Corrigendum to “Grammatical constraints on phonological encoding in speech production”

Jordana R. Heller and Matthew Goldrick

Department of Linguistics
Northwestern University
2016 Sheridan Rd
Evanston, IL 60208

Address for correspondence:
Jordana Heller
Department of Linguistics, Northwestern University
2016 Sheridan Rd
Evanston, IL 60208
Phone: (206) 650-2585
Fax: (847) 491-3770
Email: jordana.heller@gmail.com
In a previous article (Heller & Goldrick, 2014), we reported that grammatical encoding constrains phonological and phonetic processing. At both stages of processing, our analyses suggested that neighbors that share a target’s grammatical category exert a significant influence on target processing, particularly when grammatical constraints are strong (i.e., naming within a sentence context vs. bare picture naming). Here, we report an error in the statistical analyses of the acoustic data that undermine our claims regarding phonetic processing. Our conclusions regarding phonological planning (based on reaction time data) are not affected by this error.

The acoustic data for our study included vowel durations and measures of vowel space size. To reduce collinearity, analysis of vowel space sizes included a factor residualizing within-grammatical-category neighborhood density on vowel duration. This residualized factor was accidentally utilized in the analyses of vowel duration as well. Re-analysis with the non-residualized measure of neighborhood density reveal that there was no significant main effect of neighborhood density on vowel duration ($\beta = -0.011$, $SE = 0.007$, $\chi^2(1) = 2.13$, $p = 0.14$) nor an interaction of density by production context (sentence context vs. bare picture naming; $\beta = 0.0002$, $SE = 0.001$, $\chi^2(1) = 0.03$, $p = 0.86$). Thus, based on these data we do not have strong evidence that grammatical encoding constrains phonetic processing. Our reanalysis also revealed significant relationship between reaction times and vowel durations (such that trials with longer reaction times had longer vowels; $\beta = 0.0439$, $SE = 0.0163$, $\chi^2(1) = 7.13$, $p = 0.0076$). This suggests that the phonetic variation in these data is related to variation in phonological planning times.
References

Grammatical constraints on phonological encoding in speech production

Jordana R. Heller and Matthew Goldrick

Department of Linguistics
Northwestern University
2016 Sheridan Rd
Evanston, IL 60208
jordana@u.northwestern.edu, matt-goldrick@northwestern.edu
Abstract

To better understand the influence of grammatical encoding on the retrieval and encoding of phonological word-form information during speech production, we examine how grammatical class constraints influence the activation of phonological neighbors (words phonologically related to the target, e.g., moon, two for target TUNE). Specifically, we compare how neighbors that share a target’s grammatical category (here, nouns) influence its planning and retrieval, assessed by picture naming latencies, and phonetic encoding, assessed by vowel productions in picture names, when grammatical constraints are strong (in sentence contexts) vs. weak (bare naming). Within-category (noun) neighbors influenced planning time and phonetic encoding more strongly in sentence contexts. This suggests that grammatical encoding constrains phonological processing; the influence of phonological neighbors is grammatically dependent. Moreover, effects on planning times could not fully account for phonetic effects, suggesting phonological interaction affects articulation after speech onset. These results support production theories integrating grammatical, phonological, and phonetic processes.

Keywords: speech production; psycholinguistics; phonology
Grammatical Constraints on Phonological Encoding in Speech Production

1. Introduction

Sending words from idea to articulation is not a simple task. For example, when planning the sentence *The bird sings its tune*, *BIRD*, *SINGS*, and *TUNE* are activated simultaneously at the message level, although they must be produced serially (Dell, Oppenheim, & Kittredge, 2008). Competition for production of each word comes from other words in the sentence context (between *BIRD* and *SINGS* or *TUNE*) as well as other candidate words in the lexicon (e.g., *MOON* and *SOON* for target *TUNE*; Dell et al., 2008).

Grammatical constraints—which can limit the number of candidate structures and lexical items considered—may mitigate this difficulty. For example, during phrase production, semantically related distractor words that do not share a target’s grammatical category (e.g., noun vs. verb) affect target processing less than related same-category distractors (Alario, Matos, & Segui, 2004; Pechmann & Zerbst, 2002; Vigliocco, Vinson, & Siri, 2005). However, how grammatical constraints affect phonological processing remains underexplored.

Past research suggests that grammatical and phonological information mutually constrain which word speakers select for production. Lexical speech errors that are phonologically related to the intended target tend to match its grammatical category (see Goldrick, Folk, & Rapp, 2010, for a review)—an effect that is stronger in sentential contexts (Berndt, Mitchum, Haendiges, & Sandson, 1997). In non-errorful speech, phonological similarity between words in sentences can affect word choice (Jaeger, Furth, & Hilliard, 2012) and grammatical structure (Bock, 1987).
Here we examine whether grammatical-phonological interactions are limited to lexical selection, or whether grammatical constraints also condition changes in phonological activation. We index phonological activation by the effects of phonological neighbors (words in the lexicon differing by the addition, substitution or deletion of one phoneme) on target processing. Our results show that in production contexts where grammatical constraints strongly influence processing (in noun phrases in sentences), neighbors sharing the target’s grammatical category more strongly influence processing (relative to productions in the absence of a syntactic context). This suggests grammatical class strongly influences phonological activation in syntactic contexts, supporting interactive processing models.

1.1 Phonological neighbors in speech production

The number of phonological neighbors a word has (its *neighborhood density*) is correlated with multiple aspects of its processing, and can have either inhibitory or facilitatory effects. High neighborhood density has been associated with slower (Gordon & Kurczek, in press; Sadat, Martin, Costa, & Alario, 2014) or faster reaction times (Vitevitch, 2002), greater (Goldrick et al., 2010; Gordon, 2002; Harley & Bown, 1998; Vitevitch, 1997, 2002; Vitevitch & Sommers, 2003) or diminished (Newman & German, 2005) naming accuracy, and more extreme (i.e., expanded F1-F2 vowel space; Munson, 2007; Munson & Solomon, 2004; Wright, 2004) or less extreme (shorter vowel durations and contracted vowel space; Gahl, Yao, & Johnson, 2012) phonetic properties.

One approach to this range of effects incorporates a spreading-activation production model that links word-level (here, *L-level*) and phonemic representations. In the case of facilitatory effects of neighbors, activation-feedback resonance between these
two levels increases the activation of phonemes that neighbors share (Vitevitch, 2002). For example, the L-level representation *TUNE* activates phonemes /t/, /u/, and /n/, which send limited activation to neighbors *MOON, SOON*, etc., which again boost activation of shared phonemes. The greater the neighborhood density, the more reactivation shared phonemes receive. The higher the phonemes’ activation, the swifter their selection (Vitevitch, 2002) and the more extreme their phonetic manifestation (Baese-Berk & Goldrick, 2009). In contrast, inhibitory effects arise when neighbors are strongly activated, allowing lateral inhibition between lexical representations to dominate over activation-feedback resonance (Chen & Mirman, 2012). Differences in the relative activation of neighbors could arise due to a number of factors (e.g., production task).

**1.2 Phonological interaction in sentential contexts**

Does grammatical processing affect the influence of phonological neighbors? We outline two ways that grammatical constraints—which specify constraints on L-level representations—could interact with phonological processes, yielding contrasting predictions.

**1.2.1 Syntax-dependent phonological activation.**

To account for the mutual influence of grammatical and phonological information on lexical selection in syntactic contexts, we could incorporate a mechanism that, in such contexts (where grammatical constraints are strong), either inhibits L-level representations that do not match the category of the target, or, alternatively, boosts the activation of representations that match its category (Dell, Burger, & Svec, 1997; Dell et al., 2008; Gordon & Dell, 2003). This mechanism would serve to modulate the phonologically-
driven activation outlined above, allowing grammatical and phonological information to mutually influence L-level selection.

This account makes a novel prediction regarding the influence of phonological neighbors: in syntactic contexts, primarily same-category neighbors should affect phonological encoding (as reflected by facilitatory or inhibitory effects of within-category neighborhood density). During the production of TUNE in the example The bird sings its tune, neighbors such as MOON, which share TUNE’s grammatical category, would affect its activation. However, different-category neighbors such as SOON would contribute far less. In the absence of a syntactic context, however, where grammatical constraints are low, all neighbors may contribute. Thus, reaction times and phonetic outcomes would be better predicted by a target’s within-category neighborhood density in sentence contexts, but by its total neighborhood density in the absence of context. Words matched in overall but differing in within-category density should therefore behave similarly in the absence of a grammatically constraining context, but differ from one another in its presence.

1.2.2 Syntax-independent phonological activation.

Alternatively, grammatical constraints could solely influence which L-level representation is selected for production. This is consistent with production theories incorporating monitoring mechanisms that verify the properties of selected representations (see Hartsuiker, 2006, for a review). Because phonological activation influences L-level representations, such a mechanism would allow grammatical and phonological information to mutually influence L-level selection without constraining which representations become activated. Under this model, all phonological neighbors would contribute to target phonological activation—regardless of context. Therefore,
words matched in overall but differing in within-category density should exhibit similar behavior both in the presence and absence of a syntactic context.

1.3 Current Experiment

To test these predictions, we compare production of words with relatively few vs. many within-category phonological neighbors (but matched for overall number of phonological neighbors) across two conditions. Using a paradigm adapted from Griffin and Bock (1998), participants named pictures in the absence or presence of a syntactically (but not thematically) constraining sentence frame. Naming latencies and phonetic properties of vowels were used to examine how within-category density and the presence of a syntactically constraining sentence frame affect retrieval and planning vs. phonetic processes.

2. Methods

2.1 Participants

Sixty-four native English speakers (55 female) aged 18-34 with no history of speech or language deficits participated in exchange for course credit or $10.

2.2 Materials

Target words were 24 English CVC nouns (see supplementary materials), divided into two groups differing in within-category (noun) phonological neighborhood density and frequency-weighted noun phonological neighborhood density. Across groups, target words were matched for total and frequency-weighted phonological neighborhood density, vowel, onset and coda voicing, word frequency, contextual diversity, imageability, and sum uniphone and biphone phonotactic probability. Vowels included /i, e, æ, a, o, u/. Picture
name agreement for illustrations representing targets (see supplementary materials) was also matched across stimulus groups.

Four non-thematically predictive carrier sentences were designed for each target word, creating 96 target sentences (see supplementary materials). Targets appeared in sentence-final position preceded by a definite determiner (e.g., *He wanted a kick at the PEN*). Sentences were divided into four lists on which each target appeared once. These were integrated with 27 filler sentences of various syntactic structures, for which filler pictures were completions. Two pseudorandomized orderings were created for each list.

### 2.3 Procedure

Participants were first familiarized with the pictures, and then participated in both conditions: picture naming in the absence of a context (bare naming) and picture naming as sentence completion (sentence context). Condition order was counterbalanced.

During bare naming, participants were asked to name aloud pictures that appeared on the screen as quickly and accurately as possible. Participants pressed the space bar to begin each trial. A fixation cross appeared at the center of the screen for 500 ms, followed by a picture, which remained until participants pressed the space bar to move on.

In the sentence context condition, participants were asked to read aloud a sentence that appeared in the center of the screen one word at a time, and to name the picture that appeared to complete the sentence as quickly and accurately as possible. Trials began with a fixation cross (500 ms), followed by the 500 ms presentation of each word in a sentence individually in the center of the screen. In place of the sentence’s final word, its picture

---

1 Two of these sentences contained a neighbor of the target in an adjacent phrase (e.g., *He chose to look at the LOCK*). These had no detectable effect on target processing, consistent with previous work showing limited phonological effects beyond the phrase (Smith & Wheeldon, 2004; see supplementary materials for details).
appeared and remained until participants pressed the space bar. Three practice trials were followed by one of the eight sentence lists. Participants named targets once in each condition. Six comprehension questions were interspersed to encourage holistic sentence processing.

2.4 Speech Analysis

Trials with errors or disfluencies were excluded from all analyses. Two items with high error rates were excluded from all analyses; their vowel-matched counterparts were additionally excluded from vowel analyses.

2.4.1 Speech onset latencies

Reaction times (RTs) were calculated from picture presentation to speech onset. Speech onsets were detected automatically in Praat (Boersma & Weenink, 2012) using intensity thresholds, and were hand corrected.

2.4.2 Vowel duration

Vowel duration was hand-measured in Praat. Vowel onsets and offsets were marked using cues from the waveform and spectrogram (see supplementary materials). A second coder marked 25% of the data to assess measurement reliability. Measurements were well correlated ($r(627) = 0.84, p < 0.0001$).

2.4.3 Vowel space size

Measurements (in Hz) of the first and second formant of each vowel were taken from the point of maximal formant displacement (Wright, 2004) using Burg LPC automatic formant detection implemented in Praat. Vowel space size was measured using Euclidean distance in F1-F2 space by calculating the average distance of each token produced by a participant from every token produced by that participant in another vowel category.
3. Results

3.1 Speech onset latencies (RTs)

3.1.1 Analysis

Exclusions of errors and RTs shorter than 300 ms resulted in removal of 43 observations (1.5% of the data). RTs were log transformed to normalize their distribution. Linear mixed effects model-based outlier trimming further removed RTs with residual errors more than 2.5 standard deviations from the mean (Baayen, 2008), excluding 60 additional observations (2.5% of the data).

A linear mixed effects regression analysis was performed on the remaining RTs. Fixed effects of interest included noun neighborhood density, contrast-coded condition (bare naming or sentence context), and their interaction. To control for possible order effects, contrast-coded task order (block) was included, as well as its two- and three-way interactions with noun density and condition. The maximal random effects structure supported by the design (Barr, Levy, Scheepers, & Tily, 2013) included random intercepts for participant and word, as well as random slopes for noun density and condition by participant, and for condition by word. Interactions were excluded from random effect structure in order to obtain model convergence. Significance was determined by nested model comparison (Barr et al., 2013).

3.1.2 Results

Results for the effects of interest are summarized in Figure 1. Pictures were named faster in the sentence context condition than in bare naming ($\beta = -0.099$, SE = 0.019, $\chi^2(1) = 23.49$, $p < 0.0001$). No main effect of noun density was detected ($\chi^2(1) < 1$), but there was a

---

2 Similar results were found for the subset used in the vowel analyses.
significant interaction of noun density and condition ($\beta = 0.0038, SE = 0.0017, \chi^2(1) = 4.79, p = 0.029$). The effect of noun density was larger in sentence contexts than bare naming; in sentence contexts, words with higher noun density elicited significantly longer RTs than words with lower noun density.

RTs were shorter in the second block ($\beta = -0.055, SE = 0.017, \chi^2(1) = 10.374, p = 0.0013$). None of the interactions with block reached significance ($\chi^2s(1) < 2; ps > 0.1$).

![Figure 1](image.png)

**Figure 1.** RTs (latencies to speech onset) for words with many and few noun neighbors in bare naming and sentence contexts. Error bars indicate by-participant standard error.

### 3.2 Vowel duration

#### 3.2.1 Analysis

Vowel durations were log transformed. Exclusion of errors and model-based outliers removed 116 observations (4.5% of the data).

A linear mixed effects regression was performed. Again, fixed effects of interest included noun density, contrast-coded condition, and their interaction. Vowel identity and
block were included as contrast-coded control factors. Because interactions with block did not significantly predict RT, block was included here only as a main effect. Random effects were identical to those above.

3.2.2 Results

Figure 2 shows the fixed effects of interest for vowel duration. Items with higher noun density had significantly shorter vowel durations ($\beta = 0.13, \text{SE} = 0.00062, \chi^2(1) = 611.56, p < 0.0001$). Again, the effect of noun density was significantly larger in the sentence context condition than the bare naming condition ($\beta = 0.0037, \text{SE} = 0.00025, \chi^2(1) = 78.60, p < 0.0001$). Vowel ($\chi^2(5) = 2.66; p > 0.1$) and other factors ($\chi^2s(1) < 1; ps > 0.1$) were not significant predictors.

![Figure 2](image-url)

**Figure 2.** Vowel durations for words with many and few noun neighbors in bare naming and sentence contexts. Error bars indicate by-participant standard error.
3.2.3 Relationship between speech onset latency and vowel duration

To assess whether the observed effects on vowel duration could be attributed solely to differences in speech planning time, an additional regression was performed on vowel durations identical to their analysis above, with an additional fixed effect for log RT and random slope for RT by participant\(^3\). While RT was a marginally significant predictor of vowel duration ($\beta = 0.010$, $SE = 0.0058$, $\chi^2(1) = 3.10$, $p = 0.0782$), the effect of noun density ($\beta = 0.12$, $SE = 0.00081$, $\chi^2(1) = 573.66$, $p < 0.0001$) and its interaction with condition ($\beta = 0.011$, $SE = 0.00054$, $\chi^2(1) = 65.77$, $p < 0.0001$) remained significant.

3.3 Vowel space size

3.3.1 Analysis

Log-transformed vowel space size was analyzed using a linear mixed effects regression including the same factors as the duration analysis above, with additional fixed effects to control for effects of speaker gender (Simpson, 2009) and log-transformed vowel duration (Moon & Lindblom, 1994). To avoid collinearity, noun density was residualized against log-transformed vowel duration. Removal of errors and outliers excluded 135 observations (5% of the data).

3.3.2 Results

As vowel duration increased, vowel space size increased ($\beta = 0.11$, $SE = 0.036$, $\chi^2(1) = 7.83$, $p = 0.0051$). No effects of noun density, condition, or their interaction were detected ($\chi^2s(1) < 2.7$; $ps > 0.10$). Significant control predictors included vowel ($\chi^2(5) = 17.24$, $p < 0.0001$) and gender (females’ larger than males’; $\beta = 0.087$, $SE = 0.019$, $\chi^2(1) = 17.24$, $p < 0.0001$), but not block ($\chi^2(1) < 1$; $p > 0.1$).

\(^3\) This analysis relies on the assumption that RT and vowel length are linearly related. Visual data inspection did not suggest any other relationship.
4. Discussion

4.1 Syntax-dependent phonological processing

While previous research has suggested syntactic information constrains lexical access, it is unclear whether such constraints limit phonological activation or simply lexical selection. In order to adjudicate between these mechanisms, we compared how within-category (noun) phonological neighborhood density influenced speech planning and phonetic encoding when grammatical constraints were strong vs. weak. Measures of both planning and phonetic encoding support syntax-dependent phonological processing during sentence production. Controlling for overall neighborhood density, the effect of within-category density was stronger in the sentence context condition than in bare naming for both reaction times and vowel durations. This is consistent with a model in which different-category L-level representations are inhibited (or same-category representations are boosted) more strongly when grammatical constraints are strong.

These results extend previous work supporting grammatical influences on phonological processes in production. Janssen and Caramazza (2009) found that the effect of phonological similarity between contiguous words in a phrase is influenced by word order (e.g., facilitation in adjective-noun phrases, but no effect in noun-adjective phrases). Our results reveal not only how grammatical constraints influence both the processing of words present within the sentence context and the implicit activation of phonological neighbors in the lexicon, but also provide novel insight into one possible functional mechanism at this interface.

Interestingly, similar results have recently been reported in speech perception, where grammatical cues reduce competition from phonological neighbors outside of the
expected grammatical category (Strand, Simenstad, Cooperman, & Rowe, in press). A clear avenue for future work is to investigate whether these similar effects arise due to similar mechanisms.

4.2 Variable influences of neighborhood

Previous work has shown both facilitatory and inhibitory effects of neighborhood density. Here, we found the latter: higher within-category phonological neighborhood density was associated with slower RTs and shorter vowel durations (a primary predictor of vowel space size). These data are consistent with Gahl et al. (2012): In spontaneous speech (crucially, in sentential contexts), high overall neighborhood density was associated with shorter vowel durations and smaller vowel spaces.

To account for these results, we build on Chen and Mirman’s (2012) analysis showing that in interactive processing models weakly activated neighbors can cause net target facilitation, whereas strongly activated neighbors create net target inhibition. We hypothesize that when grammatical constraints are weak, as in single word naming, neighbors may be weakly activated, but when the candidate set of neighbors is reduced by increasing grammatical constraints, the relative increase in these neighbors’ activation may lead to target inhibition. It is probable that total and within-category phonological neighborhood density were correlated in the Gahl et al. (2012) dataset. Further research will clarify the interaction between these factors.

4.3 Neighborhood effects in planning and phonetic outcomes

RTs were weakly related to vowel durations; however, noun density-related factors affected vowel durations beyond what RT predicted. This suggests phonetic outcomes are not simply a by-product of planning times (c.f. Kirov & Wilson, 2013). Such effects are more
consistent with interactive models in which lexical-phonological interaction continues to affect phonetic outcomes as speech execution begins.

5. Conclusion

These results suggest that in contexts with strong grammatical constraints, phonological interaction during speech production is constrained by grammatical category. These task-dependent changes in activation throughout the production system affect not only planning, as reflected in differential naming latencies, but also phonetic encoding, as reflected in vowel durations. These results support theories integrating grammatical, lexical, phonological, and phonetic processes during sentence production.
Acknowledgments

We gratefully acknowledge support from the National Science Foundation (Grant BCS0846147 awarded to MG) and Northwestern University’s Research Grants Committee (Graduate Research Grant awarded to JRH). Thanks to Janice Johnson and Danika Paskvan for assistance running participants, Ariana Steele for assistance with acoustic reliability measurements, and the Northwestern Sound Lab for helpful comments.
References


Supplementary Materials

Grammatical constraints on phonological encoding in speech production

Jordana R. Heller and Matthew Goldrick

S1. Target word properties

Properties for all 24 target words are listed in Table S1. Noun phonological neighborhood density (ND), frequency-weighted ND, total ND, and total frequency-weighted ND were calculated from the part-of-speech-annotated SUBTLEX-US corpus of American English subtitles (Brysbaert & New, 2009; Brysbaert, New, & Keuleers, 2012). Noun ND was calculated as the number of phonological neighbors that were used as nouns in the SUBTLEX-US corpus. Frequency-weighted noun ND was calculated as the summed frequency with which a target and its neighbors were used as nouns in the corpus. Target frequency per million and contextual diversity were taken from the same corpus. Imageability data was taken from norms collected by Cortese and Fugett (2004). Picture name agreement for items used in this study was collected in a separate norming study (see supplementary section S2). Voicing of word onset and coda consonants, which are known to affect vowel duration (Peterson & Lehiste, 1960), are included. If a segment is voiced, it is listed with a value of 1 in the table; if it is voiceless, it is listed with a value of 0. Sum uniphone and biphone phonotactic probabilities were taken from the Irvine Phonotactic Online Database (IPhOD; Vaden, Halpin, & Hickok, 2009).

Due to high error rates, two items (POT, PAN) were excluded from the RT analyses. Their vowel-matched counterparts (LOCK, HAT) were additionally excluded from vowel analyses. Details of group averages for these subsets of the stimuli are listed in Table S2.
Table S1. Details of stimulus properties for high and low noun phonological neighborhood density (ND) groups.

<table>
<thead>
<tr>
<th>Target</th>
<th>Vowel</th>
<th>Noun ND</th>
<th>Frequency -weighted</th>
<th>Total ND</th>
<th>Frequency -weighted</th>
<th>Total ND</th>
<th>Frequency (per million)</th>
<th>Contextual Diversity</th>
<th>Imagery</th>
<th>Name Agreement</th>
<th>Onset Voice</th>
<th>Coda Voice</th>
<th>Sum Uniphone Phonotactic Probability</th>
<th>Sum Biphone Phonotactic Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low Noun Density Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hat†</td>
<td>æ</td>
<td>34</td>
<td>72.16</td>
<td>37</td>
<td>112.92</td>
<td>64.18</td>
<td>1728</td>
<td>6.7</td>
<td>0.77</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0778</td>
<td>0.0015</td>
</tr>
<tr>
<td>rat</td>
<td>æ</td>
<td>31</td>
<td>68.84</td>
<td>37</td>
<td>98.24</td>
<td>32.61</td>
<td>941</td>
<td>6.8</td>
<td>0.85</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0730</td>
<td>0.0031</td>
</tr>
<tr>
<td>doll</td>
<td>a</td>
<td>16</td>
<td>32.78</td>
<td>17</td>
<td>44.99</td>
<td>24.76</td>
<td>648</td>
<td>6.4</td>
<td>0.77</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.0644</td>
<td>0.0015</td>
</tr>
<tr>
<td>pot*</td>
<td>a</td>
<td>32</td>
<td>67.91</td>
<td>35</td>
<td>98.36</td>
<td>22.53</td>
<td>763</td>
<td>6.2</td>
<td>0.85</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0661</td>
<td>0.0014</td>
</tr>
<tr>
<td>head</td>
<td>e</td>
<td>20</td>
<td>44.49</td>
<td>27</td>
<td>81.18</td>
<td>371.51</td>
<td>6294</td>
<td>6.8</td>
<td>0.77</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0622</td>
<td>0.0014</td>
</tr>
<tr>
<td>bed</td>
<td>e</td>
<td>20</td>
<td>42.13</td>
<td>29</td>
<td>84.20</td>
<td>187.12</td>
<td>4113</td>
<td>6.6</td>
<td>1.00</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.0664</td>
<td>0.0014</td>
</tr>
<tr>
<td>beak</td>
<td>i</td>
<td>24</td>
<td>52.21</td>
<td>29</td>
<td>71.96</td>
<td>2.10</td>
<td>88</td>
<td>5.2</td>
<td>0.96</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0689</td>
<td>0.0012</td>
</tr>
<tr>
<td>seal</td>
<td>i</td>
<td>30</td>
<td>66.93</td>
<td>35</td>
<td>92.77</td>
<td>14.75</td>
<td>472</td>
<td>5.5</td>
<td>0.88</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.0678</td>
<td>0.0019</td>
</tr>
<tr>
<td>nose</td>
<td>o</td>
<td>14</td>
<td>27.89</td>
<td>17</td>
<td>49.09</td>
<td>69.75</td>
<td>2125</td>
<td>6.4</td>
<td>1.00</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.0566</td>
<td>0.0010</td>
</tr>
<tr>
<td>boat</td>
<td>o</td>
<td>26</td>
<td>59.76</td>
<td>28</td>
<td>81.67</td>
<td>95.78</td>
<td>1457</td>
<td>6.7</td>
<td>0.81</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0691</td>
<td>0.0010</td>
</tr>
<tr>
<td>goose</td>
<td>u</td>
<td>13</td>
<td>31.11</td>
<td>14</td>
<td>36.06</td>
<td>13.04</td>
<td>376</td>
<td>6.9</td>
<td>0.92</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0554</td>
<td>0.0005</td>
</tr>
<tr>
<td>moose</td>
<td>u</td>
<td>17</td>
<td>46.89</td>
<td>18</td>
<td>53.03</td>
<td>5.53</td>
<td>126</td>
<td>6.7</td>
<td>0.96</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0588</td>
<td>0.0006</td>
</tr>
<tr>
<td><strong>High Noun Density Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pan*</td>
<td>æ</td>
<td>27</td>
<td>69.09</td>
<td>30</td>
<td>91.11</td>
<td>12.29</td>
<td>366</td>
<td>6.3</td>
<td>0.88</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.0730</td>
<td>0.0043</td>
</tr>
<tr>
<td>tack</td>
<td>æ</td>
<td>35</td>
<td>84.84</td>
<td>36</td>
<td>100.56</td>
<td>2.12</td>
<td>73</td>
<td>5.3</td>
<td>0.31</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0689</td>
<td>0.0025</td>
</tr>
<tr>
<td>bomb</td>
<td>a</td>
<td>16</td>
<td>37.09</td>
<td>19</td>
<td>51.33</td>
<td>53.65</td>
<td>934</td>
<td>6.4</td>
<td>1.00</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.0596</td>
<td>0.0013</td>
</tr>
<tr>
<td>lock†</td>
<td>a</td>
<td>31</td>
<td>75.78</td>
<td>32</td>
<td>88.88</td>
<td>56.57</td>
<td>1866</td>
<td>5.6</td>
<td>0.92</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0634</td>
<td>0.0015</td>
</tr>
<tr>
<td>bell</td>
<td>e</td>
<td>26</td>
<td>62.30</td>
<td>28</td>
<td>87.32</td>
<td>39.33</td>
<td>1035</td>
<td>6.6</td>
<td>1.00</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.0667</td>
<td>0.0019</td>
</tr>
<tr>
<td>pen</td>
<td>e</td>
<td>26</td>
<td>52.97</td>
<td>27</td>
<td>76.85</td>
<td>24.73</td>
<td>816</td>
<td>6.6</td>
<td>1.00</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.0655</td>
<td>0.0045</td>
</tr>
<tr>
<td>heel</td>
<td>i</td>
<td>28</td>
<td>63.84</td>
<td>29</td>
<td>85.33</td>
<td>7.39</td>
<td>281</td>
<td>5.8</td>
<td>0.96</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.0650</td>
<td>0.0008</td>
</tr>
<tr>
<td>peel</td>
<td>i</td>
<td>31</td>
<td>70.10</td>
<td>32</td>
<td>85.41</td>
<td>5.35</td>
<td>189</td>
<td>5.4</td>
<td>0.65</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.0650</td>
<td>0.0013</td>
</tr>
<tr>
<td>coat</td>
<td>o</td>
<td>27</td>
<td>62.51</td>
<td>28</td>
<td>75.86</td>
<td>42.08</td>
<td>1272</td>
<td>6.7</td>
<td>0.62</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0662</td>
<td>0.0012</td>
</tr>
<tr>
<td>bowl</td>
<td>o</td>
<td>26</td>
<td>64.02</td>
<td>28</td>
<td>76.81</td>
<td>21.45</td>
<td>682</td>
<td>6.3</td>
<td>1.00</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.0661</td>
<td>0.0012</td>
</tr>
<tr>
<td>moon</td>
<td>u</td>
<td>21</td>
<td>51.35</td>
<td>23</td>
<td>66.06</td>
<td>49.96</td>
<td>1167</td>
<td>6.8</td>
<td>0.77</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.0636</td>
<td>0.0006</td>
</tr>
<tr>
<td>boot</td>
<td>u</td>
<td>28</td>
<td>69.45</td>
<td>30</td>
<td>84.57</td>
<td>11.14</td>
<td>374</td>
<td>6.9</td>
<td>0.96</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0678</td>
<td>0.0009</td>
</tr>
</tbody>
</table>

Low Average: 23.08 51.09 26.92 75.37 75.31 1594.25 6.41 0.88 0.67 0.42 0.0651 0.0014
High Average: 26.83 63.61 28.50 80.84 27.17 754.58 6.23 0.84 0.50 0.67 0.0659 0.0018

p-value (paired t-test, df = 11): 0.04 < 0.001 0.38 0.29 0.14 0.14 0.26 0.60 0.63 0.21

* Item excluded from RT analysis due to high error rate. † Vowel-matched item excluded from vowel analyses.
Table S2. Group summary statistics for stimulus subsets analyzed

<table>
<thead>
<tr>
<th></th>
<th>Noun ND</th>
<th>Frequency -weighted Noun ND</th>
<th>Total ND</th>
<th>Frequency -weighted Total ND</th>
<th>Frequency (per million)</th>
<th>Contextual Diversity</th>
<th>Imageability</th>
<th>Name Agreement</th>
<th>Onset Voice</th>
<th>Coda Voice</th>
<th>Sum Uniphone Phonotactic Probability</th>
<th>Sum Biphone Phonotactic Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group summary statistics excluding POT, PAN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Average</td>
<td>22.27</td>
<td>49.56</td>
<td>26.18</td>
<td>73.28</td>
<td>80.10</td>
<td>1669.82</td>
<td>6.43</td>
<td>0.88</td>
<td>0.73</td>
<td>0.45</td>
<td>0.0014</td>
<td>0.0650</td>
</tr>
<tr>
<td>High Average</td>
<td>26.82</td>
<td>63.11</td>
<td>28.36</td>
<td>79.91</td>
<td>28.52</td>
<td>789.91</td>
<td>6.22</td>
<td>0.84</td>
<td>0.55</td>
<td>0.64</td>
<td>0.0016</td>
<td>0.0653</td>
</tr>
<tr>
<td>p-value (paired t-test, df = 10)</td>
<td>0.09</td>
<td>0.04</td>
<td>0.45</td>
<td>0.42</td>
<td>0.15</td>
<td>0.16</td>
<td>0.23</td>
<td>0.54</td>
<td></td>
<td></td>
<td>0.52</td>
<td>0.92</td>
</tr>
<tr>
<td>Group summary statistics excluding POT, PAN, LOCK, HAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Average</td>
<td>20.00</td>
<td>44.91</td>
<td>23.78</td>
<td>66.10</td>
<td>87.15</td>
<td>1744.33</td>
<td>6.36</td>
<td>0.90</td>
<td>0.78</td>
<td>0.55</td>
<td>0.0012</td>
<td>0.0633</td>
</tr>
<tr>
<td>High Average</td>
<td>26.40</td>
<td>61.85</td>
<td>28.00</td>
<td>79.01</td>
<td>25.72</td>
<td>682.30</td>
<td>6.28</td>
<td>0.83</td>
<td>0.50</td>
<td>0.70</td>
<td>0.0016</td>
<td>0.0654</td>
</tr>
<tr>
<td>p-value (paired t-test, df = 9)</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>0.14</td>
<td>0.07</td>
<td>0.15</td>
<td>0.13</td>
<td>0.51</td>
<td>0.45</td>
<td></td>
<td></td>
<td>0.44</td>
<td>0.54</td>
</tr>
</tbody>
</table>
S2. Picture name agreement

Name agreement for pictures used in this study was assessed separately in an online norming study using a free-response paradigm.

**S2.1 Participants**

Twenty-six Northwestern University undergraduates who were native speakers of American English participated in exchange for partial course credit.

**S2.2 Materials**

Materials consisted of colored illustrations representing 37 potential experimental items (CVC nouns) and 19 multisyllabic fillers. Pictures were selected from a database of colorized Snodgrass and Vanderwart (1980) pictures (Rossion & Pourtois, 2004) when available. Otherwise, illustrations were selected from public domain clip art sources. Some of these were edited digitally by the first author.

**S2.3 Procedure**

Participants viewed pictures individually on a computer screen and were asked to type the best name for the picture into a text box below the picture. Participants saw each picture once.

**S2.4 Analysis and results**

Naming agreement was defined for each item as the fraction of name responses matching the authors’ intended target name. This was calculated by summing the number of participants who typed the intended target name and dividing this sum by the total number of participants. Thus, if all participants responded to a picture with the intended target name (as was the case for, e.g., nose, Table S1), name agreement had a value of 1. Results for items selected for use in the current study can be seen in Table S1.
S3. Sentence construction and norming

S3.1 Sentence construction

Four carrier sentences were designed for each of the 24 target nouns. These stimuli consisted of written sentences in which the last word in the sentence—the target word—was represented by a picture rather than text, following Griffin and Bock (1998). Sentences were designed not to be thematically predictive of the target word. However, all sentences contained a syntactic cue in the form of a definite determiner preceding the target. The thematic predictiveness of each target sentence was normed in a separate study (see supplementary section S3.2).

The four sentences for each target were minimally different, developed in a 2 (Prime Phonological Relation) x 2 (Prime Category Match) design: the sentence either contained or did not contain an onset-matched phonological neighbor of the target word (cohort prime), and the structure of the sentence varied minimally to allow words in this prime position to behave as either nouns (grammatical category match to the target) or verbs (grammatical category mismatch). An example set of four sentences for one target are given in 1a-d—the full set of sentences are given in Table S3, below. Phonologically related primes were chosen whose meanings—and frequencies of occurrence—as nouns and verbs were matched as closely as possible (Brysbaert, New, & Keuleers, 2012). Unrelated primes were selected to preserve the meaning of the phonologically related prime as closely as possible. Main verbs (e.g., chose) were selected to take noun and verb complements (e.g., chose a look vs. chose to look) with frequencies that were as well-matched as possible (Roland, Dick, & Elman, 2007).
1a) **RELATED/MATCH**  He chose a look at the *LOCK*.

1b) **RELATED/MISMATCH**  He chose to look at the *LOCK*.

1c) **UNRELATED/MATCH**  He chose a glance at the *LOCK*.

1d) **UNRELATED/MISMATCH**  He chose to glance at the *LOCK*.

<table>
<thead>
<tr>
<th>Target</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>beak</td>
<td>He started the/to bike/ride near the beak.</td>
</tr>
<tr>
<td>bed</td>
<td>He announced a/to bet/wager on the bed.</td>
</tr>
<tr>
<td>bell</td>
<td>She expected the/to bail/pardon at the bell.</td>
</tr>
<tr>
<td>boat</td>
<td>He hurried a/to bite/snack on the boat.</td>
</tr>
<tr>
<td>bomb</td>
<td>She promised a/to boss/captain near the bomb.</td>
</tr>
<tr>
<td>boot</td>
<td>She attempted a/to bat/punch at the boot.</td>
</tr>
<tr>
<td>bowl</td>
<td>He regretted the/to bill/charge for the bowl.</td>
</tr>
<tr>
<td>coat</td>
<td>She refused a/to comb/brush near her coat.</td>
</tr>
<tr>
<td>doll</td>
<td>She proposed a/to deal/trade with the doll.</td>
</tr>
<tr>
<td>goose</td>
<td>He tried the/to goof/joke about the goose.</td>
</tr>
<tr>
<td>hat†</td>
<td>She suggested a/to hit/punch through the hat.</td>
</tr>
<tr>
<td>head</td>
<td>He learned a/to hem/stitch for the head.</td>
</tr>
<tr>
<td>heel</td>
<td>She ordered the/to haul/load with her heel.</td>
</tr>
<tr>
<td>lock†</td>
<td>He chose a/to look/peek at the lock.</td>
</tr>
<tr>
<td>moon</td>
<td>He remembered a/to moan/sigh at the moon.</td>
</tr>
<tr>
<td>moose</td>
<td>She continued her/to move/run by the moose.</td>
</tr>
<tr>
<td>nose</td>
<td>He indicated the/to note/comment with his nose.</td>
</tr>
<tr>
<td>pan*</td>
<td>She rushed a/to pat/bang on the pan.</td>
</tr>
<tr>
<td>peel</td>
<td>He confessed a/to peep/glance at the peel.</td>
</tr>
<tr>
<td>pen</td>
<td>She wanted a/to peck/kick at the pen.</td>
</tr>
<tr>
<td>pot*</td>
<td>She began a/to pout/frown by the pot.</td>
</tr>
<tr>
<td>rat</td>
<td>She approved the/to route/walk by the rat.</td>
</tr>
<tr>
<td>seal</td>
<td>She liked the/to sail/cruise by the seal.</td>
</tr>
<tr>
<td>tack</td>
<td>He noticed a/to tap/click on the tack.</td>
</tr>
</tbody>
</table>

* Items excluded from RT analysis due to high error rate. † Vowel-matched items excluded from vowel analyses.

Because neither the prime phonological relation nor prime category match had an effect on RTs in the current data (consistent with past work showing limited phonological priming beyond the phrase, e.g., Smith & Wheeldon, 2004), these manipulations were
excluded from analysis, and data is presented collapsing across sentence type (RT condition means are provided in Table S4).

**Table S4. RT condition means for prime conditions**

<table>
<thead>
<tr>
<th></th>
<th>Phonologically Related</th>
<th>Phonologically Unrelated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun Prime</td>
<td>564 ms</td>
<td>557 ms</td>
</tr>
<tr>
<td>Verb Prime</td>
<td>562 ms</td>
<td>557 ms</td>
</tr>
</tbody>
</table>

**S3.2 Sentence norming**

In order to ensure that targets were not semantically or thematically predictable from the sentence contexts, target sentences were normed in a separate study in which participants were asked to rate the likelihood of each target word given its sentential context on a seven-point Likert scale. These ratings were compared to ratings for key words in sentences rated in past studies as thematically predictable or unpredictable.

**S3.2.1 Participants**

Twenty-one Northwestern University undergraduates who were native speakers of American English participated in exchange for partial course credit.

**S3.2.2 Stimuli**

Target stimuli consisted of the sentences described above in section S3.1, which featured target words sentence-finally. Filler stimuli were 60 sentences of various structures taken from a past study of thematic predictability (Bradlow & Alexander, 2007), to be used as a comparison set. The sentence-final words in these filler sentences had been rated as having high or low predictability, given the sentence context. For example, a sentence with a high predictability sentence-final word is *He washed his hands with soap and water*. A sentence with a low predictability sentence-final word is *We talked about the*
water. Thirty high-predictability and 30 low-predictability sentences were included. All sentences were integrated in a pseudorandom order.

**S.3.2.3 Procedure**

Sentences appeared one at a time on a computer screen. Participants were asked to rate, on a scale of one to seven, how easy it would be to guess the sentence-final word, given the information in the sentence preceding that word. Before they completed the task, they were familiarized with high and low predictability examples, as above. Participants rated each sentence once.

**S.3.2.3 Analysis and results**

A summary of results can be seen in Figure S1. Target predictability in target sentences was rated on average 1.75 out of 7. Sentence-final words previously rated as unpredictable were rated on average 2.55/7. Those previously rated as predictable were rated 5.44/7. Differences between stimulus groups were assessed using a linear mixed effects regression analysis of participants’ predictability ratings, with contrast-coded group (Target, Low Predictability, High Predictability) as the sole fixed effect. Random effects included random intercepts for participant and item, as well as a random slope for group by participant. Significance of comparisons within the sole three-level fixed effect were assessed by assuming that the distribution of t-statistics follows that of z-statistics, rather than by model comparison. Targets in both low predictability ($\beta = -0.70, SE = 0.15, t = -4.61, p < 0.0001$) and target sentences ($\beta = -1.49, SE = 0.12, t = -12.43, p < 0.0001$) were rated significantly lower than those in high predictability sentences. In a follow-up regression, targets in target sentences were shown to have a significantly lower predictability scores than the low predictability set ($\beta = -0.70, SE = 0.15, t = -4.61, p < 0.0001$). Thus, target
words in target sentences were not semantically or thematically predictable from their sentence contexts.

![Figure S1](image.png)

**Figure S1.** Mean participant rating of sentence-final word predictability across stimulus group

**S4. Speech analysis: Details of vowel measurement**

Vowel duration was hand-measured in Praat. Vowel onsets and offsets were marked using cues from the waveform and spectrogram. Vowel boundaries were defined differently depending on the voicing and manner of the consonants adjacent to each vowel. For stop and fricative onset consonants, the vowel onset was marked at the onset of clear formant structure in the spectrogram, accompanied by clear periodicity in the waveform. Following nasal onset consonants, vowel onset was marked at the rise in formant amplitude in the spectrogram indicative of oral airflow. Following liquid onsets, the vowel onset was marked at the increase in formant amplitude and onset of the third formant, which is absent during most productions. For voiced and voiceless stop coda consonants, the vowel offset was marked at the moment of oral closure. This was indicated by a
reduction in waveform amplitude and disappearance of higher formant structure for voiced stop codas and by the absence of higher frequency noise for voiceless stop codas. For fricative codas, offsets were marked at the onset of frication noise in the waveform and spectrogram. For nasal and liquid codas, vowel offsets were marked at the sharp decrease in formant amplitude in the spectrogram.

References


