relations (Gagné & Spalding 2009), the modifier’s recent usage (Gagné 2001, 2002), the head’s general and recent usage (Spalding et al. 2010), semantic information associated with the head (Spalding & Gagné 2007), and sentential and discourse context (Gagné & Spalding 2004; Gagné et al. 2005).

Competition among potential readings is a central tenet of Štekauer’s theory and this fits well with the existing psycholinguistic evidence. Thus, it seems promising to systematically examine various aspects of meaning predictability in the context of comprehension and written production to gain insight into how competition affects psycholinguistic processing and thereby improve our understanding of the language and cognitive systems. Before proceeding, we should point out that our application of meaning predictability does not map directly onto Štekauer’s theory, but rather we examine another facet of the construct of predictability. Štekauer’s theory posits the existence and use of prototypical semes which are involved in combining modifiers and heads. In this sense, his semantic analysis is more fine-grained in terms of the semantic content of words than is that of theories based on the notion of relational structures. For
example, Levi (1978) focuses more on the nature of what she calls “recoverable deletable predicates” (what we are calling relational structures) than on the particular semantics of the modifiers and heads entering into those relations. In the psycholinguistic evidence that we will discuss, the emphasis will also be on the way in which ease of processing is affected by knowledge about the types of relations that modifiers and heads occur, rather than on the semantic characteristics of those modifier and heads. Although it is true that the prototypical semes are more specific than the relations, it is also the case that relational structures embody multiple correlated semes, and, consequently, one advantage of the relational approach is that the fact that a particular word is used often with a particular relation means that this also takes account of many semes for that word.

However, this is not to say that the semantics of the words is not important (see Gagné & Spalding 2014; Spalding et al. 2010, for a discussion of this issue). It is the case, for example, that chocolate has the capacity to function as a modifier in a MADE OF relation because it is a material. Moreover, knowledge about the semantic properties of the constituent concepts (i.e., modifier and heads) as well as pragmatic information also play a role in determining which relational interpretations are possible. For example, the word mountain denotes a place and, thus, is often used with a LOCATED relation (e.g., mountain stream, mountain cabin, and mountain goat). However, for mountain planet, the LOCATED relation is not possible because planets are too big to be located in the mountains. That said, the primary aim of the current paper will be on the relations more so than on the semantics of the words that have been linked by the relation.

In the remainder of this paper, we will discuss empirical evidence that addresses various aspects of meaning predictability. We begin by examining the two subcomponents of Štekauer’s predictability rate (goodness and prevalence). Next, we examine the influence of meaning dominance on ease of comprehension. Finally, we extend our examination of meaning dominance to written production.
Influence of goodness and prevalence on the interpretation of modifier-noun phrases

The aim of the first set of experiments was to apply Štekauer’s claims about meaning predictability to psycholinguistic processing. Gagné, Marchak, and Spalding (2010) investigated how meaning predictability (measured by predictability rate) and its two components (prevalence and goodness) influenced the reaction times in a series of tasks that involved the interpretation of novel modifier-noun phrases. The items were selected from those used in Gagne and Shoben (2002). In Gagne and Shoben (2002), participants were presented with potentially ambiguous noun-noun phrases (e.g., grass net, picture lamp) and were asked to provide a paraphrase of each item. The responses were then classified (by the researchers) based on Levi’s (1978) relational categories. The percentage of responses within each relation was used to determine the relative frequency. For example, only three relations were used for cotton bag: bag MADE OF cotton was used in 67% of the responses, bag FOR cotton was used in 23% of the responses, and bag HAS cotton was used in 10% of the responses. These percentages are an indication of a particular reading’s frequency and, thus, correspond to a reading’s prevalence (or meaning dominance). To measure a reading’s goodness, a separate study was conducted using the dominant and second most dominant reading for each item. Participants rated the goodness of each reading on a 7-point rating scale. The items were counterbalanced such that each person only saw one reading for each item.

In Experiment 1 of Gagne et al. (2010), participants were presented with novel modifier-noun phrases (e.g., wool basket) that had at least two different possible interpretations (e.g., basket MADE OF wool or basket FOR wool). The interpretations corresponded to the dominant and second most dominant meanings based on the results of the Gagne and Shoben (2002) paraphrase study. The participants had to decide if each of the modifier-noun combinations had a sensible, literal interpretation, or not. The results of this sense-nonsense task showed that meaning predictability and each of its components taken separately were good predictors of reaction times. Participants took less time to decide that an item had a sensible interpretation when it had a high predictability rate, when it had high prevalence, or when it had high goodness. Also, the
results revealed that goodness was a better predictor of reaction times in this task than was prevalence.

In Experiment 2, participants were presented the same phrases, each one followed either by its dominant or subdominant meaning (e.g., *adolescent doctor* followed by a *doctor for adolescents* or a *doctor that is an adolescent*). The task consisted in deciding whether the definition provided was plausible or not. As in Experiment 1, meaning predictability and each of its components taken separately were valid predictors of reaction times. However, unlike Experiment 1, prevalence and goodness interacted. When goodness was high, participants were faster to respond to items with high prevalence than to items with low prevalence. When goodness was low, however, higher prevalence was associated with slower responses.

These results not only support the idea that meaning predictability plays a role in the processing of modifier-noun phrases, but also reveal that prevalence and goodness exert different influences depending on the type of task (sense-nonsense versus verification tasks). These findings suggest that, when examining psycholinguistic processing, it is more appropriate to evaluate goodness and prevalence separately rather than combining them into a Predictability Rate measure.

### 3 The influence of meaning dominance on the comprehension of familiar and novel modifier-noun phrases

Štekauer has shown that most items tend to yield one reading that is most predictable. In this section, we examine the psycholinguistic consequences of such meaning dominance. Gerrig (1989) found that the time required to produce an innovative reading for a compound such as *dog sled* (e.g., a sled that has a picture of dog on it) is affected by the presence of the conventional (i.e., lexicalized) meaning (e.g., a sled pulled by dogs). That is, the conventional meaning interferes with the creation of an innovative meaning. Is the reverse true — does the creation of an innovative meaning affect processing of the conventional meaning? To address this question, Gagné, Spalding, and Gorrie (2005) examined whether recent exposure to an innovative meaning influences the processing of the conventional meaning and vice-versa. The participants were native
English speakers. All were undergraduates enrolled in a first-year Psychology course and received partial course credit for their participation. Each person participated in only one experiment.

In Experiment 1, participants saw a novel modifier-noun phrase (e.g., grass net) on a computer screen and judged which of two possible definitions was best. The relation that was chosen most frequently was deemed the dominant relation and the other was the subdominant relation. In addition to allowing us to identify which meaning was dominant for a given phrase, the results of this experiment indicated that the time required to decide which definition was best was correlated with meaning dominance (i.e., the percentage of people selecting the most popular meaning), $r = -.38$. The more dominant the definition was, the less time it took to respond. This result suggests that the less competition there is among possible interpretations for a novel modifier-noun phrase, the easier it is to select an interpretation.

In Experiment 2, we examined whether sentential context affected the availability of the dominant and subdominant meanings of novel (e.g., grass net) and familiar (e.g., silk worm) noun-noun phrases. The familiar phrases are also called open-compounds in the psycholinguistic literature. Participants viewed the phrases as part of a sentence that was consistent with either the dominant meaning (e.g., “The thread that a silk worm produces is often used by Kim to make beautiful scarves”) or with the innovative meaning (e.g., “Kim decided it would be fun to make a silk worm out of the fabric she had bought”). Immediately after viewing this prime sentence, the participants viewed the target phrase (e.g., silk worm) with either the dominant definition (e.g., a worm that produces silk) or the innovative meaning (e.g., a worm made of silk) and indicated whether the definition was plausible. The experimental items include an equal number of novel phrases and established phrases. To balance the number of yes and no responses, filler items that did not have plausible definitions were also presented.

The results indicated that sentential context affected the relative availability of the dominant and subdominant meanings for both novel and known phrases. Participants took longer to indicate that the dominant definition was plausible when the sentence supported the subdominant meaning than when it supported the dominant meaning. In addition, the percentage of “plausible” responses to the dominant meaning
was reduced when the sentence used the subdominant meaning. The impact of sentential context was greater for the novel phrases than for the known phrases. However, even for the known phrases, the plausibility judgments for the dominant meaning was reduced when the sentence was biased towards the subdominant meaning. These findings suggest that the dominant (established) meaning was competing with the innovative meaning constructed in the previous sentence and this competition decreased the availability of the established meaning.

4 The influence of meaning dominance on written production

More recently, we have investigated whether the influence of meaning dominance that was observed in comprehension is also seen in written production. Previous research has used typing time to study the production of compounds. The typing task is well-suited for obtaining information about the processing of a word on a letter-by-letter basis, which allows researchers to examine differences in processing difficulty at various points in the word. That is, the data generated from this task (i.e., the time spent to type each letter of a word on a computer keyboard) makes it possible to study not only the production of the words as wholes, but also the smaller pieces that compose them. For example, Sahel, Nottbusch, Grimm, and Weingarten (2008) found that there was an elevation in typing times for German compounds at the morpheme boundary for both semantically transparent and opaque compounds, which suggests that morphology operates without recourse to the meaning of the constituents. Libben and Weber (2014) also found an increase in typing latency at the morpheme boundary for English compounds. This increase was smaller for opaque-opaque (OO, e.g., fleabag) compounds than for transparent-transparent (TT, e.g., bedroom) and opaque-transparent (OT, e.g., strawberry) compounds. These studies indicate that typing latencies are sensitive to psycholinguistic properties of a word.

We used this methodology to determine whether the influence that meaning dominance exerts on the comprehension of modifier-noun phrases is also present in production. Our experimental materials were the 68 novel modifier-noun phrases that appeared in Gagné, Spalding, and Gorrie (2005). Unlike the familiar (i.e., lexicalized)
compounds used by Libben and Webber (2014), our items were novel phrases and were written with a space between the two words. The measure of meaning dominance was obtained from this previous research and was defined as the percentage of people that selected the most preferred definition for each item. To obtain these data, Gagné et al. (2005) presented people with the two most popular paraphrases for an item, based on Gagné and Shoben’s (2002) study, and asked people to indicate which definition was the best. The participants were all native-English speakers and were enrolled in a first-year Psychology course.

In the current study, each noun-noun phrase was displayed on a computer screen and the participant pressed the space bar once he/she had read it. Next, the phrase disappeared and the participant typed the item they had just read. The time to read each phrase and the inter-key typing times (i.e., the time spent to type each of the letters of the words) were recorded by the computer program. Previous studies in our lab and in other labs (e.g., Libben & Weber 2014) using compound words (e.g., blueberry) have shown that, when analyzing typing time data, the key positions to look at are usually the first letter typed and the letters around the boundary between the two words or the words’ morphemes. In the case of our novel modifier-noun phrases, the key positions to analyze were the first letter of the modifier and the letters before and after the space between the modifier and the noun. We expected that participants would take longer to initiate the typing if the item had meanings that were approximately balanced; on the contrary, if an item had a meaning that was clearly dominant, people would start typing it faster. We also expected that this same tendency would be present at the boundary between the two words, that is to say, that the two words of phrases that were more balanced would present larger time delays between the two component words when typed.

The design was within-subjects; each person received all experimental conditions and, thus, acted as their own control. Consequently, any differences due to typing skills or other person-specific variables were constant across the various conditions. The inclusion of subjects as a random factor in our data analysis also controlled for person-specific differences.
The data was analyzed using Linear Mixed Effects (LME) models (see Pinheiro & Bates 2000) using the \textit{mixed} and \textit{contrast} functions in Stata 13. In all models, subjects and items were treated as random factors. The dependent variable, typing latency, was log-transformed to reduce skewness.

We found that meaning dominance (i.e., the percentage of participants that produced the most dominant meaning) was a valid predictor of the time taken to initiate the typing task (i.e., the time to type the first letter of the phrase), \( z = -2.08, p = .04 \). Typing time was faster for items with high dominance than for items with lower dominance. This finding suggests that the more ambiguous the meaning of a modifier-noun phrase is, the slower people are to initiate the production of that phrase, which may reflect a difficulty in the attempt to settle on a meaning before starting to produce the phrase.

To examine whether there was evidence of the use of a constituent structure during written production, we compared the typing time for the letter before and after the space (e.g., for the phrase \textit{wool basket}, we compared the times for typing the letters \textit{l} and \textit{b}). We also compared the typing time at each of these two positions to the space itself. The typing latency for the first letter of the second constituent (\( M = 219 \) ms, \( SD = 148 \)) was longer than typing latency for the last letter of the first constituent (\( M = 152 \) ms, \( SD = 86 \)), \( z = 21.40, p < .001 \). In addition, the time to type the space between the two constituents (\( M = 157 \) ms, \( SD = 86 \)) was also longer than the typing time for the last letter of the first constituent. Although small in magnitude, this difference was statistically significant, \( z = 2.34, p = .02 \). These results indicate that when typing a modifier-noun phrase, people did not output it as a single entity, but rather the entire phrase appears to be stored in memory as two parts and each part is used as a planning unit. Additional time is needed after the output of the first unit to access the planning unit for the second part of the phrase.

Next, we included dominance in our analysis to see whether this variable influenced typing time. Meaning dominance interacted with letter position, \( \text{Chi}(2) = 8.24, p = .02 \); the differences in typing time that were observed for the letter before, at, and after the space were influenced by dominance. We conducted two further analyses to understand the nature of this interaction. One analysis focused on the space and the
last letter of the first constituent and the second analysis focused on the space and the first letter of the second constituent. The influence of dominance differed for the space and the last letter of the first constituent, $z = 2.97$, $p = .003$; typing time at the end of the first constituent increased as dominance increased, whereas typing time at the space decreased (i.e., typing became faster) as dominance increased.

Tests of the differences in the slopes for the two letter positions across the various values of dominance revealed that the difference in typing latencies at these two positions emerged only when dominance was above 70%. When dominance was above 70%, typing times for the space were longer than typing times for the last letter of the first constituent and this difference increased as dominance increased. To illustrate, producing the last letter of the first word of a phrase with a highly dominant meaning (e.g., rubber scoop, for which dominance is 96%) took less time than did producing the space. However, this difference was not observed for a phrase with a less dominant meaning (e.g., wool basket, for which dominance is 51%). In contrast, the second set of analyses revealed that the influence of dominance was constant for the space and the first letter of the second constituent (i.e., the two variables did not interact), $z = .51$, $p = .61$. The results of a model that did not include this interaction term revealed that at these two positions (i.e., the space and the first letter of the second constituent), dominance had no effect, $z = .59$, $p = .56$, though as in the previous analysis without dominance as a predictor, the space was typed much more quickly than the first letter of the second word, $z = 18.77$, $p < .0001$.

In sum, these findings support the idea that meaning dominance impacts not only the comprehension but also the written production of novel word combinations, especially in key positions, such as the beginning of the first constituent, the space between the two constituents, and the letters surrounding the space. Our findings also show that the impact of meaning dominance is not constant across the production of the word; the influence of meaning dominance occurs during the production of the first constituent of a phrase but not during the production of the second constituent.
5 Discussion and conclusions

Štekauer (2005) argued that many theories of word interpretation did not try to identify which of the many possible readings for a word was most predictable, and proposed that these theories could be improved by considering more carefully the word formation process, and specifically by looking at the ways meaning predictability affects that process. Our aim in this paper has been to present evidence that meaning predictability is viable as a theoretical construct not only in the analysis of the formation of new meaning units, but also in the psycholinguistic processing of modifier-noun phrases, both known (i.e., established compounds) and novel.

Štekauer’s idea of meaning predictability inherently includes an aspect of competition among the possible readings, and this competition among various readings was one of the factors that had been neglected in the literature. Although not exactly the same as meaning predictability, the related notion of competition among relational interpretations of modifier-noun phrases and compounds has been investigated in a large number of studies (see Gagné & Spalding 2014 for an overview). In the current paper, we have brought these approaches together, and have shown that competition among relational interpretations affects the processing of both known (e.g., silk worm) and unknown (e.g., cream cloud) modifier-noun phrases. Moreover, the streams of research that we discussed show that such phrases are affected by meaning predictability in its various forms (e.g., goodness, prevalence, and dominance). Importantly, the different measures of meaning predictability do not always affect processing in the same way, suggesting the importance of maintaining the various measures as separate theoretical constructs. Perhaps most interesting is that meaning dominance affected not only the comprehension of the phrases, but also the ease of production. This finding is particularly interesting because presumably the phrase had already been interpreted prior to production and thus the meaning of the phrase was available prior to typing. The influence of dominance on the output of the first constituent of a phrase indicates that the meaning resolution process has effects that last beyond the time-frame of interpreting a compound or modifier-noun phrase.
Taken together, the results of these streams of research indicate that the notion of predictability in its various forms is a viable psycholinguistic construct and, as such, an improved understanding of the effects of meaning predictability provides valuable insight into the structure of human language and into the organization and functioning of human language system.

References


Gagné, Christina L. & Spalding, Thomas L. 2009. Constituent integration during the processing of compound words: Does it involve the use of relational structures? In *Journal of Memory & Language*, vol. 60. 20-35.


Christina Gagné  
Department of Psychology  
University of Alberta  
P217 Biological Sciences Building  
Edmonton Alberta Canada  
T6G 2E9  
cgagne@ualberta.ca

Thomas Spalding  
Department of Psychology  
University of Alberta  
P217 Biological Sciences Building  
Edmonton Alberta Canada  
T6G 2E9  
spalding@ualberta.ca

Juana Park  
Department of Psychology  
University of Alberta  
P217 Biological Sciences Building  
Edmonton Alberta Canada  
T6G 2E9  
juana@ualberta.ca