The effects of frequency and predictability on the representation of multi-word phrases in English

Recent research has suggested that learners of a language may store more than just words, including multi-word phrases (e.g., *I don't know*; Bybee & Scheibman, 1999). However, it's still unclear what drives the storage of multi-word phrases. While frequency has often been cited as the driving factor behind storage (e.g., Bybee & Scheibman, 1999), it's also possible that predictability may drive the storage of multi-word phrases. For example, predictability has been shown to be predictive of both phoneme-learning and word-learning (Olejarczuk et al., 2018; Ramscar et al., 2013).

Moreover, the nature of the internal representation of stored items remains unclear. One possible consequence of storage is that stored multi-word phrases may retain their internal structure, but it's also possible that they lose at least part of their internal structure (see the right side of Figure 1). For example, in Kapatsinski & Radicke (2009) participants were tasked with pressing a button as soon as they heard the segment *up.* They found that for English verb+*up* collocations (e.g., *pick up*), reaction time as a function of frequency followed a u-shaped curve: as frequency of the phrasal verb increased, participants were quicker to recognize the morpheme *up*, except for the highest frequency collocations, where they were actually *slower* to recognize *up*. They argued that this slowdown for the highest frequency verbs occurs because these verb+*up* combinations are stored as holistic "prefabs" which interfere with the recognition of their component parts (because recognition of the prefab either competes with or supersedes the recognition of its components). They thus argue that the highest frequency phrasal verbs are stored holistically with at least partial loss of their internal structure.

In the present study, we examine the factors that drive storage and the internal representation of stored items. We do so by replicating and extending Kapatsinski & Radicke (2009) to examine the effects of frequency as well as predictability on the recognizability of V+*up* phrases. Our stimuli consisted of 200 V+*up* phrases that varied across frequency and predictability (see Example 1) and 200 matched sentences that didn't contain V+*up* phrases. Predictability was operationalized as the odds ratio of the probability of *up* occurring immediately after the verb to the probability of any other word occurring, calculated using the Google *n*-grams corpus.

We recruited 350 participants for our experiment, although 44 were excluded for having an accuracy below 70%, leaving 306 participants altogether. We analyzed our data using Generalized Additive Mixed Models which allow our fixed-effects to vary nonlinearly. Our dependent measure was reaction time and our fixed effects were the interaction between frequency and predictability. We also used maximal random effects.

Our GAM model suggested that there were non-linear effects of both frequency and predictability. There was no significant interaction between predictability and frequency. In order to follow-up on this, we then ran a Bayesian quadratic regression model that included frequency, predictability, and quadratic terms for both (Table 1). The results of the Bayesian model confirmed that there was an effect of the quadratic term for predictability and for frequency (Figure 2).

Our results suggest a U-shaped effect of both predictability and frequency: as predictability or frequency increases, participants become faster at recognizing the segment *up*, except for the highest predictability or frequency items where there is actually an increase in reaction times. This suggests that verb phrases may lose some amount of their internal representation if they are high enough predictability or frequency. Since holistic storage in some form is a prerequisite for loss of internal structure, our results further suggest that frequency and predictability contribute independently to storage of multi-word phrases. Overall, these results suggest that stored multi-word phrases may have varying internal representations mediated by both predictability and frequency.

Example 1: Example stimuli. For the sake of demonstrating the range of our materials, we show stimuli that represent "high" to "low" frequency/predictability, but note that both frequency and predictability were measured on a continuous scale. The position of the verb phrase varied across sentences such that sometimes they were at the beginning of the sentence, the middle of the sentence, or towards the end of the sentence.

1. Low-frequency, high-predictability: Finally, he **wised up** and realized the truth.
2. Low-frequency, low-predictability: They **lightened up** the mood after the tense argument.
3. High-frequency, high-predictability: He **picked up** the phone and answered the call.
4. High-frequency, low-predictability: They **made up** after their argument and became friends again.

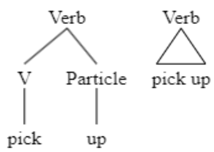

Figure 1: A visualization of multi-word storage with the left picture representing a completely intact internal structure and the right picture representing a complete loss of internal structure, i.e. a loss of internal morphological representation.

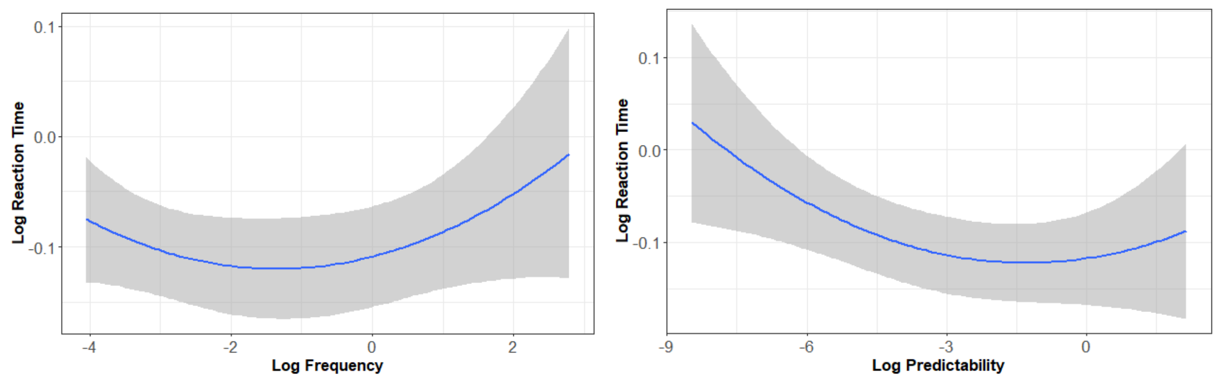

Figure 2: The predicted effects of log frequency (left) and log predictability (right) on log reaction times. The model suggests a u-shaped effect for both on the recognizability times of the segment *up*.

	Estimate	Est.Error	Q2.5	Q97.5
Intercept	-0.102	0.029	-0.161	-0.046
log-frequency	0.019	0.011	-0.002	0.041
log-predictability	0.009	0.011	-0.013	0.032
duration	-0.135	0.098	-0.328	0.057
log-predictability^2	0.003	0.002	-0.000	0.007
log-frequency^2	0.005	0.004	-0.002	0.012

Table 1: Bayesian quadratic model results. Note that while the CIs for log(predictability^2) and log(frequency^2) cross zero, over 96% of the posterior samples for log(predictable^2) were greater than zero, and nearly 95% of the posterior samples for log(frequency^2) were greater than zero.

References:

1. Bybee, J. and Scheibman, J. (1999). The effect of usage on degrees of constituency: the reduction of don't in english. 37(4):575–596. 2. Chomsky, N. (1975). Reflections on language. 3. Kapatsinski, V. and Radicke, J. (2009). *Formulaic language, 2:499–520*. 4. Olejarczuk, P., Kapatsinski, V., & Baayen, R. H. (2018). Distributional learning is error-driven. *Linguistics Vanguard*, 4(s2), 20170020. 5. Ramscar, M., Dye, M., & Klein, J. (2013). Children value informativity over logic in word learning. *Psychological science, 24(6)*, 1017-1023.