NORTHWESTERN UNIVERSITY

Contextual Constraints on Phonological Activation during Sentence Production

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Linguistics

By

Jordana Robin Heller

EVANSTON, ILLINOIS

August 2014

ABSTRACT

Contextual Constraints on Phonological Activation during Sentence Production

Jordana Robin Heller

A great deal of research has shown that during speech production many representations in addition to those of the current word in an utterance are also activated. Representations that are phonologically and/or semantically related to the current word (*phonological* and *semantic neighbors*, respectively), as well as other representations corresponding to other words planned in the utterance become concurrently active. The current project investigates the extent to which the predictability of a word from contextual cues affects this co-activation. Specifically, I investigate how grammatical constraints and meaning-related constraints from the utterance context affect the activation of phonological neighbors during phonological processing. The existence and nature of context-dependent phonological processing mechanisms is assessed in three speech production experiments performed in the laboratory, using an experimental paradigm in which participants name pictures in sentence and discourse contexts. In these investigations, both latency to speech onset and phonetic measures are analyzed, allowing comparison of the effects of contextual constraint on speech planning and phonetic processing. The relationship of context-dependent and –independent phonological processing mechanisms is further examined in an investigation of phonetic properties of words in spontaneous speech.

The results of the present investigations suggest that grammatical constraints, but not meaning-related constraints, affect phonological neighbor activation in sentence contexts. Phonological neighbors that share the grammatical category of a target word (*within-category neighbors*) influenced planning time and phonetic encoding more strongly in sentence contexts, supporting a grammatically dependent phonological processing mechanism. Meaning–related constraints independently affected target planning and phonetic encoding. However, meaning-related constraints did not reliably affect the activation of phonological neighbors over and above effects of grammatical constraints, suggesting a far smaller role for meaning-related constraint on phonological neighbor activation. In the present investigation of spontaneous speech, independent roles were found for grammatically dependent and -independent phonological processing mechanisms, suggesting both are at work during everyday speech processing. Implications of these results for theories of speech production are discussed.

ACKNOWLEDGMENTS

It could be said that this dissertation is a creation of my own design. This would be true in some ways, but in the deepest sense, it is not. Instead, its creation has been a community effort, and I have had the great fortune to work with and receive support from a great many wonderful people during its creation. I am so grateful to have had the opportunity to do this work and to create this manuscript, and I am deeply indebted to many who touched it along the way.

Many people contributed very concretely to this project. Antonia Agbeh, Janice Johnson, Danika Paskvan, and Ariana Steele helped with running participants in the lab. I owe particular thanks to Ariana Steele, who also measured many vowels. The corpus investigation discussed in Chapter 4 could not have been possible without the kindness of Yao Yao, Susanne Gahl, and Keith Johnson, who shared their dataset with me. I am also particularly indebted to the University Grants Committee at Northwestern, who awarded me a Graduate Research Grant, which helped to defray the costs of paying participants.

Thanks to everyone who provided feedback and support during my dissertation process. Sound Lab members all contributed helpful feedback during lab meetings—you're the reason that lab meeting is my favorite thing. Thanks also to the Syntax Lab. Many many thanks to my roommate and colleague, Mike Frazier, for many late-night discussions of life, the universe, and the finer details of linguistic theory—you are amazing. To Charlotte Vaughn, my good friend and long-time office mate, I owe many coffees—your support has meant so much to me. Thanks also to Melissa Baese-Berk, who always made me feel like finishing good work was possible even when the end didn't seem to be anywhere in sight. I also received helpful feedback from innumerable linguists and cognitive scientists at home at Northwestern and away at conferences. To those of you who made me think more deeply about my work, thank you.

To my family, who never doubted my ability to dissertate, I'm relieved to say you were right. What a delight it was especially to have my parents and brother come to a poster presentation of mine—you are all brilliant, I love you, and I appreciate your engagement. To my friends outside of the field, and especially to Chris, thank you for keeping me sane, grounded, and happy during this process. You may not know it, but I couldn't have done this without you. To my undergrad mentors, Jen Pardo and Robert Remez, thank you for continuing to keep me under your wing when I needed support and advice, even after I left the nest. To my thesis committee members, Ann Bradlow and Masaya Yoshida, thank you for your feedback and especially for your encouragement.

Lastly, and definitely most importantly, I'd like to thank my advisor, Matt Goldrick, without whom my dissertation would have been a very different beast. His generosity with his enthusiasm, encouragement, trust, time, expertise, and editing ability is, to my knowledge, unparalleled in the world of advisors. Beyond the scope of this thesis, you have sharpened me into a better thinker (and a kinder thinker), and I think it is that for which I owe the greatest debt—I will carry that with me always. Thank you, thank you. Thanks also to the National Science Foundation, whose grant to Matt helped to enable his generosity.

Chapter 2 is a revised version of Heller and Goldrick (in press). The dissertation author was the primary investigator and author of this paper.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Due to the complexity of the speech production system—its many processing levels, the interactive connections between these levels, and the decisions the system must make at each level—sending a word from idea to articulation is not a simple task. For example, good deal of research shows that, for the production of any given word, phonologically and semantically related words (phonological and semantic neighbors), as well as other words to be produced at other points in the utterance, may also be activated, creating a diffuse pattern of activation over the lexicon and increasing competition between these representations (Dell, Oppenheim, & Kittredge, 2008). However, normal speakers manage this complexity easily in everyday speech. One potential aid to the efficient management of this task is the ability to utilize (when available) the predictability of upcoming linguistic units to limit the number of candidate structures considered while making each decision. This predictability may be used to limit which units are eligible to be selected by the system at a given processing level, or, in a more extreme scenario, used to limit which units are eligible to be activated before selection occurs.

For example, when planning to produce the sentence *The bird sings its tune*, the semantic representations of *BIRD, SINGS,* and *TUNE* are activated in parallel at the message level, causing competition between lexical representations during production (e.g., Dell et al., 2008). This competition comes from at least two sources—from other words in the current utterance (e.g., between *BIRD* and *SINGS*) as well as from other candidate words in the lexicon (e.g., between *BIRD* and *SPARROW*—a semantic associate—and/or *WORD*—a phonologically similar representation; Dell et al., 2008). These sources of competition may be reduced by capitalizing on sources of word predictability in a given context.

For example, when planning the first noun (*BIRD*), the knowledge that a noun must occur soon after *the* would reduce competition from the verb *SINGS*, as well as from other words in the lexicon that do not share the grammatical category of noun. Recent work indicates that syntactic constraints limit the spread of semantic activation during lexical selection in production, such that semantically-related words of the same grammatical category interfere more with each other's processing than the processing of related words from other grammatical categories (Alario, Matos, & Segui, 2004; Pechmann & Zerbst, 2002; Vigliocco, Vinson, & Siri, 2005). These processes limit competition between words that differ in grammatical category but not competition from words that share a grammatical category. That is, the activation of *BIRD* would cause more interference for the processing of *TUNE* during the production of the sentence, and vice versa, than either noun would cause for the processing of *SINGS*, even though all three words are highly associatively related.

Meaning-related predictability may be able to speed processing over and above grammatical predictability by further limiting competition. While grammatical predictability limits competition from words of other grammatical categories, meaning-related predictability may be able to limit competition within a grammatical category, to words that are topically appropriate given the context. That is, during the production of our example sentence, *The bird sings its tune*, a phonological associate of *TUNE* such as *TUBE* may be activated. While *TUBE* shares *TUNE*'s grammatical category, it is less plausible in context, and so it is possible that its activation is limited in context by meaning-related predictability mechanisms.

The extent to which information about upcoming linguistic elements affects the flow of information throughout the linguistic processing system during sentence production—that is, how it affects phonological and phonetic processing, specifically—remains underexplored. In speech perception, a number of studies have documented patterns of interaction that may be complementary to potential effects in production: higher-level predictability constrains the bottom-up activation of phonologically related lexical items (e.g., Dahan, Swingley, Tanenhaus, & Magnuson, 2000; Dahan & Tanenhaus, 2004; Hartmann, 2004; Magnuson, Tanenhaus, & Aslin, 2008; Sekerina, Brooks, & Kempe, 2006; Sekerina, 2008; Sommers & Danielson, 1999; Strand, Simenstad, Cooperman, & Rowe, in press; but c.f. Spinelli, Meunier, & Seigneuric, 2006). These studies have shown that semantic, pragmatic, or grammatical (e.g., gender-marked) contexts that strongly predict upcoming linguistic units can constrain the search through the phonological space to items fit for that context. That is, the use of contextual information appears to lessen phonological competition from other words in the lexicon (although limitation of the influence of other words in the present linguistic context may also be possible). The use of contextual constraint in perception to limit phonological competition may closely mirror its use in production, or these constraints may operate under more distinct principles in the two modalities.

The current project will investigate how predictive processes at higher levels of planning affect phonological processing in speech production. This investigation will consist of three types of investigation. First, I will examine whether activation of phonologically related words (phonological neighbors) in the lexicon is constrained by grammatical encoding processes during sentence production. This will be accomplished by manipulating the presence or absence of sentential contexts of target words. I will then examine how these factors modulate the influence of phonological neighbors on processing of target words. Second, I will investigate whether phonological processing is constrained by predictive meaning-related processes, over and above constraints available from grammatical cues, during sentence and discourse production. This will be evaluated by manipulating the meaning-related predictability of target words in sentence and discourse contexts. I will then examine how high and low levels of meaning-related predictability modulate activation of targets' phonological neighbors. Third, in order to more fully understand the relationship between context-independent components of phonological processing with grammatically- and meaning-dependent components of phonological processing, if any, I will evaluate how context-dependent and context-independent factors affect phonetic outcomes of phonological processing in a corpus of spontaneous speech.

These investigations will add to the growing body of evidence about how predictive processes modulate linguistic processing during speech production. Moreover, data collected in these experiments will allow assessment of the relationship between the effects of contextual constraint on the planning and the execution of speech production processes by comparing speech planning (as assessed via latencies to speech) and phonetic outcomes, which reflect both these higher-level processes and articulatory execution.

Below, I will describe in more detail the structure of the speech production system, how phonological neighbors affect production, and how constraints on lexical activation and lexical selection comparatively affect speech production. Next, I will briefly describe the current investigations proposed to fulfill the three aims stated above. Finally, I will describe the implications and broader impacts of these experiments.

1.1.1 Structure of the speech production system

Lexical access in word production is largely held to be a two-step process (e.g., Schwartz, Dell, Martin, Gahl, & Sobel, 2006; see Fig. 1.1). The first step in the production process is the link between semantic features and word-level representations (*L-level* representations; Rapp & Goldrick, 2000), which are associated with grammatical information such as part of speech. The second step selects the corresponding phonemes. This has been modeled as a connectionist network with cascading activation between network levels and limited feedback between phonological and L-level representations (e.g., Rapp & Goldrick, 2000). The process begins with the activation of semantic features from the speaker's intended message, and this activation spreads throughout the system. If the intention is to utter the word *CAT* (/kæt/), then the features *small, furry, pet,* and *feline* might become activated. Activation from these semantic nodes cascades down to the L level, activating their associated L-level representations. For example, the features *small, furry,* and *pet* would cause coactivation of the L-level representations *DOG, CAT,* and *RAT*, among others*.* However, *feline* only sends activation to *CAT* in this candidate set. *CAT* therefore receives the most activation and as a result receives an extra boost of activation. L-level activation continues to cascade down to the phoneme level, where /k/, /æ/, and /t/ are activated. Activated phonemes can send some activation back up to the L-level. This feedback causes coactivation of formally similar words (phonological neighbors). For example, because of the feedback activation from /æ/ and /t/, *MAT*, which shares no semantic features with *CAT*, nonetheless enters the competition for phonological selection. *MAT* sends activation to /m/ and increases the activation of /æ/ and /t/. Finally, the phonemes with the highest activation are selected and sent to the articulatory system to be produced.

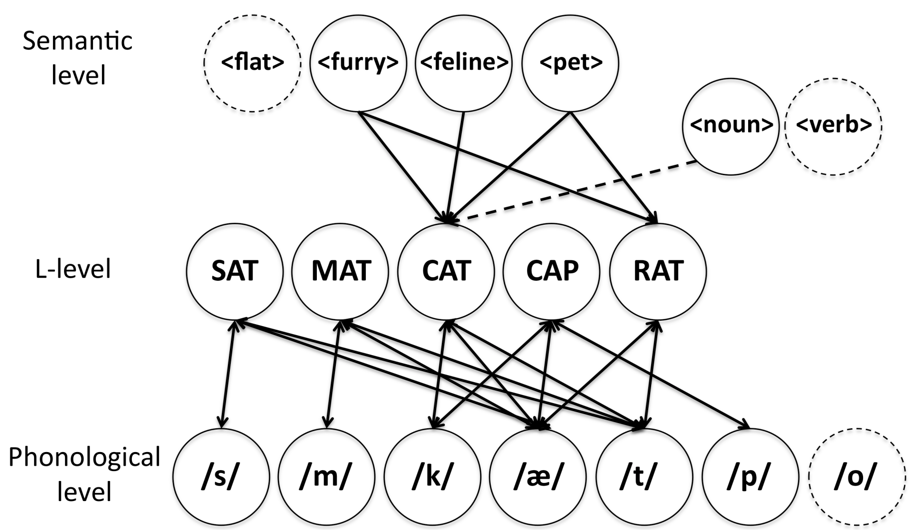

Fig. 1.1. Two-step production model structure for single word naming

In everyday usage, this word production system operates within higher-level discourse-structural and grammatical planning systems. These systems guide the structure and agreement of both topical and syntactic processing. A common mechanism utilized in production theories uses grammatical phrasal frames that include open slots for content words (e.g., Garrett, 1975) and function words (Bock, 1989; but c.f. Garrett, 1975). These slots are marked as to which grammatical classes can fill them. As discussed in more detail below, these frames influence L-level selection, ensuring that a grammatically appropriate lexical item is selected for phonological encoding. Fig. 1.1 shows a syntactic frame utilized for single word naming (e.g., bare noun naming of a picture). Evidence varies as to how much of a frame is planned at one time (deep structure clause: Ford & Holmes, 1978; linear phrase: Gillespie & Pearlmutter, 2011; Smith & Wheeldon, 2004; sentence: Griffin & Bock, 2000; Rapp & Samuel, 2002; Smith & Wheeldon, 2004; for a review, see Levelt & Meyer, 2000); most research agrees that more than one word but not more than one sentence is being processed at once (e.g., Rapp & Samuel, 2002).

1.1.2 Phonological neighbors in speech production

As mentioned above, words that are phonologically related to a word that a speaker intends to utter become activated during speech production. This activation affects target word processing, either in a facilitatory or inhibitory manner (which may reflect the relative activation of non-target representations; Chen & Mirman, 2012). Investigations of these effects are typically operationalized in terms of the number of a word's phonological neighbors, or the number of words differing from a target word by the addition, deletion, or substitution of a single phoneme (e.g., *CAP*, *SAT*, *CUT*, *AT, SCAT…* for target *CAT*). The denser a word's phonological neighborhood (i.e., the more phonological neighbors the word has), the slower (Gordon & Kurczek, 2014; Sadat, Martin, Costa, & Alario, 2014) or faster the reaction times (Vitevitch, 2002), greater (Goldrick, Folk, & Rapp, 2010; Gordon, 2002;

Harley & Bown, 1998; Vitevitch, 1997, 2002; Vitevitch & Sommers, 2003) or diminished (Newman & German, 2005) the naming accuracy, and more extreme (i.e., expanded F1-F2 vowel space; Kilanski, 2009; Munson, 2007; Munson & Solomon, 2004; Scarborough, 2010, 2012, 2013; Scarborough & Zellou, 2013; Wright, 2004) or less extreme (shorter word durations and contracted vowel space; Gahl, Yao, & Johnson, 2012) the phonetic properties.

One account of these effects attributes these effects to variation in the activation of phonological representations during lexical access. Under this account, reduction in response latency and increase in phonetic extremity is attributed to the increase in activation for phonemes that are shared by neighbors due to the activation-feedback resonance between phonemes and L-level representations (Baese-Berk & Goldrick, 2009; Vitevitch, 2002), which increases as neighborhood density increases when neighbor activation is relatively low. However, under this account, when activation of neighbors is relatively high (due to, e.g., task demands), inhibitory connections between L-level representations dominate the system dynamics (Chen & Mirman, 2012), leading to a reduction in relative target activation. This leads to an increase in response latency and reduction in phonetic extremity as neighborhood density increases when neighbor activation is high. Thus, under the two-step model of word production, changes in behavior related to phonological neighborhood density and changes in neighbor activation are driven by changes in L-level activation. Therefore, observing changes in phonological neighborhood effects is a useful tool for investigating constraints on lexical-phonological activation, regardless of effect direction. We take this model as the basis for the hypotheses tested in this paper. However, two other possible accounts of phonological neighborhood density effects are outlined below, and implications for these theories based on results from the current experiments will also be discussed in the concluding chapter of this dissertation.

1.1.2.1 Other explanations of phonological neighborhood density effects

Two other accounts for neighborhood effects in speech production have been put forward. These accounts are based on evidence from single word production showing that vowels in words with a greater number of overall phonological neighbors show expanded vowel spaces. Whereas the mechanisms in the model discussed above are entirely internal to the production system, both alternative accounts rely on perceptual components. One account holds that this added phonetic extremity is for the listener's benefit (e.g., Scarborough, 2010; Scarborough & Zellou, 2013; Wright, 2004). Words with more phonological neighbors have more words that sound similar to them, making them harder to identify during perception. As such, a speaker's increase in phonetic extremity might ease the listener's task of selecting among these relatively many alternatives. In contrast, when there are fewer neighbors, phonetic extremity is less necessary for the listener to select an appropriate representation; thus, the speaker produces less phonetic extremity. Another account holds that in order for the speaker themselves to develop a phonetic representation of a word, they must first successfully classify instances of the word perceptually (Pierrehumbert, 2002). Again, as listeners, they classify words with fewer competitors more easily, even if they contain less phonetic extremity, whereas more instances of words with more competitors will be successfully classified if they contain more phonetic extremity. Therefore, speakers develop word-specific phonetic representations. For example, the speaker-internal phonetic representation for the quality of the vowel in the word *CAT*, which has many phonological neighbors and therefore many perceptual competitors, will contain more phonetic extremity than the stored vowel quality for *HAT*, which has fewer neighbors in English and is therefore easier to identify. The speaker reproduces these differences in perceptual encoding when they produce words based on these stored phonetic representations.

1.1.3 Moving beyond context-independent processing of phonological neighbors

The studies mentioned above have assumed that all phonological neighbors play a role in the processing of each target word, no matter what the context of its appearance. However, there is reason to believe that phonological neighbors do not function interchangeably. For example, when a lexical speech error is made that is phonologically related to the intended word, the part of speech of the unintended word is more likely than chance to match the part of the speech of the intended word (e.g., Nooteboom, 1969; see Goldrick et al., 2010, for a recent review). This effect is even stronger when the words are produced in sentential contexts (e.g., Berndt, Mitchum, Haendiges, & Sandson, 1997). This suggests that syntactic processing influences the processing of phonologically related words, and moreover that the constraint is stronger the stronger the syntactic context. Similarly, word-level errors that are both semantically and phonologically related to the intended word (e.g., *CAT* → *RAT* vs. *CAT* → *DOG* or *CAT* → *CAB*) are made more often than would be expected by chance (see Goldrick, 2006, for a recent review). This *mixed error effect* suggests that semantic processing may modulate phonological activation during word production, as well.

1.1.4 Contrasting context-dependent and context-independent activation mechanisms

Two potential mechanisms could allow syntactic and thematic constraints—which are stronger in sentential and highly predictive contexts—to conspire with phonological processing to produce the speech error results discussed above. In the first, lexical-phonological activation is modulated by contextual constraints to increase the relative influence of those representations that are appropriate in the context—phonological activation proceeds in a context-dependent manner. In the second, lexical-phonological activation proceeds in a context-independent manner but its output is checked by a contextually-sensitive monitoring system.

1.1.4.1 Context-dependent phonological activation

To account for the mutual influence of grammatical, meaning-related, and phonological information on lexical selection in syntactic contexts, where contextual constraints are strong, a mechanism could be incorporated that in such contexts either inhibits L-level representations that do not match the grammatical or semantic 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, meaning-related, and phonological information to mutually influence L-level selection. Specifically, context-dependent activation from the grammatical and/or semantic processing system(s) may change the pattern of activation in the lexicon, which in turn would change the pattern of lexical-phonological feedback activation that drives phonological neighborhood effects.

This account makes a novel prediction regarding the influence of phonological neighbors: in syntactic contexts, primarily same-category neighbors should affect phonological encoding. 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. A similar mechanism may limit the activation of L-level representations not related to the target word in meaning, allowing a similar effect of meaning-related constraints on phonological neighborhood processing, especially when meaning-related word predictability is high. Thus, under this account, reaction times and phonetic outcomes would be 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 neighborhood density but differing in within-category density should therefore behave differently from one another in constraining contexts.

1.1.4.2 Context-independent phonological activation

Alternatively, grammatical and meaning-related 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). 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.

1.1.4.3 Relationship between context-dependent and –independent mechanisms

In the case that results of the current experiments show support for context-dependent phonological processing, several hypothesis could be outlined about the relationship between context-dependent phonological processing mechanisms and the context-independent mechanism that has gained much support in the single word production literature. On the one hand, phonological processing in sentence contexts, such as during spontaneous speech, may be entirely determined by grammatically- and meaning-dependent mechanisms; context-independent processing may be reserved for contexts in which grammatical and meaning-related constraints are not present. In other words, effects of grammatical and meaning-related constraints may dominate phonological processing when these constraints are available, virtually eliminating any effects of neighbors that do not share the target's category. The fact that previous studies (e.g., Gahl et al., 2012) have observed effects of overall neighborhood density in spontaneous speech does not rule out this account; overall neighborhood density may be highly correlated with within-category neighborhood density, especially with respect to grammatical processing. That is, a word with many neighbors may be more likely to have more within-category neighbors, as well. Therefore, previous analyses reporting effects of overall density on processing during spontaneous speech may have been inadvertently detecting effects of within-category density. If this account is correct, overall neighborhood density should have no reliable effect on phonological processing when within-category density is controlled.

On the other hand, it is possible that both context-independent and -dependent phonological processing mechanisms are at work during fluent production. For example, all of a word's neighbors may play some role in its phonological processing, but within-category neighbors may play an additional (or different) role, beyond that contributed by other phonological neighbors. This would occur if grammatical and meaning-related constraints on L-level activation are gradient rather than absolute. If this is the case, both overall neighborhood density and within-category density should have independent effects on phonological processing when the influence of each is simultaneously brought under statistical control.

1.1.4.4 Separating grammatically dependent and meaning-dependent mechanisms

To contrast these accounts, this project examines the modulation of effects of phonological neighbors by context with respect to two sets of constraints on lexical items: grammatical and meaning-related constraints. The former type is governed by rules of the grammatical system, while the latter is constrained by associative and structural semantic and thematic relations. Predictability of upcoming linguistic units due to both kinds of constraints has been shown to affect production (meaning-related: e.g., Clopper & Pierrehumbert, 2008; Gregory, Raymond, Bell, Fosler-Lussier, & Jurafsky, 1999; Griffin & Bock, 1998; Lieberman, 1963; Scarborough, 2010; syntactic: Jaeger, 2010; Tily, Gahl, Arnon, Snider, Kothari, & Bresnan, 2009). However, it is not clear that both types of constraints have the same relationship to lexical activation and selection. Therefore, the role of each type of constraint in activation and selection must be examined separately.

The context-independent account of phonological processing claims that only selection is constrained by higher-level (i.e., thematic and grammatical) constraints; activation is unconstrained. This predicts that all phonological neighbors become activated and all contribute to the phonological activation of the target word's phonological representation, no matter what the grammatical or meaning-related constraint. For example, if activation is not limited to nouns in a context where nouns would be produced, *SOON* will be able to influence the processing of *TUNE* in both unconstrained as well as sentential contexts.

In contrast, according to the context-dependent account of phonological processing, contextual constraints limit the activation of phonological neighbors. This predicts that the effect of phonological density on target processing should be influenced by the context in which a word is produced. If the context-dependent account holds only for syntactic constraints but not for thematic constraints, any neighbor of the appropriate grammatical category, regardless of thematic fit, would contribute to phonological activation. In this case, the activation of /t/, /u/, and /n/ in the perception or production of *TUNE* in the example sentence *The bird sang its tune* would be affected by neighbors such as *MOON, TONE,* and *TUBE*, which share the grammatical category of the target. However, phonological activation would not be affected by the neighbor *SOON*, which differs in grammatical category from the target (n.b. *MOON* in place of *TUNE* would be thematically anomalous but syntactically legal).

If, on the other hand, the context-dependent account holds for both grammatical and meaning-related constraints, only words that are both thematically and syntactically plausible in context would be eligible for activation. In this case, *TONE* might contribute to phonological activation of *TUNE* (to the extent that it is thematically likely in that context), but neither *MOON, SOON*, or any other anomalous word should contribute phonological activation in this context.

It should be noted that for the purposes of the present investigation, grammatical constraints and meaning-related constraints are construed quite broadly. Grammatical constraints are assumed to arise in syntactic contexts where there is some cue to upcoming parts of speech (e.g., *the…* cues noun occurrence soon after). However, no assumptions are made about whether these constraints arise due to linear predictability or structural predictability. Similarly, meaning-related constraints are assumed to arise when a context is predictive of a certain word based on semantic relationships that word has to previous words in the sentence. However, multiple types of relationship exist (e.g., selectional/pragmatic relations between a verb and its object, such requirement of a liquid object for the verb *drink*, versus, e.g., associative relations between words of the same category such as *sailor* and *mariner*); no one relationship in particular is investigated here, although it is possible that differences between these types of relationships affect processing differently.

1.2 Current investigation

The first experiment, outlined in Chapter 2 (Heller & Goldrick, in press) investigates the role of grammatical constraint on the limitation of activation and selection in production of grammatically (but not thematically) constrained contexts (versus the absence of a syntactic context). The second set of experiments, detailed in Chapter 3 (Heller

& Goldrick, in preparation b) investigate the role, if any, of meaning-related predictability on these processes in the production of sentences and short discourses. In these investigations, both latency to speech onset and phonetic measures are analyzed, allowing comparison of the effects of contextual constraint on speech planning and phonetic processing. The third investigation, detailed in Chapter 4 (Heller & Goldrick, in preparation a) assesses the relationship between context-independent and -dependent mechanisms in an analysis of word duration and phonetic properties of vowels in a corpus of spontaneous speech.

To preview the results of the current investigation, grammatical constraints, but not meaning-related constraints, affect phonological neighbor activation in sentence contexts. Within-category neighbors influenced planning time and phonetic encoding more strongly in sentence contexts, supporting a grammatically dependent phonological processing mechanism. Meaning–related constraints independently affect target planning and phonetic encoding. However, meaning-related constraints do not reliably affect the activation of phonological neighbors over and above effects of grammatical constraints, suggesting a far smaller role for meaning-related constraint on phonological neighbor activation. In an analysis of word productions from spontaneous speech, independent roles were found for grammatically dependent and -independent phonological processing mechanisms, suggesting both are at work during everyday speech processing. Results are summarized in Chapter 5, where implications of these results for theories of speech production are also discussed.

This project is designed to investigate how predictive processes modulate phonological activation and selection in speech production. Specifically, this work provides evidence about how speakers leverage grammatical and meaning-related constraints to manage the activation of phonological neighbors, thereby streamlining the sentence production process. By investigating both reaction times and phonetic outcomes, it provides evidence about the extent to which this activation spills over into phonetic and articulatory processing. The current results not only shed light on the details of processing throughout the speech production system, but also establish a theoretical connection between existing speech production theories and predictive processes, making strides towards the goal of accounting for production behavior during everyday speech.

CHAPTER 2

Grammatical constraints on phonological encoding in speech production

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) versus 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 that phonological interaction affects articulation after speech onset. These results support production theories integrating grammatical, phonological, and phonetic processes.

2.1 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 nonerrorful 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 that grammatical class strongly influences phonological activation in syntactic contexts, supporting interactive processing models.

2.1.1 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 longer (Gordon & Kurczek, in press; Sadat, Martin, Costa, & Alario, 2014) or shorter (Vitevitch, 2002) reaction times (RTs), 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*, and so forth, 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).

2.1.2 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.

2.1.2.1 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, RTs 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.

2.1.2.2 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 in both the presence and absence of a syntactic context.

2.1.3 Current investigation

To test these predictions, we compare production of words with relatively few versus 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 versus phonetic processes.

2.2 Methods

2.2.1 Participants

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

2.2.2 Materials

Target words were 24 English CVC nouns (see Appendix A), 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, ɛ, æ, a, o, u/. Picture name agreement for illustrations representing targets (see Appendix A) was also matched across stimulus groups.

Four non-thematically predictive carrier sentences were designed for each target word, creating 96 target sentences (see Appendix A[1]). 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.

2.2.3 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.

[1] Two of these sentences contained a neighbor of the target in an adjacent phrase (e.g., *He chose to <u>look</u> 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).

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 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.

2.2.4 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.

2.2.4.1 Speech onset latencies

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.

2.2.4.2 Vowel duration

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

2.2.4.3 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.

2.3 Results

2.3.1 Speech onset latencies (RTs)

2.3.1.1 Analysis

Exclusions of errors and RTs shorter than 300 ms resulted in the 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).

2.3.1.2 Results

Results for the effects of interest are summarized in Fig. 2.1.[2] Pictures were named faster in the sentence context condition than in bare naming (β = -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 significant interaction of noun density and condition (β = 0.0038, SE = 0.0017, $\chi^2(1)$ = 4.79, p = 0.029). The effect of noun density was larger in sentence contexts than in bare naming; in sentence contexts, words with higher noun density elicited significantly longer RTs than did words with lower noun density.

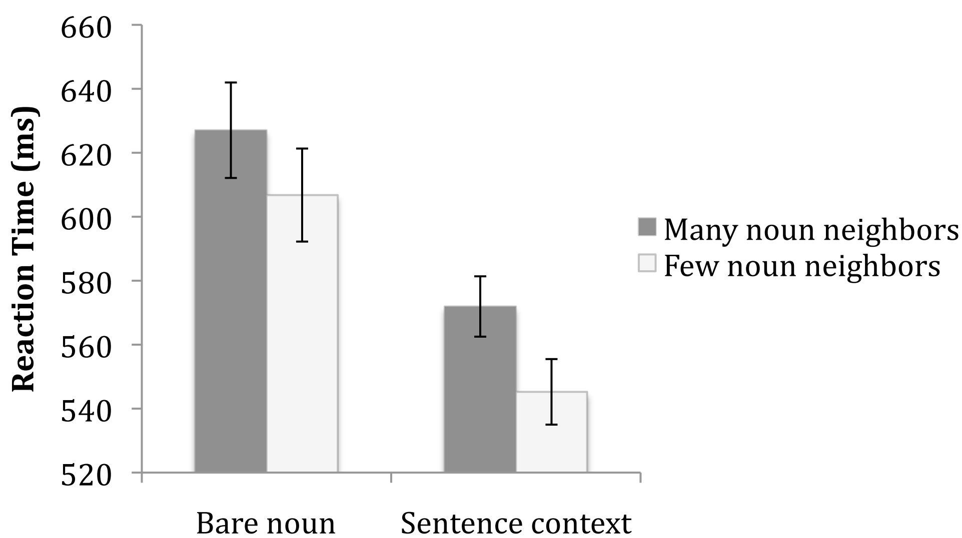

Fig. 2.1. Reaction times (latencies to speech onset) for words with many and few noun neighbors in bare naming and sentence contexts. Error bars indicate by-participant standard error.

[2] Similar results were found for the subset used in the vowel analyses.

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

2.3.2 Vowel duration

2.3.2.1 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.

2.3.2.2 Results

Fig. 2.2 shows the fixed effects of interest for vowel duration. Items with higher noun density had significantly shorter vowel durations (β = 0.13, 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 in the bare naming condition (β = 0.0037, SE = 0.00025, $\chi^2(1)$ = 78.60, p < 0.0001). Vowel ($\chi^2(5)$ = 2.66; p > 0.1) and other factors (χ^2s(1) < 1; ps > 0.1) were not significant predictors.

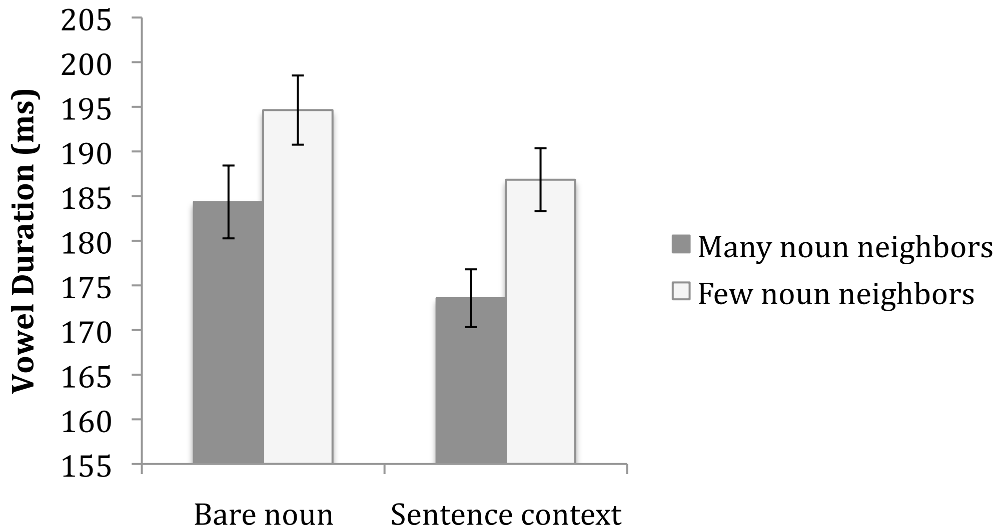

Fig. 2.2. Vowel durations for words with many and few noun neighbors in bare naming and sentence contexts. Error bars indicate by-participant standard error.

2.3.2.3 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 (β = 0.010, SE = 0.0058, $\chi^2(1)$ = 3.10, p = 0.0782), the effect of noun density (β = 0.12, SE = 0.00081, $\chi^2(1)$ = 573.66, p < 0.0001) and its interaction with condition (β = 0.011, SE = 0.00054, $\chi^2(1)$ = 65.77, p < 0.0001) remained significant.

[3] This analysis relies on the assumption that RT and vowel length are linearly related. Visual data inspection did not suggest any other relationship.

2.3.3 Vowel space size

2.3.3.1 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).

2.3.3.2 Results

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

2.4 Discussion

2.4.1 Syntax-dependent phonological processing

While previous research has suggested that 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 versus 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 RTs 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 not only reveal 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.

2.4.2 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 longer 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.

2.4.3 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 that 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.

2.5 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 in running participants, Ariana Steele for assistance with acoustic reliability measurements, and the Northwestern Sound Lab for helpful comments.

CHAPTER 3

Meaning-related constraints modulate speech planning and phonetic processing but not phonological neighborhood activation during speech production

ABSTRACT

Many studies suggest that the predictability of words in context influences both speech planning and phonetic processing. In order to better understand the mechanisms through which predictability affects activation in the speech production system, we examine what effect, if any, meaning-related predictability has on phonological processing during sentence production. In two experiments, we compare how contexts with high and low meaning-related constraint on target identity in sentence and short discourse contexts affect the activation of phonological neighbors (words phonologically related to the target) over and above effects of grammatical constraint. Results suggest that meaning-related constraint independently affects target planning, as reflected in latencies to picture naming, and phonetic encoding, as reflected in phonetic properties of vowels in picture names. Grammatical constraints modulate the activation of phonological neighbors, replicating previous findings. However, meaning-related constraints do not reliably affect the activation of phonological neighbors over and above effects of grammatical constraints. These results refine our understanding of the effects of contextual predictability on speech production in sentence contexts.

3.1 Introduction

When a speaker produces a word in a sentence, many pieces of linguistic knowledge must be coordinated. These include the speaker's intended message and the meaning of the current word to be produced, the syntactic structure of the sentence and the syntactic features of the current word, as well as the phonological and phonetic properties of that word, at a minimum. A good deal of research shows that phonologically and semantically related words (phonological and semantic *neighbors*), as well as other words to be produced at other points in the utterance, may also be activated during this process, creating a diffuse pattern of activation over the lexicon and increasing competition between these representations (Dell, Oppenheim, & Kittredge, 2008). Thus, selecting an appropriate word for production at the appropriate time is not trivial.

A number of studies suggest that word predictability may streamline this process, speeding speech planning and phonetic processing. For example, words that are high in frequency are produced faster (Griffin & Bock, 1998), with shorter pronunciation durations (e.g., Bell, Brenier, Gregory, Girand, & Jurafsky, 2009), and with vowels exhibiting less extreme spectral properties (Munson & Solomon, 2004) than lower frequency words. Grammatical predictability also limits effects of simultaneously activated words: in a phrasal context, words that are semantically related to a target word that are not syntactically appropriate affect processing time less than words that share the grammatical category of the word to be produced (Alario, Matos, & Segui, 2004; Pechmann & Zerbst, 2002; Vigliocco, Vinson, & Siri, 2005). Phonologically related words that do not share the target word's grammatical category have limited influence on target processing during sentence production, as well (Heller & Goldrick, in press). Words that are predictable due to the meaning of the sentence they occur in are produced faster (Griffin & Bock, 1998), with shorter durations (Lieberman, 1963; but c.f. Hunnicutt, 1987), and may contain vowels with more spectral reduction (Scarborough, 2010; but c.f. Clopper & Pierrehumbert, 2008). Similarly, words that are more predictable given the directly preceding word or syllables may be produced with shorter durations (Aylett & Turk, 2004; Bell et al., 2009; Jurafsky, Bell, Gregory, & Raymond, 2001) and with vowels that exhibit less extreme spectral properties (Aylett & Turk, 2006; Jurafsky et al., 2001). Words that are predictable in virtue of having been previously evoked in the speech context are also produced with shorter durations (Fowler & Housum, 1987; see Lam & Watson, in press, for a recent review).

Although it is clear that many types of predictability speed processing during speech production, work remains in the investigation of the mechanisms by which these diverse types of predictability act on the speech production system. The current paper aims to increase our understanding of how predictability affects information flow throughout the linguistic processing system by assessing how meaning-related predictability affects phonological and phonetic processing. Specifically, we investigate how meaning-related word predictability in a sentence or discourse context affects word processing in the phonological neighborhood, over and above effects of syntactic constraints on these processes. In this endeavor, we will investigate effects of meaning-related contextual constraint on phonological neighborhood effects in speech planning processes (as assessed via latencies to speech execution) as well as phonetic outcomes (which reflect both planning and articulatory execution). Furthermore, we examine the relationship between these effects on speech planning and phonetic processing.

3.1.1 A model of single word production

Word production is largely held to be a two-step process (e.g., Schwartz, Dell, Martin, Gahl, & Sobel, 2006). The first step links semantic features and word-level representations (*L-level* representations; Rapp & Goldrick, 2000), which are associated with grammatical information such as part of speech. The second step selects the corresponding stored sound structures (e.g., segmental representations) at the phonological level. This process has been modeled as a connectionist network with cascading activation between network levels and limited feedback between phonological and L-level representations (e.g., Rapp & Goldrick, 2000). The process begins with the activation of semantic features from the speaker's intended message, and this activation spreads throughout the system. If the intention is to utter the word *CAT* (/kæt/), then the features *small*, *furry*, *pet*, and *feline* might become activated. Activation from these semantic nodes cascades down to the L-level, activating associated L-level representations. For example, the features *small*, *furry*, and *pet* might activate the L-level representations *DOG*, *CAT*, and *RAT*, causing co-activation of semantically related representations (semantic neighbors). However, among that candidate set, the feature *feline* activates only *CAT,* which therefore receives the most activation and subsequently receives an additional activation boost. L-level activation continues to cascade down to the phonological level, where /k/, /æ/, and /t/ are activated. Activated phonemes send a limited amount of activation back up to the L-level. This feedback causes co-activation of formally similar words (phonological neighbors). For example, because of the feedback activation from /æ/ and /t/, the L-level representation *MAT*, which shares no semantic features with *CAT*, nonetheless becomes partially activated. *MAT* sends activation back down to /m/ and increases the activation of /æ/ and /t/. Finally, the phonemes with the highest activation are selected and sent to the articulatory system to be produced.

3.1.2 Phonological neighborhood effects in word production

As mentioned above, words that are phonologically related to a word that a speaker intends to utter become activated during speech production. This activation affects target word processing, either in a facilitatory or inhibitory manner (which may reflect the relative activation of non-target representations; Chen & Mirman, 2012). Investigations of these effects are typically operationalized in terms of the number of a word's phonological neighbors, or the number of words differing from a target word by the addition, deletion, or substitution of a single phoneme (e.g., *CAP*, *SAT*, *CUT*, *AT, …* for target *CAT*). The denser a word's phonological neighborhood (i.e., the more phonological neighbors the word has), the slower (Gordon & Kurczek, 2014; Sadat, Martin, Costa, & Alario, 2014) or faster the reaction times (Vitevitch, 2002), greater (Goldrick, Folk, & Rapp, 2010; Gordon, 2002; Harley & Bown, 1998; Vitevitch, 1997, 2002; Vitevitch & Sommers, 2003) or diminished (Newman & German, 2005) the naming accuracy, and more extreme (i.e., expanded F1-F2 vowel space; Munson & Solomon, 2004; Scarborough, 2010; Wright, 2004) or less extreme (shorter word durations and contracted vowel space; Gahl, Yao, & Johnson, 2012) the phonetic properties.

One account of these effects attributes these effects to variation in the activation of phonological representations during lexical access. Under this account, reduction in response latency and increase in phonetic extremity is attributed to the increase in activation for phonemes that are shared by neighbors due to the activation-feedback resonance between phonemes and L-level representations (Baese-Berk & Goldrick, 2009; Vitevitch, 2002), which increases as neighborhood density increases when neighbor activation is relatively low. However, under this account, when activation of neighbors is relatively high (due to, e.g., task demands), inhibitory connections between L-level representations dominate the system dynamics (Chen & Mirman, 2012), leading to a reduction in relative target activation. This leads to an increase in response latency and reduction in phonetic extremity as neighborhood density increases when neighbor activation is high. Thus, under the two-step model of word production, changes in behavior related to phonological neighborhood density and changes in neighbor activation are driven by changes in L-level activation. Therefore, observing changes in phonological neighborhood effects is a useful tool for investigating constraints on lexical-phonological activation, regardless of effect direction. We take this model as the basis for the hypotheses tested in this paper; we return to other possible accounts of phonological neighborhood density effects in the discussion.

3.1.3 Contextual constraints on neighborhood activation during sentence production

Recent research suggests that activation in the phonological neighborhood during speech production may be modulated by constraints on L-level properties in a given sentential context. Specifically, Heller and Goldrick (in press) showed that when the total number of phonological neighbors was controlled, the density of neighbors sharing a target word's grammatical category (noun) influenced processing more strongly when immediately preceded by a definite determiner in a sentence than in the absence of a syntactic context. That is, when *CAT* occurs in the absence of a syntactic context (e.g., bare picture naming), both neighbors that share its part of speech, such as *MAT,* and neighbors that do not, such as *SAT*, appear to participate in lexical-phonological activation feedback resonance, as described in the single word production model above. However, in a more constrained syntactic context, such as "*… the CAT*", neighbors like *SAT*, which do not share the target's part of speech, appear to contribute less to target activation. Specifically, when the word was produced in a sentence context, an increase in within-category neighborhood density was associated with longer reaction times and shorter vowel durations. Moreover, changes in vowel duration could not be predicted by changes in reaction times. These results are consistent with a model in which grammatical processing constrains activation in the phonological neighborhood in a context-dependent manner and continue to affect production after speech is initiated. For example, under the activation-based model sketched above, different-category L-level representations may be inhibited—or same-category representations may be boosted—when grammatical constraints are strong. This decreases the influence of different-category neighbors during phonological processing, resulting in changes to the relative activation of the phonological representations of the target, thus altering response times and phonetic properties.

Might context-specific activation of L-level representations from the semantic level due to meaning-related word predictability further limit phonological activation during sentence production? This would require semantic-lexical and lexical-phonological processing (i.e., steps 1 and 2 in the two-step model outlined above) to overlap in time in order to interact at the L-level. Many studies have attempted to elucidate the nature of semantic-phonological interactions during speech production (particularly during single word production), although fewer have focused specifically on the effect of meaning-related predictability on target processing during sentence production.

Evidence supporting some level of mutual influence of semantic and phonological levels on L-level selection comes from both investigations of speech errors and from chronometric investigations of speech production (see Goldrick, 2006, for a review). For example, in speech errors, L-level errors that are both semantically and phonologically related to the intended word (e.g., *CAT* → *RAT* vs. *CAT* → *DOG* or *CAT* → *CAB*) are made more often than would be expected by chance, suggesting that these two types of information must be acting in concert (Rapp & Goldrick, 2000). Similar interactions have been shown in chronometric investigations of priming during speech production. For example, priming for words phonologically related to near-synonyms (e.g., priming for *SODA* from *COUCH*, which has the near-synonym *SOFA*; Peterson & Savoy, 1998) suggests that semantically-related L-level representations can influence phonological encoding. However, at least in these single word production contexts, semantic associates weaker than near-synonyms do not give rise to phonological priming effects (e.g., no reliable priming for *BET* from *COUCH,* which has the category associate *BED*; Levelt, Schriefers, Vorberg, Meyer, Pechmann, & Havinga, 1991; Peterson & Savoy, 1998). This suggests that, in single word production, semantic influences on L-level activation for semantically related words such as category associates are too weak to detect when semantic relationships are relatively weak (versus strong, such as in near-synonymy).

Because words to be produced in an utterance may be activated concurrently at the message level (see Dell et al., 2008, for a review), one might expect meaning relationships to have more robust effects on phonological processing during the production of words in sentences, particularly for words whose meanings are highly related to—and predictable from—the meaning of other words in the utterance. Fewer studies have specifically investigated the role of meaning-related predictability on phonological processing during sentence production. Using a picture naming paradigm for target words in semantically constraining and unconstraining sentence contexts, Griffin and Bock (1998) investigated the interaction between target word frequency—thought to affect phoneme selection processes—and semantic constraint. They found an interaction between target frequency and constraint condition, such that the frequency effect (high frequency words being named faster than low frequency words) disappeared when words were placed in high constraint contexts. In the activation-based model discussed above, such an effect would be predicted if activation cascades down from the semantic to the lexical and phonological levels, enhancing the activation of predictable target words. Because high frequency lexical and/or phonological representations are already easy to retrieve, they benefit less from this cascading activation than low frequency words. However, because such an account refers only to the activation of target representations, it does not reveal how meaning-related activation influences the activation of non-target phonological neighbors during sentence production.

In a study of vowel production, Scarborough (2010) found no evidence that the processing of non-target neighbors is influenced by semantic processing. The effect of frequency-weighted phonological neighborhood density on vowel realization was not statistically different in high and low semantic constraint contexts, although both neighborhood density and semantic constraint showed significant effects. Thus, although meaning-related predictability may increase activation of the target representation at the L-level and phonological level, it may not influence the activation of neighbors.

3.1.4 Current investigation

In the current investigation, we build on this previous research by asking specifically whether context-specific meaning-related activation—e.g., due to word predictability provided by the meanings of the words that precede it—influences phonological encoding beyond effects associated with grammatical processing. Moreover, we specifically compare effects on speech planning (latencies to speech during picture naming) and phonetic outcomes (durational and spectral properties of vowels in target words); in contrast, previous work has typically focused on one type of measure (e.g., Griffin and Bock (1998) focus only on latencies to speech, while Scarborough (2010) focuses only on phonetic outcomes). We manipulate the number of neighbors that share the target's grammatical category (i.e., words that are grammatically plausible alternatives for the target word; here, nouns), while controlling for targets' total neighborhood density, in two types of context with higher or lower levels of meaning-related constraint on target word identity. We first manipulate constraint by preceding a non-thematically constrained (but grammatically constrained) target sentence with a context sentence designed to make each target either more predictable *than its neighbors* or no more predictable than its neighbors (Experiment 1). The non-thematically constrained target sentences used in this study are identical to those investigated in the previous investigation of grammatical-category-specific neighborhood density (Heller & Goldrick, in press), allowing direct comparison of the effects of grammatical and meaning-related constraint. Second, we test increased meaning-related constraints on target identity in single sentences, where, in the high meaning-related constraint condition, targets are the most probable sentence completion rather than simply being more probable than their neighbors, versus, in the low constraint condition, sentences where targets are plausible but highly unpredictable (Experiment 2).

Given the evidence that semantic-phonological interaction occurs in some production contexts, and taking the two-step model of speech production as a basis, we outline hypotheses about how meaning-related predictability might interact with phonological processing during word production in sentence contexts. One possible mechanism would increase the activation of L-level representations sharing semantic features with a target word to a greater extent when the target is predictable given the meanings of the words preceding it. Alternately, L-level representations that do not share the semantic features of the target may be inhibited, decreasing their activation.

These two hypothesized mechanisms lead to at least two possible predictions. If semantic associates have increased activation at the L-level, which cascades down to the phonological level, we might expect a very small change in lexical-phonological interaction in the phonological neighborhood. This is because few of a word's phonological neighbors, which participate in its phonological processing, are also semantic neighbors. For example, when *CAT* is produced, the activation of *RAT* (which shares several semantic features) and its associated segmental representation may also be raised, especially in predictable contexts, but the activation of *MAT*, *CAP*, *CUT*, etc., (and their phonological representations) is not affected by the increased activation from semantic features related to *CAT*. Therefore, although semantic-phonological interaction may be occurring, any changes in phonological activation may be small and difficult to detect, even when semantic activation of the target and its semantic associates are boosted due to meaning-related predictability in the sentence context. In this case, the number of neighbors that share a target's grammatical category should have a roughly equivalent effect on reaction times and phonetic properties whether the target appears in a highly predictable context or an unpredictable context.

In contrast, if the L-level representations that are not semantically related to the target are inhibited, a larger change in lexical-phonological processing is predicted. This is because, as noted above, most of a word's phonological neighbors are not semantic associates. Following the example above, when *CAT* is produced, the L-level representations for *MAT*, *CAP*, *CUT*, etc., would be inhibited, although the activation of *RAT* and its associated segmental representation may also be relatively unchanged. This effect would emerge more strongly when activation of *CAT* and its associated semantic features is relatively high (i.e., when the word is predictable given its context). This would functionally reduce the contribution of most phonological neighbors to target processing in highly predictable contexts; if L-level representations of neighbors are inhibited, they cannot participate in the lexical-phonological interaction that is thought to drive phonological neighborhood effects. Therefore, under this hypothesis, the number of a target's within-category neighbors should have less (or no) impact on speech planning and phonetic processing when target likelihood is high, but should have a greater impact on processing time when target likelihood is low (or more matched to that of its neighbors).

The results suggest that meaning-related constraint independently affects target planning, as reflected in latencies to picture naming, and phonetic encoding, as reflected in phonetic properties of vowels in picture names. Grammatical constraints modulate the activation of phonological neighbors, replicating previous findings. However, meaning-related constraints do not reliably affect the activation of phonological neighbors over and above effects of grammatical constraints. We conclude that a model under which neighbors that do not share a target's semantic features are inhibited during lexical-phonological interaction is inconsistent with the current data.

3.2 Experiment 1

3.2.1 Methods

3.2.1.1 Participants

Sixty-four native English speakers (38 female) aged 18-23 with no history of speech or language deficits participated in exchange for course credit or $10.

3.2.1.2 Materials

Stimuli consisted of two-sentence pairs: a context sentence that was either thematically predictive or unpredictive of the upcoming target's identity (depending on condition), followed by a thematically unpredictive but grammatically predictive target sentence which contained the target word as its final word.

Target words were 20 CVC nouns (stimuli included in vowel analyses in Heller and Goldrick, in press), 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, sum uniphone and biphone phonotactic probability, vowel identity (vowels included /i, ɛ, æ, a, o, u/) and picture name agreement (see Appendix B).

The non-thematically predictive target sentences were the carrier sentences for target words included in vowel analyses in Heller and Goldrick (in press). Targets appeared in sentence-final position preceded by a definite determiner (e.g., *She wanted a peck at the PEN*). In order to maximize possible phonological interaction, each target sentence contained a prime word whose phonological onset and part of speech (noun) matched that of the target word (e.g., *peck*/*PEN*).

Two context sentences were designed to serve as a preamble to the target sentence, creating a total of 40 sentence pairs. One version of this context sentence was designed to be a plausible preamble to the target sentence, but not related in any way to the target word itself. Thus, the only clues to the word's identity in context were grammatical cues from its position in the target sentence (Grammatical Constraint condition); many of a target's neighbors that share its grammatical category would be plausible sentence pair completions. The other version of the context sentence was thematically related to the target word, designed to make a target's appearance more likely in the target sentence than its neighbors (Meaning + Grammatical Constraint condition). Example sentence pairs can be seen in 1a-1b.

1a) Grammatical Constraint

Annie was shocked that David guessed correctly what she'd brought to work that day. He confessed a peep at the *PEEL*.

1b) Meaning + Grammatical Constraint

Annie was shocked that David guessed correctly that she'd had a banana for lunch. He confessed a peep at the *PEEL*.

The likelihood of the target (versus a high frequency noun neighbor, used as a proxy for all neighbors) appearing as a sentence completion given its two-sentence context frame was normed using subjective likelihood ratings in a separate study (see Appendix B). Sentence pairs were divided into two lists of 20 items on which half of the targets appeared in each context (Grammatical Constraint vs. Meaning + Grammatical Constraint) and each target appeared only once. Each of these were integrated with 20 filler sentence pairs of various structures and various meaning-related constraint strengths. Two pseudorandomized orderings were created for each list, creating a total of four lists.

Sentence pairs appeared as printed words, with the exception of the target word, which appeared as a colored illustration. Pictures presented are described in Heller and Goldrick (in press).

3.2.1.3 Procedure

Participants were first familiarized with the pictures used in the experiment, and then familiarized with the sentence reading and picture naming task by completing three practice trials. Participants were asked to read aloud the sentence pair and then name the picture that appeared to complete the sentence pair as quickly and accurately as possible. Participants clicked a mouse to begin each trial. First, the entire context sentence (either the Grammatical Constraint or Meaning + Grammatical Constraint sentence) appeared centered on the screen, where it remained until participants completed reading the sentence aloud and clicked a mouse key to advance to the target sentence. A fixation cross appeared for 500 ms in the center of the screen, followed by the presentation of the target sentence one word at a time in the center of the screen. Each word appeared centered on the screen for 500 ms; participants read each word aloud as it appeared. In place of the target sentence's final word, its picture appeared and remained until participants clicked the mouse[4]. Participants saw one stimulus list; they named pictures in both the Grammatical and Meaning + Grammatical conditions but saw each target only once. Six comprehension questions, which focused on information from the context sentence, were interspersed to encourage integration of information across sentences.

---

[4] In a pilot experiment, both sentences were presented in a word-by-word manner. This led to very long picture naming latencies and disrupted all effects of meaning-related constraints. This is consistent with evidence that while sentences presented in a word-by-word manner are well comprehended and remembered, paragraphs presented in this manner are less well comprehended and remembered (Potter, Kroll, & Harris, 1980).

3.2.1.4 Speech Analysis

Trials with errors or disfluencies were excluded from all analyses. Trials in which participants began producing a picture name prior to 300 ms after picture presentation were also excluded.

3.2.1.4.1 Speech onset latencies

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

3.2.1.4.2 Vowel duration

Vowel duration was hand-measured in Praat. Vowel onsets and offsets were marked using cues from the waveform and spectrogram (see Heller & Goldrick, in press, [Appendix A] for details). A second coder marked 7% of the data to assess measurement reliability. Measurements were well correlated ($r(88) = 0.87$, $p < 0.0001$).

3.2.1.4.3 Vowel space size

Measurements (in Hz) of the first and second formant of each vowel were taken from the vowel midpoint 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.

3.2.2 Results

3.2.2.1 Speech onset latencies (RTs)

3.2.2.1.1 Analysis

Because no participant missed more than one comprehension question, data from all participants was analyzed. Exclusions of errors and RTs shorter than 300 ms resulted in removal of 68 observations (5.3% 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 32 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 (Grammatical Constraint or Meaning + Grammatical Constraint), and their interaction. The maximal random effects structure supported by the design (Barr, Levy, Scheepers, & Tily, 2013) included random intercepts for participant and target word, as well as random slopes for noun density and condition by participant, and for condition by target. Interactions were excluded from random effect structure in order to obtain model convergence. Significance was determined by nested model comparison (Barr et al., 2013).

3.2.2.1.2 Results

RT results are summarized in Fig. 3.1. Pictures were named significantly faster in the Meaning + Grammatical Constraint condition than the Grammatical Constraint only condition (β = -0.046, SE = 0.011, $\chi^2(1)$ = 12.84, p = 0.0003). Targets with higher noun density had were named at significantly longer latencies than words with lower noun density (β = 0.006, *SE* = 0.003, $\chi^2(1)$ = 3.85, p = 0.0497). However, there was no reliable interaction between these two factors ($\chi^2(1)$ = 0.43, p > 0.1).

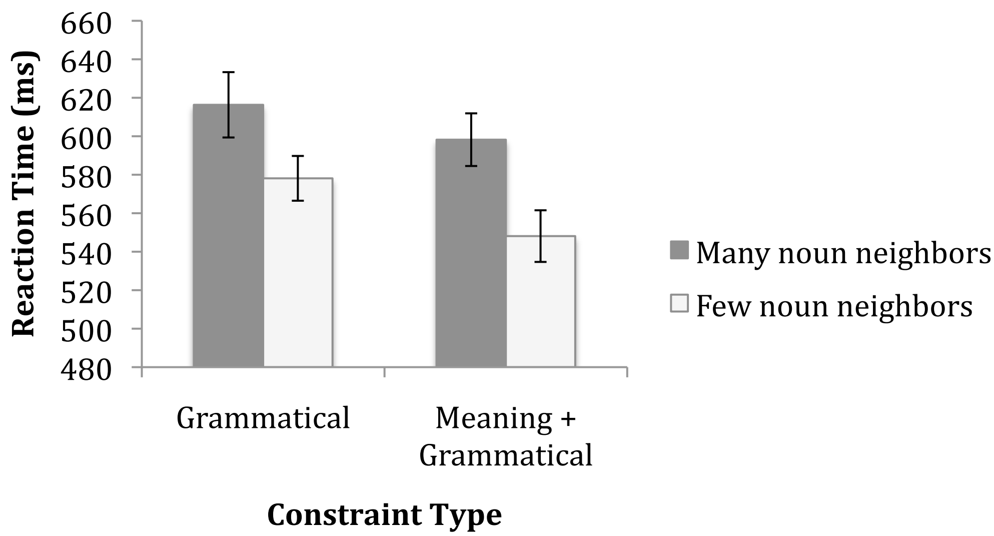

Fig. 3.1. RTs (latencies to speech onset) for words with many and few noun neighbors in Grammatical Constraint and Meaning + Grammatical Constraint contexts in Experiment 1. Error bars indicate by-participant standard error.

3.2.2.2 Vowel duration

3.2.2.2.1 Analysis

Vowel durations were log transformed. One participant's data was excluded from this analysis and subsequent vowel analyses because breathy voice prevented accurate identification of vowel onsets and offsets. Exclusion of errors and model-based outliers removed an additional 88 observations (7.0% of the data).

A linear mixed effects regression was performed, in which fixed effects again included noun density, contrast-coded condition, and their interaction. Vowel identity was included as a contrast-coded control factor. Random effects were identical to those above.

3.2.2.2.2 Results

Fig. 3.2 summarizes the vowel duration results. Targets with higher noun density had significantly shorter vowel durations (β = -0.022, *SE* = 0.009, $\chi^2(1)$ = 4.72, p = 0.0298). There was no reliable difference between vowel durations for targets in Grammatical Constraint versus Meaning + Grammatical Constraint contexts ($\chi^2(1)$ = 0.05; p > 0.1), and no reliable effect of the interaction of condition and noun density on vowel duration ($\chi^2(1)$ = 0.14; p > 0.1). The effect of vowel identity on vowel duration also did not reach statistical significance ($\chi^2(5)$ = 4.64; p > 0.1).

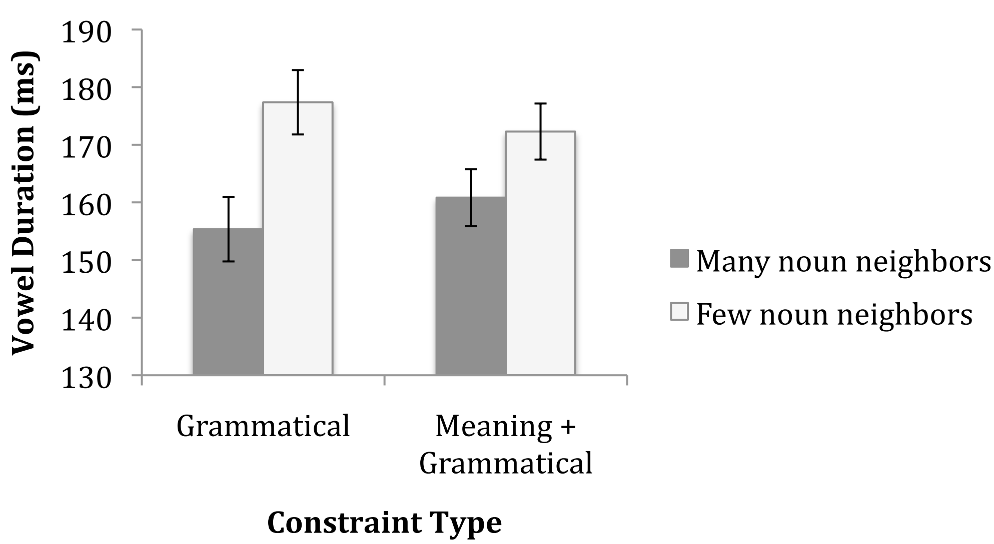

Fig. 3.2. Vowel durations for words with many and few noun neighbors in Grammatical Constraint and Meaning + Grammatical Constraint contexts in Experiment 1. Error bars indicate by-participant standard error.

3.2.2.2.3 Relationship between speech onset latency and vowel duration

To assess whether changes in vowel duration could be attributed to changes in speech planning time, an additional regression was performed. A linear mixed effects model identical to the analysis above, with an additional fixed effect for log RT and random slope for log RT by participant, was fit to the vowel duration data. Again, increased noun density was a significant predictor of decreased vowel duration (β = -0.022, SE = 0.009, $\chi^2(1)$ = 4.84, p = 0.0278). However, RT, was not a significant predictor of vowel duration ($\chi^2(1)$ = 0.17, p > 0.1). As above, constraint condition, the interaction of constraint and noun density (χ^2s(1) < 1, ps > 0.1) and vowel ($\chi^2(5)$ = 5.03, p > 0.1) were not significant predictors of vowel duration.

3.2.2.3 Vowel space size

3.2.2.3.1 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). Removal of errors and outliers excluded 110 observations (8.7% of the data).

3.2.2.3.2 Results

Vowels spaces were smaller for targets in Meaning + Grammatical Constraint contexts than in contexts with Grammatical constraints only (β = -0.052, SE = 0.021, $\chi^2(1)$ = 5.67, p = 0.0173). As vowel duration increased, vowel space size also increased (β = 0.054, SE = 0.019, $\chi^2(1)$ = 7.72, p = 0.0055). However, noun density ($\chi^2(1)$ = 0.38; p > 0.1) and the interaction of noun density and constraint condition ($\chi^2(1) = 0.03$; $p > 0.1$) did not contribute significantly to vowel space size. Vowel identity also reliably predicted vowel space size ($\chi^2(5) = 30.30$; $p < 0.0001$), and females exhibited significantly larger vowel spaces than males ($\beta = 0.096$, $SE = 0.017$, $\chi^2(1) = 26.45$, $p < 0.0001$).

3.2.3 Discussion

In this study, meaning-related constraint was manipulated by manipulating the thematic relation between a context sentence and a target word in a second sentence, which itself had no cues to target meaning but did place grammatical constraints on target identity. Grammatical constraints reliably affected reaction times and vowel durations (which in turn were a significant predictor of vowel space size), as indicated by a significant effect of noun density (while overall neighborhood density was controlled), replicating previous results. Specifically, the more within-category neighbors a target had, the longer the reaction times and shorter the vowel durations. Effects of within-category density on vowel durations could not be accounted for by effects on reaction times, replicating previous work, and suggesting that grammatical constraints on phonological processing continue to affect processing after speech onset.

The presence of a preceding context sentence related in meaning to the target word clearly influenced processing of the target; it sped reaction times and resulted in contracted vowel spaces relative to target productions where contextual cues to target meaning were absent. However, no detectable interaction between meaning-related constraint and grammatical constraint was found for reaction time, vowel duration, or vowel space size. This suggests that although both meaning-related constraint and grammatical constraint affect both speech planning and phonetic encoding, the decrease in response latency and vowel space size associated with higher meaning-related constraint was not attributable to a change in neighborhood activation. Together, these results support cascading activation model of speech production: both grammatical and meaning-related constraints have independent effects on speech planning and phonetic encoding. However, these results are inconsistent with the model under which meaning-related constraints have a large impact on phonological activation during sentence production (e.g., due to the inhibition of unrelated L-level representations). In this type of production context, only grammatical constraints appear to modulate activation of phonological neighbors during speech production.

One may note, however, that the source of meaning-related constraints on target identity in this experiment is relatively distant from the target—it occurs in the previous sentence. One explanation for differences between effects of grammatical and meaning-related constraints on phonological processing might be that only more local cues affect phonological processing. Moreover, target predictability, even in the high meaning-related constraint condition, was not strong. Context sentences in the Meaning + Grammatical Constraint condition were designed to make a target's appearance more likely in the target sentence than its highest frequency noun neighbor (used as a proxy for all neighbors) but did not specifically make the target itself predictable. It is possible that an effect of meaning-related constraint on phonological processing may emerge if stronger meaning-related cues were used. To test these hypotheses, a second study was designed, in which meaning-related constraint was increased (assessed using a test of cloze probability) and made more local (by using single-sentence stimuli rather than sentence pairs). We also included 14 additional target items in this second study, to test whether effects generalize to an expanded stimulus set.

3.3 Experiment 2

<u>3.3.1 Methods</u>

3.3.1.1 Participants

Thirty-four Northwestern University undergraduates (20 female) who were native speakers of American English, aged 18-22, with no history of speech or language deficits, participated in exchange for course credit or $10.

3.3.1.2 Materials

Target words were 34 English CVC nouns (see Appendix B), consisting of the 20 target words used in Experiment 1 and 14 additional items. These items were again divided into two groups differing in within-category (noun) phonological neighborhood density and frequency-weighted noun phonological neighborhood density, and were matched on control factors identical to those used in Experiment 1.

Targets appeared embedded in a single-sentence context, and always appeared in sentence-final position, preceded by a definite determiner. Two versions of this sentence frame were designed for each target word: a sentence for which a target word was either a highly predictable completion (e.g., *When you eat a banana, don't eat the PEEL*; Meaning + Grammatical Constraint) or an unpredictable completion (e.g., *He confessed a peep at the PEEL*; Grammatical Constraint), creating 68 total stimulus sentences. The predictability of sentence-final targets was normed in a separate study using a cloze sentence completion task (see Appendix B for details of stimulus construction and norming). These sentences were divided into two block lists on which half of the items came from each condition and each target appeared only once. These were integrated with an equal number of filler sentences with varying meaning-related constraints on the sentence-final word. Two pseudorandomized orderings were created for each list.

Sentence frames were presented as printed text and targets were presented as colored illustrations. Illustrations were taken from the Rossion and Pourtois (2004) colored line drawing database and supplemented with public domain clip art. Name agreement for these illustrations was normed in a previous study (which examined a total of 37 potential target pictures; see Heller & Goldrick, in press, supplementary materials [Appendix A]).

3.3.1.3 Procedure

Procedure was similar to Experiment 1. Participants were first familiarized with the pictures. Then, participants completed two blocks of sentences, so that every participant saw every item in both conditions. Block presentation order was counterbalanced across participants.

Participants were asked to read aloud each sentence, which 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 appeared and remained until participants clicked the mouse. Twelve comprehension questions were interspersed to encourage holistic sentence processing.

3.3.1.4 Speech Analysis

Trials with errors or disfluencies were excluded from all analyses. Because of the high predictability of targets in the Meaning + Grammatical Constraint condition, many participants began to complete these sentences without waiting for the picture to appear. This provides clear evidence of the effectiveness of the predictability manipulation in increasing the likelihood of the target relative to Experiment 1. Excluding productions beginning earlier than 300 ms after picture presentation resulted in a loss of 567 RT observations, amounting to 25% of total RT observations, or approximately half of the observations in the Meaning + Grammatical Constraint condition. Because of this, RTs are not analyzed here.

Instead, we analyze vowel durations and vowel space sizes in all productions, including those beginning earlier than 300 ms after picture presentation, with the caveat that some productions may not reflect a picture-cued word retrieval process, but rather a word retrieval process cued by information from the sentence context. Vowel duration and vowel space size were determined using the same criteria as in Experiment 1.

3.3.2 Results

3.3.2.1 Vowel Duration

3.3.2.1.1 Analysis

Exclusion of errors and model-based outliers removed 129 observations (5.6% of the data). A linear mixed effects regression analysis of log-transformed vowel durations included noun neighborhood density, contrast-coded condition (Meaning + Grammatical Constraint or Grammatical Constraint), and their interaction as fixed effects of interest. As above, vowel identity was included as a control factor. To control for possible order effects, contrast-coded list order (first versus second block) was also included. The random effects again included random intercepts for participant and word, as well as random slopes for noun density and condition by participant, and for condition by word.

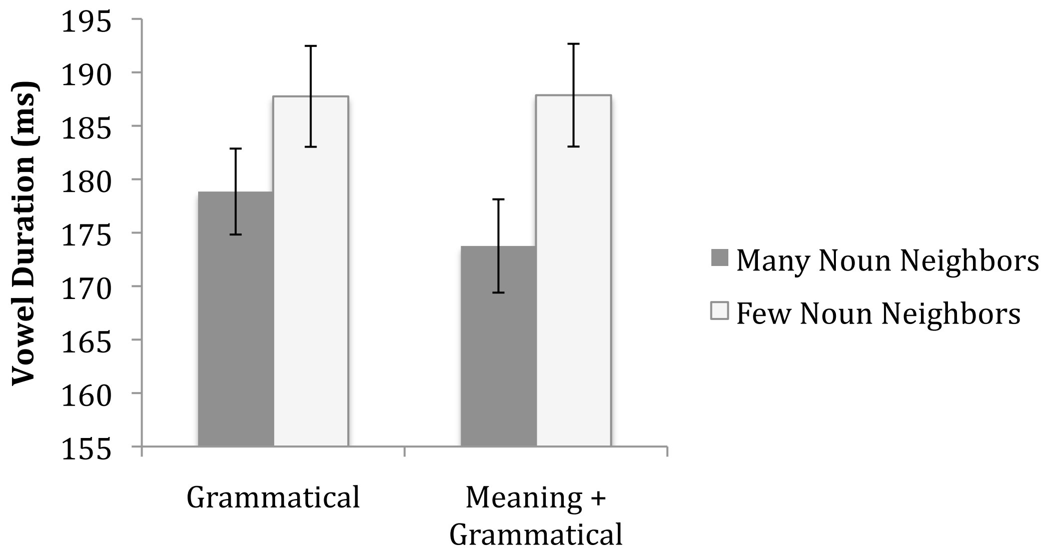

Fig. 3.3. Vowel durations for words with many and few noun neighbors in Grammatical Constraint and Meaning + Grammatical Constraint contexts (collapsing across block) in Experiment 2. Error bars indicate by-participant standard error.

3.3.2.1.2 Results

Fixed effects of interest in the analysis of vowel duration results are summarized in Fig. 3.3. Vowels in target words with higher noun density had shorter vowels (β = -0.013, SE = 0.005, $\chi^2(1)$ = 5.49, p = 0.0191). Again, there was no reliable difference between vowel durations for targets in Grammatical Constraint versus Meaning + Grammatical Constraint contexts ($\chi^2(1) = 0.37$; $p > 0.1$), and no reliable effect of the interaction of condition and noun density on vowel duration ($\chi^2(1) = 2.55$; $p > 0.1$). Participants exhibited reliably shorter vowels in the second block ($\beta = -0.012$, $SE = 0.004$, $\chi^2(1) = 11.49$, $p = 0.0007$). Vowel identity was also a significant predictor of vowel length ($\chi^2(5) = 22.90$; $p = 0.0004$).

3.3.2.3 Vowel space size

3.3.2.3.1 Analysis

A linear mixed effects regression analysis of log-transformed vowel space size was performed including the same factors as the duration analysis, again with additional control factors of speaker gender and log-transformed vowel duration. Removal of errors and outliers excluded 138 observations (6.0% of the data).

3.3.2.3.2 Results

Vowels spaces were smaller for targets in Meaning + Grammatical Constraint contexts than in contexts with Grammatical constraints only ($\beta = -0.020$, $SE = 0.010$, $\chi^2(1) = 4.56$, $p = 0.0327$). Noun density ($\chi^2(1) = 2.05$; $p > 0.1$) and the interaction of noun density and constraint condition ($\chi^2(1) = 0.073$; $p > 0.1$) again failed to reach significance. Vowel duration ($\chi^2(1) = 2.30$; $p > 0.1$) and block ($\chi^2(1) = 0.14$; $p > 0.1$) did not reliably predict vowel space size, but vowel identity ($\chi^2(5) = 77.26$; $p < 0.0001$) and gender ($\beta = 0.096$, $SE = 0.019$, $\chi^2(1) = 21.48$, $p < 0.0001$) did.

3.3.3 Discussion

In this study, meaning-related constraint was increased in strength and came from sources more local than those in Experiment 1. The increase in the number of trials in which participants began producing the correct picture name before picture processing completed (relative to Experiment 1) is an indicator that these manipulations indeed increased meaning-related target predictability for participants.

Overall, results for vowel productions were similar to Experiment 1. Grammatical constraint showed a significant influence on vowel productions; higher noun density was associated with a decrease in vowel duration, and was marginally associated with a decrease in vowel space size even when vowel duration was controlled. The addition of meaning-related constraint was associated with contracted vowel space size. However, no interaction of meaning-related constraint and noun density was detected for either vowel duration or vowel space size.

These results support a cascading activation model; both grammatical and meaning-related constraints affected phonetic outcomes. However, even when meaning-related constraints are relatively strong and local, they do not modulate phonological neighborhood activation. As in the previous experiment, these results are inconsistent with a model under which L-level representations that are not semantically related to the target are inhibited, which would have predicted a larger change in phonological neighborhood effects in highly predictive versus unpredictive contexts.

3.4 General Discussion

Recent research suggests that contextual constraints such as grammatical constraints change the way that phonological encoding proceeds during sentence production; grammatical constraints increase the relative effects of phonological neighbors that share a target's grammatical category. Here, in two experiments, we tested how meaning-related constraints on word identity modulate these effects. This was tested by observing how target words with many versus few within-category (noun) neighbors influenced reaction times when meaning-related constraints from the preceding context were present versus absent. In Experiment 1, meaning-related constraint was implemented as presence versus absence of a thematic relation between a context sentence and a target word in a second sentence that had no cues to word meaning, making the target word more predictable in context than its neighbors. In Experiment 2, meaning-related constraint was strengthened; in single-sentence contexts, targets were either unpredictable or the most likely sentence completion, as assessed in a cloze task.

Results were largely consistent across experiments. When overall neighborhood density was controlled, noun density significantly predicted reaction times and vowel durations in all analyses, such that higher noun density predicted longer reaction times and shorter vowel durations, replicating previous results suggesting that grammatical constraint affects phonological processing by increasing the processing relevance—for both speech planning and phonetic encoding—of words that share a target's grammatical category. Moreover, these effects were not unitary; although noun density was a significant predictor of both RT and vowel duration, changes in RT could not account for changes in vowel duration, suggesting continued effects of grammatical constraints on phonetic encoding after speech initiation. These findings replicate those of Heller and Goldrick (in press).

The addition of meaning-related constraint decreased reaction times and vowel space sizes in all analyses, replicating previous results showing that meaning-related constraint affects speech planning and phonetic outcomes (e.g., Aylett & Turk, 2006; Bell et al., 2009; Griffin & Bock, 1998; Scarborough, 2010). However, the interaction of meaning-related constraint and within-category neighborhood density was not found in any analysis. That is, it is not the case that when cues to target meaning are available in context, thematically-irrelevant neighbors (*HAT, MAT,* etc., for target *CAT* in the sentence *Give a saucer of milk to the CAT*) play less of a role in phonological processing than when only grammatical context is available and meaning-related cues are absent. These results were consistent when meaning-related constraints were relatively weak and distant from the target word (Experiment 1) as well as relatively strong and proximal to the target word (Experiment 2). Results were also consistent when fourteen additional target words were added to the design (Experiment 2).

These results build on previous investigations of semantic-phonological interactions, and shed light on the mechanisms by which increased word predictability affects processing throughout the speech production system during sentence production. While grammatical constraint appears to change the activation of representations that share a target's grammatical class (including the target word itself) at the L-level, affecting lexical-phonological interaction, meaning-related constraint appears to have its largest effect on the target L-level representation itself, at least at the moment of lexical-phonological interaction. This finding is inconsistent with a model under which L-level representations that do not share semantic features with the target are inhibited. In that case, most, if not all, of a target's phonological neighbors would be inhibited (to a larger degree the higher the activation of the target, such as in predictable contexts), causing a larger change in lexical-phonological interaction. In contrast, if semantic activation of L-level representations is only facilitatory—if only the target L-level representation and its semantic neighbors receive facilitatory activation from the semantic level—little effect on the activation of phonological neighbors is predicted. Since few phonological neighbors generally share semantic features with the target word (whereas many phonological neighbors share grammatical features), most neighbors would be unaffected by this semantic activation.

An alternative hypothesis that is consistent with the current data is that because message processing precedes grammatical processing in time, the effects of semantic processing on semantically related L-level representations may have faded by the moment of lexical-phonological interaction. However, such an account would have difficulty explaining the well-documented interactions between semantic and phonological processes during speech production (see Goldrick, 2006, for a review).

3.4.1 Other explanations of phonological neighborhood density effects

It should be noted that at least two other accounts for neighborhood effects in speech production have been put forward. These accounts are based on findings from single word production showing that vowels in words with a greater number of overall phonological neighbors show expanded vowel spaces. Whereas the model discussed above is entirely internal to the production system, both alternative accounts rely on perceptual components. One account holds that this added phonetic extremity is for the listener's benefit (Wright, 2004). Words with more phonological neighbors have more words that sound similar to them, so a speaker's increase in phonetic extremity might ease the listener's task of selecting among these relatively many alternatives. In contrast, when there are fewer neighbors, phonetic extremity is less necessary for the listener to select an appropriate representation; thus, the speaker produces less phonetic extremity. Another account holds that in order for the speaker themselves to develop a phonetic representation of a word, they must first successfully classify instances of the word perceptually (Pierrehumbert, 2002). Again, as listeners, they classify words with fewer competitors more easily, even if they contain less phonetic extremity, whereas more instances of words with more competitors will be successfully classified if they contain more phonetic extremity. Therefore, speakers develop word-specific phonetic representations. For example, the speaker-internal phonetic representation for the quality of the vowel in the word *CAT*, which has many phonological neighbors and therefore many perceptual competitors, will contain more phonetic extremity than the stored vowel quality for *HAT*, which has fewer neighbors in English. The speaker reproduces these differences in perceptual encoding when they produce words based on these stored phonetic representations.

These models nicely account for the evidence they are based on. However, they struggle to account for recent results, replicated here. In particular, results showing that words with more phonological neighbors that share their grammatical category (here, nouns), which may be plausible paradigmatic alternatives in grammatical contexts, are associated with a *reduction* in the phonetic extremity of vowels is problematic for these theories. One would imagine that if a speaker were modeling a listener's needs, the listener would need *more* phonetic information to differentiate between words with many neighbors sharing their grammatical category, rather than less. Indeed, an investigation of the impact of grammatical contexts on speech perception showed that grammatical contexts that reduced the number of phonological neighbors that were grammatically plausible in context sped recognition (Strand, Simenstad, Cooperman, & Rowe, in press). Similarly, because words with more within-category neighbors are more difficult to perceive, one might imagine that these words would be stored with more phonetic extremity under a word-specific-phonetics account. In fact, data are to the contrary.

Because these alternative models could not account for effects of grammatical constraints on phonological processing, we did not pursue them in our investigation of meaning-related constraints on phonological processing here. However, it seems almost certain that a combination of production-internal and listener-modeling systems leads to everyday speech behavior. For example, a speaker will speak louder when addressing a listener in a noisy environment (Lombard, 1911), or use more phonetic extremity when addressing a non-native speaker (Uther, Knoll, & Burnham, 2007). Further research is needed to elucidate the interactions between production-internal and listener-focused modulations to speech (see Scarborough & Zellou, 2013, for recent work in this area).

3.5 Conclusion

Together, the results of these experiments suggest that meaning-related constraint on target word identity in sentences and discourses independently affects target planning, as reflected in latencies to picture naming, and phonetic encoding, as reflected in phonetic properties of vowels in picture names. However, although grammatical constraint changes the way that lexical-phonological interaction proceeds in the phonological neighborhood during sentence production, meaning-related constraints do not reliably affect lexical-phonological interaction over and above effects of grammatical constraint, ruling out a model under which L-level representations that are not semantically related to the target are inhibited concurrent with lexical-phonological activation. These results help refine our knowledge of how different types of contextual predictability affect processing during speech production.

Acknowledgments

We are grateful for support from the National Science Foundation (Grant BCS0846147 awarded to MG) and Northwestern University's Research Grants Committee (Graduate Research Grant awarded to JH). Many thanks to Ariana Steele for acoustic measurement, to Antonia Agbeh, Janice Johnson, and Ariana Steele for assistance running participants, and to the Northwestern SoundLab for helpful comments

CHAPTER 4

Grammatically dependent and grammatically independent effects on phonological processing during spontaneous speech

Abstract

During speech production, words whose forms are phonologically related to a word a speaker is currently uttering—its phonological neighbors—become concurrently activated and influence its processing. Recent laboratory experiments suggest that during sentence production (i.e., when grammatical constraints are relatively high) phonological neighbors that share the grammatical category of a target word (*within-category* neighbors) have an additional influence on word production, when overall number of neighbors is controlled (Heller & Goldrick, in press), suggesting a grammatically dependent component to phonological processing. The current study investigates the relationship between overall neighborhood size and within-category neighborhood size in an analysis of word duration and phonetic properties of vowels in spontaneous speech. Results suggest that overall and within-category neighborhood size have independent effects on word duration and vowel production, supporting both grammatically dependent and -independent components to phonological processing during spontaneous speech production.

4.1 Introduction

In many theories of speech production, words whose forms are phonologically related to a word a speaker is currently uttering become concurrently activated and influence its processing (Dell, 1986; Goldrick, Folk, & Rapp, 2010; Vitevitch, 2002). In investigations of these effects, the number of a word's phonologically related forms is most often approximated by counting the number of words in the lexicon that differ from that word by the addition, deletion, or substitution of one phoneme. These *phonological neighbors* can either inhibit or facilitate target processing (a difference that may be attributable to task demands; see below and Chen & Mirman, 2012, for discussion). The more phonological neighbors a word has—the higher its *neighborhood density*—the longer (Gordon & Kurczek, 2014; Sadat, Martin, Costa, & Alario, 2014) or shorter (Vitevitch, 2002) the latency to speech during naming. Higher neighborhood density has also been associated with greater (Goldrick, Folk, & Rapp, 2010; Gordon, 2002; Harley & Bown, 1998; Vitevitch, 1997, 2002; Vitevitch & Sommers, 2003) or diminished (Newman & German, 2005) the naming accuracy, as well as and more extreme (i.e., expanded F1-F2 vowel space; Munson & Solomon, 2004; Wright, 2004; see Scarborough & Zellou, 2013, for a review) or less extreme (shorter word durations and contracted vowel space; Gahl, Yao, & Johnson, 2012) the phonetic properties.

The above studies have assumed that the mechanisms that give rise to effects of phonological neighbors during word form retrieval and production are context independent—that all of a word's phonological neighbors participate in its processing, regardless of context. For example, during the production of the target word *RUN*, these context-independent mechanisms will activate neighbors such as *FUN*, *RUM,* and *RAIN*, which affect its processing whether it is used as a noun—*the RUN*—or a verb—*to RUN.* However, recent studies of speech production in the laboratory suggest that when the total neighborhood density is brought under experimental control phonological neighbors that share a target's grammatical category (*within-category neighbors*) influence processing when grammatical constraints are relatively strong (Heller & Goldrick, in press). This suggests that grammatically dependent mechanisms also contribute to the activation of neighbors.

Specifically, Heller and Goldrick (in press) investigated how neighbors that can be used as nouns affected processing when pictures representing nouns were named as sentence completions preceded by a definite determiner (e.g., *He regretted the bill for the BOWL*) and when pictures were named in the absence of a sentence context. Experimental items were carefully matched for overall neighborhood density, but differed in the number of those neighbors that could be used as nouns. They found that noun neighbors influenced speech planning and phonetic processing (as assessed via latencies to speech and vowel durations) in sentence contexts but not in the absence of sentence contexts. That is, *FUN* and *RUM*, which can be used as nouns but not as verbs, appear to influence the processing of *RUN* more when it is used as a noun than as a verb, whereas *RAIN*, which can be used as both a noun and a verb, might influence its processing in either context. A greater number of within-category neighbors was associated with slower reaction times and shorter vowel durations in sentence contexts. These effects were subsequently replicated in two follow-up studies (Heller & Goldrick, in preparation b). This suggests a context-dependent (specifically, grammatically dependent) mechanism affecting phonological processing.

Here, we explore the relationship between grammatically independent phonological processing—as indexed by effects of overall neighborhood density—and grammatically dependent phonological processing—as indexed by effects of within-category neighborhood density—in an analysis of word duration and phonetic properties of vowels for content words in spontaneous speech. Utilizing spontaneous speech allows us to extend the investigation of within-category neighborhood density. Our previous studies (Heller & Goldrick, in press, in preparation b) relied on picture naming to elicit target words, limiting our analysis to highly imageable nouns. More naturalistic speech allows us to analyze other types of nouns and content words in other grammatical categories (verbs, adjectives, and adverbs).

This investigation will help us distinguish between several hypotheses about the relationship between grammatically independent and -dependent phonological processing. On the one hand, phonological processing in sentence contexts, such as during spontaneous speech, may be entirely determined by grammatically dependent mechanisms; grammatically independent processing may be reserved for contexts in which grammatical constraints are not present. The fact that previous studies (e.g., Gahl et al., 2012) have observed effects of overall neighborhood density in spontaneous speech does not rule out this account; overall neighborhood density may be highly correlated with within-category neighborhood density. That is, a word with many neighbors may be more likely to have more within-category neighbors, as well. Therefore, previous analyses reporting effects of overall density on processing during spontaneous speech may have been inadvertently detecting effects of within-category density. Consistent with this possibility, Gahl et al. (2012) reported that higher overall neighborhood density is associated with a decrease in phonetic extremity in spontaneous speech—similar to the effect of within-category density within sentence contexts (Heller & Goldrick, in press, in preparation b). If this account is correct, overall neighborhood density should have no reliable effect on word durations, vowel durations, or vowel space sizes when within-category density and other control factors are accounted for.

On the other hand, it is possible that both grammatically independent and -dependent phonological processing mechanisms are at work during sentence production. For example, all of a word's neighbors may play some role in its phonological processing, but within-category neighbors may play an additional (or different) role, beyond that contributed by other phonological neighbors. If this is the case, both overall neighborhood density and within-category density should have independent effects on word durations and vowel productions when the other is brought under statistical control.

In the present study, we used a dataset of monosyllabic CVC content words from the Buckeye Corpus of Conversational Speech (Pitt, Dilley, Johnson, Kiesling, Raymond, Hume, & Fosler-Lussier, 2007). In a previous investigation of this dataset, Gahl et al. (2012) found that higher overall neighborhood density was associated with shorter word durations and smaller vowel space sizes, when other factors known to affect word duration and vowel dispersion were brought under statistical control. In order to investigate the relationship between overall neighborhood density and within-category neighborhood density, we built directly on the analyses of Gahl et al. (2012) by using the same dataset and the same basic model structures, additionally assessing effects, if any, of within-category neighborhood density over and above overall neighborhood density and other control factors. We first investigated content words (nouns, verbs, adjectives, and adverbs) together. Then, in order to assess the influence of within-category density outside of the previously investigated category of nouns, we analyzed the non-noun subset of these words.

The results support both grammatically dependent and -independent components to phonological processing during spontaneous speech. Overall and within-category neighborhood density show independent effects on word duration and vowel productions during spontaneous speech. Specifically, a higher overall neighborhood density was associated with shorter word and vowel durations as well as smaller vowel space sizes, whereas a higher within-category density was associated with longer word and vowel durations.

4.2 Method

4.2.1 The Buckeye Corpus

The Buckeye Corpus of Conversational Speech (Pitt et al., 2007) contains approximately 40-60 minutes of spontaneous speech from each of 40 adult (20 female) participants. All participants were from central Ohio. Half were more than 40 years old.

4.2.2 The Gahl et al. (2012) dataset

Here, we analyze word durations, vowels durations, and vowel space size for vowels in the 9075 total tokens of 414 monomorphemic CVC content word types included in the vowel analysis in Gahl et al. (2012). This is comprised of words containing vowel productions longer than 25 ms from 39 speakers (errors in the transcription for one speaker's recording made its usage problematic). Vowels included /a, æ, ɛ, e, ɪ, i, o, ʊ, u/. An additional analysis of 803 tokens of similar words containing /ʌ/ was completed to calculate the center of each speaker's vowel space.

Word duration, vowel duration, and vowel space size were calculated in a manner identical to Gahl et al. (2012). Word and vowel durations were taken from word- and phone-level segmentation available in the Buckeye Corpus. Vowel spaces were calculated as the Euclidean distance of each token in F1-F2 space from that speaker's vowel space center, which was calculated as the average F1 and F2 for their productions of the central vowel /ʌ/ in monomorphemic CVC content words. Vowel spaces were z-normalized by participant.

4.2.3 Model construction

Linear mixed effects models were fit separately to log-transformed word durations, log-transformed vowel durations, and vowel space size. Models for each data type were based on the final models for word duration and vowel space size in Gahl et al. (2012), and will be described in more detail below. Full models contain as fixed effects all factors included in the final models of Gahl et al. (2012), with the addition of within-category neighborhood density. Values for all numerical factors were centered prior to model construction. In order to assess the influence of overall neighborhood density and within-category neighborhood density using the maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013), we also added random slopes for overall phonological neighborhood density and within-category density by participant, in addition to the random intercepts for word and participant included in previous analyses. Significance of factors was assessed using nested model comparison (Barr et al., 2013).

All control factors discussed were calculated in a manner identical to in Gahl et al. (2012), with the exception of one factor: In all models, we replace the measure of overall neighborhood density used by Gahl et al. (2012) with a highly correlated measure calculated from the SUBTLEX-US database (Brysbaert, & New, 2009; Brysbaert, New, & Keuleers, 2012). This was done because within-category neighborhood density was also calculated from SUBTLEX-US. Details of these calculations are below.

4.2.4 Overall neighborhood density

Overall neighborhood density for each word was defined as the number of monomorphemic (as determined from CELEX; Baayen, Piepenbrock, & van Rijn, 1993) word types in the SUBTLEX-US database that differed from that word by the substitution, addition, or deletion of one phoneme (as determined from the words' CMU Pronouncing Dictionary transcriptions listed in the IPhOD database; Vaden, Halpin, & Hickok, 2009). This measure was chosen over the measure of overall neighborhood density in the analyses of Gahl et al. (2012), which was calculated from the Hoosier Mental Lexicon (Nusbaum, Pisoni, & Davis, 1984), because within-category neighborhood density was calculated using SUBTLEX-US occurrences (see below). The Hoosier Mental Lexicon and SUBTLEX neighborhood density measures were well correlated for the 414 word types included in our dataset, $r(412) = 0.88$, $p < 0.0001$.

4.2.5 Within-category neighborhood density

Part of speech (POS) of each token of each word in the corpus was determined automatically using the CLAWS4 part of speech tagger (Garside & Smith, 1997). This differs from the POS information used in the analyses of Gahl et al. (2012), where each word's most frequent grammatical category in the CELEX database (Baayen et al., 1993) was used for all instances of a word, rather than each token's part of speech in context. As in the Gahl et al. analysis, tokens fell into four grammatical categories: nouns (N = 3036), verbs (N = 4102), adjectives (N = 1419), and adverbs (N = 518). Within-category neighborhood density for each item was calculated as the number of phonological neighbors that appeared in a context where it matched that item's grammatical category in the POS-tagged SUBTLEX-US corpus (also tagged using CLAWS; Brysbaert et al., 2012). The correlation between overall neighborhood density and within-category neighborhood density for tokens in the present dataset was modest but significant, $r(9073) = 0.40$, $p < 0.0001$.

4.3 Results

4.3.1 Word duration

4.3.1.1 Analysis

The final model for word duration in Gahl et al. (2012) included as control factors fixed effects for factors known to influence word duration, including rate of speech (log-transformed speech rate for words respectively preceding and following the current word in the current stretch of speech [syllables per second], as well as a squared factor[5] for speech rate preceding the current word), predictability of the word in context (log-transformed word frequency [from SUBTLEX-US]; log-transformed preceding and following word bigram probability [calculated from the Buckeye corpus]), and other lexical factors (baseline word duration [the sum of the average lengths of each phone in the word, calculated from the Buckeye corpus]; the word's most frequent part of speech [from CELEX]). This baseline model also included overall neighborhood density.

Thus, a linear mixed effects model was fit to log-transformed word durations with the above factors, as well as within-category neighborhood density, as fixed effects. Linear mixed effects model-based outlier trimming removed word durations with residual errors more than 2.5 standard deviations from the mean (Baayen, 2008), excluding 186 observations (2.0% of the data).

4.3.1.2 Results

Results are summarized in Table 4.1. Within-category neighborhood density was a significant predictor of word duration when overall neighborhood density and other model factors were controlled, such that as within-category density increased, so did word duration. As overall neighborhood density increased, word duration decreased.

Effects of all control factors replicated findings of Gahl et al. (2012). As word frequency increased, word duration decreased. Part of speech significantly predicted word

---

[5] Gahl et al. (2012) were interested in exploring nonlinear relationships between predictors and outcomes. Included in the current models are the nonlinear factors that they selected for their models, on the basis of having improved the fit of their models.

duration.; verbs were significantly shorter than nouns, whereas adjectives were longer. Adverbs did not differ significantly in duration from nouns. As baseline duration increased, word duration also increased. As the probability of the current word given the following word increased, word duration decreased, and it also decreased as the probability of the current word given the previous word increased. Word duration decreased with higher speech rates preceding and following the current word. There was also a significant effect of preceding rate squared, such that as squared rate increased, duration decreased.

Table 4.1. Summary of fixed effects in the model of word durations

Variable Name	β	SE	χ^2	$p(\chi^2)$
Frequency	-0.031	0.005	38.36	< 0.0001
POS				
verb (vs. noun)	-0.136	0.040	11.36	< 0.0001
adjective (vs. noun)	0.090	0.046	3.86	0.0496
adverb (vs. noun)	-0.007	0.099	0.01	> 0.1
Baseline Duration	2.617	0.230	111.55	< 0.0001
Bigram After	-0.024	0.001	255.50	< 0.0001
Bigram Before	-0.014	0.002	69.87	< 0.0001
Rate After	-0.137	0.009	255.49	< 0.0001
Rate Before	-0.079	0.008	100.99	< 0.0001
Rate Before^2	-0.030	0.011	6.64	0.0100
Overall density	-0.006	0.001	21.96	< 0.0001
Within-category density	0.003	0.001	7.16	0.0075

4.3.2 Vowel duration

4.3.2.1 Analysis

Although Gahl et al. (2012) did not investigate vowel duration directly, we take as a basis their final model for vowel space size, as an increase in vowel duration is often associated with an increase in vowel space size (Moon & Lindblom, 1994). This model included a linear and a squared factor of overall neighborhood density, factors affecting the phonetic context of the vowel (manner of articulation of the consonant following the vowel; place of articulation for the consonant preceding the vowel), speech rate (a linear and a squared factor for log consonant duration in the word [total word duration minus vowel duration]; log-transformed speech rates [syllables per second] of the stretches of speech directly preceding and following the current word), and factors affecting word predictability in context (linear and squared factors for the log bigram probability of the current word given the preceding word). The Gahl et al. vowel space model also included factors of vowel duration; these are omitted here but included in vowel space analyses below.

A linear mixed-effects model including these factors, as well as within-category neighborhood density, was fit to the log-transformed vowel duration data. Model-based outlier trimming resulted in the exclusion of 187 observations (2.0% of the data).

4.3.2.2 Results

Results for the vowel duration model are summarized in Table 4.2. Within-category neighborhood density was a significant predictor of vowel duration, when overall density and other control factors included in the model were accounted for: As within-category density increased, vowel duration also increased. As overall phonological density increased, vowel duration decreased. When within-category density and overall density were accounted for in the model, overall density squared was not a significant predictor of vowel duration.

Control factors that remained significant showed effects in the expected direction based on results reported in Gahl et al. (2012) for vowel space size. Manner of the following consonant was a significant predictor of vowel duration (approximants led to shorter vowels than obstruents, while nasals led to longer vowels), but place of articulation of the preceding consonant was not. Bigram probability and squared bigram probability of the preceding word also failed to reach significance. As consonant duration in a word and consonant duration squared increased, so did vowel duration. Similarly, as speech rate before and after a word increased, vowel duration decreased.

Table 4.2. Summary of fixed effects in the model of vowel durations

Variable Name	β	SE	χ^2	$p(\chi^2)$
Manner After				
approximant (vs obstruent)	-0.3717	0.0570	40.68	< 0.0001
nasal (vs obstruent)	0.2603	0.0622	17.13	< 0.0001
Place Before	-0.0320	0.0217	2.15	> 0.1
Bigram Before	0.0017	0.0023	0.50	> 0.1
Bigram Before^2	0.0007	0.0009	0.61	> 0.1
Consonant Duration	0.2937	0.0109	690.75	< 0.0001
Consonant Duration^2	0.0616	0.0061	99.86	< 0.0001
Rate After	-0.0516	0.0115	20.03	< 0.0001
Rate Before	-0.0541	0.0106	25.87	< 0.0001
Overall density	-0.0058	0.0025	5.06	0.0245
Overall density^2	-0.0004	0.0003	2.65	> 0.1
Within-category density	0.0058	0.0014	17.73	< 0.0001

4.3.3 Vowel space size

4.3.3.1 Analysis

A linear mixed-effects regression analysis of vowel space size was performed. Again, we use the final model from the Gahl et al. (2012) analysis of vowel space as the basis of our model, with the addition of within-category neighborhood density. This included all factors in the vowel duration model above, with the addition of linear, quadratic, and cubic factors for log-transformed vowel duration. Model-based outlier trimming resulted in the exclusion of 179 observations (2.0% of the data).

4.3.3.2 Results

Results are summarized in Table 4.3. Within-category neighborhood density was not a significant predictor of vowel space size when effects of overall density, vowel duration, and other model factors were controlled. An increased overall density was a significant predictor of decreased vowel space size, as was an increase in squared overall density.

Table 4.3. Summary of fixed effects in the model of vowel space sizes

Variable Name	β	SE	χ^2	$p(\chi^2)$
Manner After				
approximant (vs obstruent)	0.105	0.121	0.75	> 0.1
nasal (vs obstruent)	0.186	0.132	1.98	> 0.1
Place Before	0.189	0.046	16.54	< 0.0001
Bigram Before	-0.010	0.005	4.23	0.0398
Bigram Before^2	0.002	0.002	0.85	> 0.1
Consonant Duration	0.214	0.025	75.93	< 0.0001
Consonant Duration^2	0.075	0.013	30.78	< 0.0001
Rate After	-0.063	0.025	6.41	0.0113
Rate Before	0.035	0.023	2.30	> 0.1
Vowel Duration	0.242	0.032	58.46	< 0.0001
Vowel Duration^2	0.086	0.025	11.54	0.0007
Vowel Duration^3	-0.125	0.033	14.18	0.0002
Overall density	-0.012	0.006	4.49	0.0340
Overall density^2	-0.001	0.001	6.26	0.0124
Within-category density	-0.003	0.003	0.94	> 0.1

Effects of significant control predictors again aligned with results reported in Gahl et al. (2012). Increases in vowel duration and squared vowel duration were associated with an increase in vowel space size, whereas an increase in cubed vowel duration was associated with a decrease in vowel space size. Manner of articulation for the following consonant was not a significant predictor of vowel space size.. The place of articulation of the preceding consonant was a significant predictor, with back consonants associated with larger vowel spaces than front consonants. The higher the probability of the current word given the previous word, the smaller the vowel space; however, the quadratic factor failed to reach significance. Longer consonant durations and squared consonant durations were associated with larger vowel spaces. A higher speech rate in the stretch of speech following a word was associated with a smaller vowel space size; however, speech rate preceding a word did not reliably predict vowel space size.

4.3.4 Within-category density beyond nouns

4.3.4.1 Analysis

In order to ascertain whether effects of within-category density observed here depend on effects on the previously investigated category of nouns, we refit the models of word duration, vowel duration, and vowel space size described above to the subsets of word tokens that were not nouns. The non-noun subset was comprised of 6039 verb, adverb, and adjective tokens. The correlation between overall density and within-category density in this subset was again modest but significant, $r(6037) = 0.35$, $p < 0.0001$. An investigation of the subset of noun tokens present in the corpus revealed an extremely high correlation between overall density and within-category density in this subset, $r(3034) = 0.97$, $p < 0.0001$. Because the extreme correlation between overall density and noun density in this subset makes them impossible to tease apart statistically, a model was not assessed for nouns only.

4.3.4.2 Results

Increased within-category density was a marginally significant predictor of an increase in word duration for the subset of tokens that were verbs, adjectives, and adverbs (β = -0.004, SE = 0.002, $\chi^2(1)$ = 3.33, p = 0.0679). It was a significant predictor of vowel duration for the non-noun subset; as in the entire dataset, increased within-category density was associated with an increase in vowel duration (β = -0.007, SE = 0.003, $\chi^2(1)$ = 6.15, p = 0.0131). Again, within-category density was not a significant predictor of vowel space size in this subset ($\chi^2(1)$ = 0.32, p > 0.1).

4.4 Discussion

Most previous investigations of phonological processing during speech production, both in laboratory and spontaneous speech, have assumed grammatically independent activation of a target's phonological neighbors. However, recent laboratory experiments suggest the existence of a grammatically dependent phonological processing mechanism. In sentence contexts, phonological neighbors that share a target's grammatical category affect processing when overall phonological density is controlled. Do these two mechanisms play independent roles in spontaneous speech, or is speech production in fluent contexts dominated by grammatically dependent processing mechanisms? To address this, we examined phonetic outcomes (word durations, vowel durations, and vowel space sizes) in spontaneous content word productions when other factors known to affect phonetic outcomes were controlled. Grammatically independent phonological processing mechanisms were investigated via effects of overall neighborhood density on phonetic outcomes, and grammatically dependent phonological processing mechanisms were investigated via effects of within-category neighborhood density. The results suggest that both overall and within-category neighborhood size have independent effects on word durations and vowel productions, supporting both grammatically dependent and grammatically independent components in phonological processing during spontaneous speech. Specifically, both overall neighborhood density and within-category neighborhood density were significant predictors of word and vowel durations when the other was controlled. However, overall density, but not within-category density, was a significant predictor of vowel space size when the effect of vowel duration was controlled. As such, we see effects of both overall neighborhood density and within-category density on phonetic processing both at the whole-word level and in phonetic properties of vowels.

These results provide further evidence for grammatical influences on phonological processing. Although effects of within-category density had previously only been investigated in two studies of highly imageable nouns, we show here that effects of within-category neighbors hold in a larger and more diverse dataset, including a more varied set of nouns, as well as verbs, adjectives, and adverbs. Results were similar when noun tokens were removed from analysis, although effects on word duration were less reliable. Moreover, grammatically dependent mechanisms influence phonological processing not only in carefully controlled laboratory experiments, but also during everyday linguistic processing, such as during spontaneous speech production.

Across analyses in the current paper, overall neighborhood density and within-category density were associated with opposite effects. Whereas an increase in overall density was associated with a decrease in word duration and vowel space size (replicating the results of Gahl et al., 2012) an increase in within-category density was associated with an increase in word and vowel durations, moderating the effect of overall density. Opposing effects of overall phonological density and within-category density are in line with available evidence from laboratory speech studies; as reviewed above, higher overall density is associated with an increase in the extremity of vowel productions (e.g., Wright, 2004), whereas higher within-category density is associated with a decrease in the extremity of vowel productions (e.g., Heller & Goldrick, in press). A summary of effects of increasing overall neighborhood density and within-category density in the current investigation and other investigations can be viewed in Table 4.4.

Table 4.4. Summary of effects of increasing overall and within-category density across studies

	Increase in overall neighborhood density		Increase in within-category neighborhood density	
Spontaneous speech				
Word duration	Decrease	Present paper; Gahl et al. (2012)	Increase	Present
Vowel duration	Decrease	Present	Increase	Present
Vowel space size	Decrease	Present; Gahl et al. (2012)	No effect	Present
Laboratory speech				
Vowel duration	No effect	e.g., Munson & Solomon (2004)	Decrease	Heller & Goldrick (in press, in prep b)
Vowel space size	Increase	e.g., Munson & Solomon (2004)	No effect	Heller & Goldrick (in press, in prep b)

Together, the data presented in Table 4.4 pose challenges for all current theoretical accounts of neighborhood density effects in speech production. Theories that rely on perceptual principles (e.g., Pierrehumbert, 2002; Wright, 2004), designed to account for overall neighborhood density effects found in laboratory speech, state that words with higher density are produced with expanded vowel spaces because words with more perceptual competitors are difficult to perceive and categorize relative to words with fewer competitors. Therefore, increased phonetic extremity is utilized to ease this perceptual process in contexts where more competitors exist. Recent results showing decreased phonetic extremity associated with higher within-category density in laboratory speech, as well as the reduction of phonetic extremity associated with higher overall density in spontaneous speech, are inconsistent with these theories. Moreover, production-internal theories (e.g., Chen & Mirman, 2012; Vitevitch, 2002), which posit that phonologically related forms become active during the production process and affect target activation, struggle to account for flipped effects of both overall and within-category density between laboratory and spontaneous speech contexts. Chen and Mirman's (2012) model of inhibitory and facilitatory effects of neighbors posits that when neighbors become highly activated they incur a net inhibitory effect on target processing. For this model to account for the data, outside-category neighbors would need to be boosted in activation relative to within-category neighbors—a difference that is not only intuitively unlikely but also inconsistent with other findings suggesting an advantage for within- vs. outside-category neighbors during sentence production (e.g., Berndt et al., 1997). As such, the current results, in combination with previous findings, call for major modifications to current theories of phonological processing during speech production. Further investigation of this topic is greatly needed.

4.5 Conclusion

Words whose forms are phonologically related to words currently being uttered become concurrently activated and influence processing. Results of the current investigation suggest that both grammatically independent factors—that is, all of a word's phonological neighbors—and grammatically dependent factors—neighbors that match a word's grammatical category in context—have independent (and opposite) effects on phonological processing during spontaneous speech production. While the current investigation replicates previous findings regarding the effects of overall neighborhood density on phonetic outcomes in spontaneous speech, it also extends previous findings regarding the effects of within-category neighborhood density from a small set of items in laboratory speech to a more diverse spontaneous speech dataset, highlighting differences between laboratory and spontaneous production processes. These results help refine our knowledge of the degree to which contextually dependent grammatical processing affects phonological processing during speech production.

Acknowledgments

We are deeply indebted to Yao Yao, Susanne Gahl, and Keith Johnson for sharing the dataset analyzed in Gahl et al. (2012). 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 JH), and thank Northwestern's Sound Lab for helpful feedback.

CHAPTER 5

Conclusion

5.1 Introduction

A great deal of research has shown that during speech production many representations in addition to those of the current word in an utterance are also activated. Representations that are phonologically and/or semantically related to the current word, as well as other representations corresponding to other words planned in the utterance become concurrently active. The current project has investigated the extent to which the predictability of a word from contextual cues affects this co-activation. Specifically, I have investigated how grammatical constraints and meaning-related constraints from the utterance context affect the activation of phonological neighbors during phonological processing.

Effects of contextual constraints on phonological neighborhood activation on speech planning (assessed via latency to speech onset) and phonetic encoding (assessed via phonetic properties of vowels) were assessed using a picture naming paradigm in the laboratory (Chapters 2 and 3). The results of these studies suggest that grammatical constraints, but not meaning-related constraints, reliably affect phonological neighborhood activation during speech production. The relationship of this grammatically dependent phonological processing mechanism to previously investigated grammatically independent phonological processing mechanisms was investigated in an evaluation of word duration and vowel properties in a corpus of spontaneous speech (Chapter 4). The results of this investigation suggest that both grammatically dependent and -independent phonological processing mechanisms are present during spontaneous speech. Results are summarized below, followed by a discussion of their theoretical implications and suggestions for further possible avenues for investigation.

5.2 Summary of current investigations

5.2.1 Grammatical constraints on phonological neighbor activation (Chapter 2)

Results from the speech error literature suggest that word selection is mutually influenced by grammatical and phonological information: speech errors that are phonologically related to the word the speaker intended to utter match the grammatical category of the intended word more often than would be predicted by chance (see Goldrick, Folk, & Rapp, 2010, for a recent review). Here, I compared two hypotheses about the nature of this mutual influence. Under one account, activation of phonological neighbors could be influenced by a grammatical processing mechanism in a context-dependent manner. Alternatively, phonological processing could proceed in a grammatically independent manner, and selected representations could be subsequently checked by a monitoring mechanism to ensure their contextual appropriateness.

In order to assess the influence of grammatical constraints on phonological neighbor activation during speech production, I compared how neighbors that share a target's grammatical category (in this study, nouns) influenced its planning and phonetic encoding when grammatical constraints were strong or weak. Participants named pictures embedded in grammatically (but not thematically) constraining sentence contexts (where grammatical constraints were strong) and in a bare picture naming task (where grammatical constraints were weak). Picture names were carefully matched for overall neighborhood density but differed in the number of phonological neighbors that shared the category of noun. Latencies to speech onset (RTs) and phonetic properties of vowels were evaluated for target picture name productions in both conditions.

Within-category (noun) neighbors influenced planning time and phonetic encoding more strongly in sentence contexts. Specifically, an increase in within-category density was associated with an increase in RT and a decrease in vowel duration in sentence contexts. This suggests that grammatical encoding constrains phonological processing; the influence of phonological neighbors is grammatically dependent.

5.2.2 Meaning-related constraints on phonological neighbor activation (Chapter 3)

Further results from the speech error literature suggest that semantic and phonological information also mutually affect lexical selection: mixed errors, which are both phonologically related and semantically related to the word that the speaker intended to utter, are more commonly produced than would be expected by chance (see Goldrick, 2006, for a review). Here I contrasted two hypotheses about how this mutual influence may play out during speech production. Under one account, words that are semantically related to the speaker's intended word are boosted in activation. During lexical-phonological processing, this would increase the influence of phonological neighbors that are also semantic neighbors. However, because few words are both lexical and phonological neighbors, this would result in only a small change in the phonological activation of the target. Under a second account, words that are not semantically related to the speaker's intended word are inhibited. This would result in a larger change in phonological activation, since many more of a word's phonological neighbors would be affected. When meaning-related predictability of the target is high given its context, semantic activation would be increased, magnifying these effects.

In order to differentiate between these hypotheses, I compared how sentence and discourse contexts that had stronger or weaker meaning-related constraints on target identity affected the influence of neighbors over and above effects of grammatical constraints. In two experiments, participants named pictures embedded in one- or two-sentence contexts in which targets were predictable to a greater or lesser extent given meaning-related cues in these contexts. In Experiment 1, meaning-related constraint was implemented as presence versus absence of a thematic relation between a context sentence and a target word in a second sentence that had no cues to word meaning, making the target word more predictable in context than its neighbors. In Experiment 2, meaning-related constraint was strengthened; in single-sentence contexts, targets were either unpredictable or the most likely sentence completion, as assessed in a cloze task. In the first experiment, targets were identical to those analyzed in the assessment of grammatical constraints, above, and in the second experiment, an expanded stimulus list was used. RTs and phonetic properties of vowels were evaluated for target picture name productions in conditions with high and low meaning-related constraint.

The results suggest that while grammatical processing constrains neighbor activation, meaning-related constraints do not reliably affect the activation of phonological neighbors. When overall neighborhood density was controlled, higher noun density was associated with longer reaction times and shorter vowel durations, replicating results presented in Chapter 2. The addition of meaning-related constraint decreased reaction times and vowel space sizes. However, meaning-related constraint did not modulate the effects of within-category neighbors. These results were consistent when meaning-related constraints were relatively weak and distant from the target word (Experiment 1) as well as relatively strong and proximal to the target word (Experiment 2), as well as when fourteen additional target words were added to the design (Experiment 2). These results suggest that rather than significantly affecting phonological neighbor activation, meaning-related constraints have their largest effect on the target representation itself. Because few phonological neighbors are also semantic neighbors, these results are inconsistent with a model under which representations that do not share semantic features with the target are inhibited.

5.2.3 Relationship between grammatically dependent effects on RTs and phonetic properties

Previous investigations of effects of phonological neighbors and predictability have focused on their effects on either speech planning (RTs) or phonetic processing. Here, I have examined both, as well as their relationship. Analyses of the relationship between effects of grammatically dependent phonological processes on RTs and phonetic properties of vowels in both Chapters 2 and 3 reveal that changes in phonetic properties of vowels cannot be accounted for by changes in speech planning times. Across analyses of this relationship, within-category density was a significant predictor of vowel duration when RT was controlled. Moreover, RTs did not exert a reliable influence on vowel durations when within-category density was accounted for. These results suggest that these effects are separable; neighbor activation continues to affect speech production after speech has been initiated.

5.2.4 Relationship between grammatically dependent and –independent effects (Chapter 4)

The results discussed above support the existence of a grammatically dependent phonological processing mechanism. In order to assess whether phonological processing in sentence contexts, such as during spontaneous speech, is entirely dependent on grammatically dependent mechanisms, or whether grammatically dependent and grammatically independent mechanisms play simultaneous independent roles, I compared effects of total neighborhood density and within-category neighborhood density on word duration and phonetic properties of vowels in a corpus of spontaneous speech. When both factors were brought under statistical control, they exerted independent (and opposite) effects. While total neighborhood density was associated with shorter word durations and reduced vowels, within-category density was associated with longer word and vowel durations. These analyses support the existence of both context-dependent (specifically, grammatically dependent) and context-independent components to phonological processing during spontaneous speech production.

5.2.5 Relationship between neighborhood effects in laboratory and spontaneous speech

Previous research has reported differential effects of overall neighborhood density on phonetic properties of vowels in laboratory and spontaneous speech. In laboratory speech, an increase in overall neighborhood density is typically associated with an increase in phonetic extremity of vowels (Kilanski, 2009; Munson, 2007; Munson & Solomon, 2004; Scarborough, 2010, 2012, 2013; Scarborough & Zellou, 2013; Wright, 2004), while in spontaneous speech, decreases in phonetic extremity have been reported (Gahl, Yao, & Johnson, 2012). Here, we also find opposing effect directions with respect to within-category neighbors in laboratory and spontaneous speech productions. An increase in within-category density was associated with a decrease in vowel duration in laboratory speech (Chapters 2 and 3), but with an increase in vowel duration in spontaneous speech (Chapter 4). Although these effects are consistent with the opposing directions across tasks previously reported for overall phonological neighborhood density, they present challenges for existing theories of phonological processing designed to account for phonological neighborhood effects. These will be discussed below.

5.3 Implications for theories of neighborhood effects in speech production

The results from the current investigations pose challenges for all current theories of phonological neighborhood effects. Implications for the three major theories designed to account for these effects will be discussed below. These theories fall into two classes: perception-driven and production-driven. Implications will be discussed for the perception-driven models first, followed by a discussion of the implications for production-driven models, from which hypotheses for the current investigations were derived.

5.3.1 Implications for perception-driven theories of phonological neighborhood effects

Two of the three major theories designed to account for phonological density effects include perceptual components. The *listener modeling* theory (e.g., Scarborough, 2010; Scarborough & Zellou, 2013; Wright, 2004) is based on the idea that words with more phonological neighbors are more difficult to identify during perception by virtue of having relatively many similar-sounding competitors. As such, a speaker might increase phonetic extremity as the number of these similar-sounding competitors increases in order to aid the listener in the selection of the appropriate representation. The *exemplar* account of phonological neighborhood effects (Pierrehumbert, 2002) is based on the idea that the speaker is also a listener. This account posits that speakers must first successfully classify instances of a word perceptually in order to develop phonetic representations for that word. Because words with fewer competitors can be identified more easily with less phonetic extremity, they will be encoded in the listener-speaker's mind with less phonetic extremity. Conversely, because instances of words with more competitors can only be successfully classified if they contain more phonetic extremity, they will be encoded with more phonetic extremity. These differences in perceptual encoding, then, are reflected in the listener-speaker's subsequent word productions.

These models were designed to account for data from laboratory studies showing that vowels in words with a greater number of overall phonological neighbors show expanded vowel spaces. However, it is not clear how they could be modified to account for the current results with respect to grammatical and meaning-related constraints. Because words with many within-category neighbors have many competitors that can fulfill the same function in a sentence, listeners' ability to perceive words should be negatively impacted by having large numbers of within-category neighbors. Indeed, a recent investigation of speech perception showed that perception was facilitated by grammatical contexts that reduced the number of phonological neighbors that were grammatically plausible (Strand, Simenstad, Cooperman, & Rowe, in press). Perception-driven theories would therefore predict that words with more versus less within-category neighbors should be produced with more phonetic extremity. In contrast, data presented in Chapters 2 and 3 showed that, in laboratory speech, higher within-category density is associated with shorter vowel durations is inconsistent. Following similar logic, data presented in Chapter 4, showing that higher overall neighborhood density is associated with shorter word durations and less phonetic extremity in spontaneous speech also challenges these theories. Moreover, these data cannot account for the flipping of effect directions for overall and within-category density in laboratory versus spontaneous speech; although these theories might predict variation in the degree to which listeners hyperarticulate across tasks, density is predicted to always be associated with more phonetic extremity.

Another challenge for the perception-based theory is the lack of effects of meaning-related constraints on phonological neighbor activation. In highly predictable contexts, less phonetic detail is required for listeners to correctly identify a word (e.g., Lieberman, 1963). As such, this theory would predict diminished influence of phonological neighborhood density on phonetic extremity in highly predictable contexts. However, although we see an overall reduction in phonetic extremity for words in more predictable contexts, the current results show that predictability effects do not interact with effects of within-category density—it is not the case that phonological neighbors influence processing less when the target is predictable, and therefore easily identifiable, in its context.

Taken together, the current results are wholly inconsistent with perception-driven accounts of phonological neighborhood density. It should be noted, however, that listener-modeling theories account for many speech phenomena, and as such should not be wholly discounted. For example, speakers produce louder speech in noisy environments (Lombard, 1911), and use more phonetic extremity when addressing non-native listeners (Uther, Knoll, & Burnham, 2007). Indeed, in a recent investigation of phonological neighborhood effects in speech perception and production, Scarborough and Zellou (2013) showed that, when an interlocutor was present, speakers produced more phonetic extremity than in the absence of an actual listener. Furthermore, this increase in phonetic extremity indeed made words of high overall density, which would otherwise be relatively difficult to perceive, easier to perceive for listeners. As such, it appears that listener modeling mechanisms can be utilized to a greater or lesser extent depending on context and task demands. Therefore, it is beyond doubt that modulations based on listener modeling act in concert with those mechanisms that produce density-related effects in order to create everyday speech behavior.

5.3.2 Implications for production-internal theories of phonological neighborhood effects

Hypotheses tested in the current investigations have been based on a production-internal account of phonological neighborhood effects. This account holds that phonological neighbor representations become active during word production due to feedback between phonological and L-level representations. This concurrent activity is thought to modulate both speech planning (Vitevitch, 2002) and phonetic outcomes (Baese-Berk & Goldrick, 2009). When neighbors are relatively weakly activated, they contribute net facilitation to target processing via a lexical-phonological feedback mechanism, but when they are strongly activated relative to target activation, they exert a net inhibitory influence on target processing via inhibitory connections between L-level representations (Chen & Mirman, 2012). Chen and Mirman (2012) speculate that neighbors are often facilitatory during speech production (but c.f., e.g., Gordon & Kurczek, 2014; Sadat, Martin, Costa, & Alario, 2014, for recent evidence that neighbors are not necessarily facilitatory during production) because target activation is relatively high compared to the activation of neighbors due to planning at the message level, which increases target activation.

First, results suggesting a grammatically dependent phonological processing mechanism imply that some mechanism must exist that increases the relative contribution of within-category neighbors during target processing. This is consistent with proposals suggesting a “syntactic traffic cop” mechanism, dynamically directing activation to words of appropriate grammatical categories as sentence production proceeds (Dell, Oppenheim, & Kittredge, 2008; Gordon & Dell, 2003). Moreover, the finding that within-category density had separable effects on RTs and vowel durations suggests that this dynamic activation cascades down into phonetic processing even after the production of a word has been initiated.

More specifically, under the account outlined by Chen and Mirman, current results (Chapters 2 and 3) showing that a higher within-category density is associated with an increase in reaction time and a decrease in vowel duration suggest that this grammatical encoding mechanism boosts the activation of within-category neighbors (rather than decreasing the activation of out-of-category neighbors), pushing these representations across the activation threshold from facilitation into net target inhibition.

Results indicating that meaning-related constraints do not reliably change the influence of within-category phonological neighbors are also consistent with a model under which semantically-related forms receive increased activation, rather than unrelated forms receiving inhibition. However, because so few phonological neighbors are also semantically related to an average target word, effects of this increased activation may be difficult to detect.

However, the findings from investigations of spontaneous speech (Chapter 4; Gahl et al., 2012) are more difficult for this theory to account for. Under this model, it is entirely possible to account for the idea that within-category density and overall density have separable, opposite effects—inhibition and activation are passed through different avenues in Chen and Mirman's model (inhibition between L-level representations, activation via lexical-phonological feedback). However, it is extremely difficult to explain why overall neighbors would cause net target inhibition (to the extent that the link between long reaction times and short vowels in the laboratory investigations presented here is processing difficulty, and can be linked to the current data), but within-category neighbors would contribute net facilitation. For this model to account for the data, outside-category neighbors would need to be boosted in activation relative to within-category neighbors—a difference that is not only intuitively unlikely but also inconsistent with other findings suggesting an advantage for within- vs. outside-category neighbors during sentence production (e.g., Berndt et al., 1997). Alternatively, this model would need to postulate a mechanism that allows the flipping of effect direction from one speech task (e.g., speech cued by laboratory materials) to another (e.g., speech produced spontaneously) that is separable from the relative activation of relevant representations. Overall, this theory of phonological neighborhood effects is the most promising account of the data presented in the current project, but major modifications will need to be included in order to account for the spontaneous speech data.

5.4 Avenues for future research

The current investigation has provided data crucial to our understanding of how context-related activation affects activation flow throughout the speech production system. However, several issues remain outstanding, and the data presented here raises new questions. Further investigation of context-dependent phonological processing will help refine our understanding of the speech production system and guide the construction of improved models.

5.4.1 Delving into grammatically dependent phonological processing mechanisms

5.4.1.1 Assessing the impact of continuous grammatical predictability

In the current investigation, grammatical constraints were operationalized broadly. In Chapters 2 and 3, I observed properties of word production when a noun was being processed in an environment that was predictive of a noun The appearance of a definite determiner (e.g., in the stimulus sentence *He confessed a peep at the PEEL*) was utilized as a cue to subsequent noun appearance. However, one may note that the appearance of a definite determiner is also consistent with subsequent adjective appearance (e.g, *the slippery peel*) or, less frequently, adverb appearance (*the very slippery peel*). Is the grammatical effect on phonological processing stronger when there is more certainty about the grammatical category of an element before it appears? The effect could also arise from category of most frequent use, or of current use, regardless of its grammatical predictability. These hypotheses could be tested using the paradigm used in the present laboratory experiments (picture naming in a sentence context), varying the strength with which the previous portion of the sentence predicts noun appearance. This could be additionally investigated (via an investigation of vowel properties) in a corpus of speech.

5.4.1.2 Comparing grammatically dependent and –independent phonological processing in laboratory speech

In the laboratory studies conducted in the present investigation, effects of grammatically dependent phonological processing were tested while effects of grammatically independent phonological processing were brought under careful experimental control. The relationship between these two processing mechanisms was investigated in spontaneous speech, rather than in the laboratory. However, spontaneous speech can be highly statistically noisy, due to many uncontrolled factors. Because of the somewhat unexpected (at least with respect to current theories) results of the corpus investigation, it may be helpful to assess the relationship between grammatically dependent and –independent processing mechanisms by designing a laboratory experiment in which target words vary in overall and within-category density independently. This will give a clearer view of this relationship, and may help illuminate a relationship that had been difficult to identify in the spontaneous speech data.

5.4.1.3 Cross-language validation and separation of overall and within-category density

The current project investigated only effects of contextual constraints on phonological processing in English. Although in the laboratory studies it was possible to design a set of noun stimuli in which within-category density varied while overall density was controlled, the high correlation between overall density and within-category density in the noun subset of the corpus data suggests that these factors are generally highly correlated among nouns in English. This high correlation is consistent with the fact that English is quite permissive as to which words may be used as nouns. The investigation of the relationship between overall neighborhood density and within-category density (as well as the assessment of the importance of this relationship to speech production processing as a whole) may be further illuminated by the investigation this relationship in a language that is not so grammatically permissive (which, in turn, may strengthen grammatical constraints), where these two factors may be more easily separated, and where their impact may be more easily visible in the wild. If effects differ across languages, this may make strides towards unifying differences recently reported for, e.g., overall neighborhood effects on picture naming times in Spanish (Sadat, Martin, Costa, & Alario, 2014) versus English (Vitevitch, 2002).

5.4.2 Delving into meaning-related influences on phonological processing

5.4.2.1 Investigating a facilitatory effect of semantic predictability on phonological processing

Results from the present investigation of meaning-related constraints on phonological processing showed that meaning-related target predictability did not have a reliable impact on the effect of phonological neighbors. These results are not consistent with a model under which neighbors that are not semantically related to the target word are inhibited during speech production. As noted above, they are more consistent with a model under which neighbors that are semantic neighbors of the target receive a boost in activation; however, the current study has not provided positive evidence for such facilitation of non-target representations. A study similar to those performed here could be designed to test this possibility, systematically varying the number of neighbors a target has that fulfill all of the following criteria: 1) share a grammatical category, 2) are phonological neighbors, and 3) are semantic neighbors. This carefully designed study may help provide evidence for (or against) the hypothesis that semantic information and phonological information do indeed mutually influence processing in non-errorful speech in a facilitatory manner.

5.5 Conclusion

The current project has investigated the extent to which the predictability of a word from contextual cues affects co-activation of representations in the lexicon that are grammatically, semantically, and phonologically related to a word a speaker is currently uttering. Specifically, I have investigated how grammatical constraints and meaning-related constraints from the utterance context affect the activation of phonological neighbors during phonological processing. The results suggest that grammatical constraints, but not meaning-related constraints, reliably affect phonological neighbor activation in sentence contexts. Phonological neighbors that share the grammatical category of the target influenced planning time and phonetic encoding more strongly in sentence contexts, supporting a grammatically dependent phonological processing mechanism. However, meaning-related constraints do not reliably affect the activation of phonological neighbors over and above effects of grammatical constraints, suggesting a far smaller role for meaning-related constraint on phonological processing. When grammatically dependent and -independent mechanisms were compared via an investigation of overall and within-category phonological neighborhood density in a corpus of spontaneous speech, independent roles were found for both processing mechanisms.

This evidence bears on theories of phonological processing during speech production and theories about the manner in which context-dependent predictive processes modulate activation during speech production. Specifically, this work provides evidence about how the speech production system leverages grammatical and meaning-related constraints to streamline production. Failures of perception-based theories of phonological neighborhood effects to account for the current data highlight differences in the treatment of predictability in the perceptual and productive systems. Revisions to current theories accounting for this data will make great strides towards the goal of accounting for production behavior during everyday speech.

REFERENCES

Alario, F.-X., Matos, R. E., & Segui, J. (2004). Gender congruency effects in picture naming. *Acta Psychologica, 117,* 185-204.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence and duration in spontaneous speech. *Language and Speech, 47,* 31–56.

Aylett, M., & Turk, A. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *Journal of the Acoustical Society of America, 119,* 3048–3058.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R.* New York: Cambridge University Press.

Baayen, H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database* (CD-ROM).

Baese-Berk, M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes, 24,* 527-554.

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,* 255-278.

Bell, A., Brenier, J., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language, 60,* 92–111.

Berndt, R. S., Mitchum, C. C., Haendiges, A. H., & Sandson, J. (1997). Verb retrieval in aphasia. *Brain and Language, 56,* 68-106.

Bock, K. (1989). Closed-class immanence in sentence production. *Cognition, 31,* 163-186.

Boersma, P., & Weenink, D. (2012). Praat: Doing phonetics by computer [Computer program]. Version 5.3.15, retrieved 6 April 2013 from http://www.praat.org/

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *Journal of the Acoustical Society of America, 121,* 2339-2349.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41,* 977-990.

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods, 44,* 991–997.

Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: Toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological Review, 119,* 417-430.

Clopper, C. G., & Pierrehumbert, J. B. (2008). Effects of semantic predictability and regional dialect on vowel space reduction. *Journal of the Acoustical Society of America, 124,* 1682–1688.

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers, 36,* 384-387.

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language, 42,* 465-480.

Dahan, D., & Tanenhaus, M. K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 498–513.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93,* 283-321.

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review, 104,* 123-147.

Dell, G. S., Oppenheim, G. M., & Kittredge, A. K. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language and Cognitive Processes, 23,* 583-608.

Ford, M., & Holmes, V. M. (1978). Planning units and syntax in sentence production. *Cognition, 6,* 35-53.

Fowler, C. A., & Housum, J. (1987). Talkers' signaling of ''new'' and ''old'' words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language, 26,* 489–504.

Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language, 66,* 789-806.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 133-175). San Diego: Academic Press.

Garside, R., & Smith, N. (1997). A hybrid grammatical tagger: CLAWS4. In R. Garside, G. Leech, & A. McEnery (Eds.), *Corpus Annotation: Linguistic Information from Computer Text Corpora* (pp. 102-121). London: Longman.

Gillespie, M., & Pearlmutter, N. J. (2011). Hierarchy and scope of planning in subject-verb agreement production. *Cognition, 118,* 377-397.

Goldrick, M. (2006). Limited interaction in speech production: Chronometric, speech error, and neuropsychological evidence. *Language and Cognitive Processes, 21,* 817-855.

Goldrick, M., Folk, J. R., & Rapp, B. (2010). Mrs. Malaprop's neighborhood: Using word errors to reveal neighborhood structure. *Journal of Memory and Language, 62,* 113-134.

Gordon, J. K. (2002). Phonological neighborhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain and Language, 82,* 113-145.

Gordon, J. K., & Dell, G. S. (2003). Learning to divide the labor: An account of deficits in light and heavy verb production. *Cognitive Science, 27,* 1-40.

Gordon, J. K., & Kurczek, J. C. (2014). The ageing neighborhood: Phonological density in naming. *Language, Cognition, and Neuroscience, 29,* 326-344.

Gregory, M. L., Raymond, W. D., Bell, A., Fosler-Lussier, E., & Jurafsky, D. (1999). The effects of collocational strength and contextual predictability in lexical production. In *Proceedings of the Chicago Linguistic Society, vol. 35* (pp. 151-166).

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38,* 313-338.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11,* 274-279.

Harley, T. A., & Bown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *British Journal of Psychology, 89,* 151-174.

Hartmann, N. (2004). *Processing Grammatical Gender in German: An Eye-Tracking Study on Spoken-Word Recognition.* BSc Thesis. Univ. of Osnabrück, Cognitive Science. Retrieved from http://www.cogsci.uni-osnabrueck.de/~CL/download/Hartmann_Gram Gender.pdf

Hartsuiker, R. J. (2006). Are speech error patterns affected by a monitoring bias? *Language and Cognitive Processes, 21,* 856-891.

Heller, J. R., & Goldrick, M. (in press). Grammatical constraints on phonological encoding in speech production. *Psychonomic Bulletin & Review.*

Heller, J. R., & Goldrick, M. (In preparation a). Grammatically dependent and grammatically independent effects on phonological processing during spontaneous speech.

Heller, J. R., & Goldrick, M. (In preparation b). Meaning-related constraints modulate speech planning and phonetic processing but not phonological neighborhood activation during speech production.

Hunnicutt, S. (1987). Acoustic correlates of redundancy and intelligibility. *Speech, Music and Hearing Quarterly Progress and Status Report, 28,* 7-14.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology, 61,* 23-62.

Jaeger, T. F., Furth, K., & Hilliard, C. (2012). Phonological overlap affects lexical selection during sentence production. *Journal*

Janssen, N., & Caramazza, A. (2009). Grammatical and phonological influences on word order. *Psychological Science, 20,* 1262-1268.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.) *Frequency and the Emergence of Linguistic Structure* (pp. 229–254). Amsterdam: Benjamins.

Kilanski, K. J. (2009). *The effects of token frequency and phonological neighborhood density on native and non-native English speech production.* Ph.D. dissertation, University of Washington. Retrieved from http://www.proquest.com (Publication No. AAT 3370508).

Kirov, C., & Wilson, C. (2013). Bayesian speech production: Evidence from latency and hyperarticulation. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 788-793). Austin, TX: Cognitive Science Society.

Lam, T. Q., & Watson, D. G. (In press). Repetition reduction: Lexical repetition in the absence of referent repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* DOI: 10.1037/a0035780

Levelt, W. J. M., & Meyer, A. S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Psychology, 12,* 433-452.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98,* 122-142.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech, 6,* 172-187.

Lombard, E. (1911). Le signe de l'elevation de la voix. *Annales de Maladies de L'oreille et du Larynx, 37,* 101–119.

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N. (2008). Immediate effects of form-class constraints on spoken word recognition. *Cognition, 108,* 866-873.

Moon, S.-J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America, 96,* 40-55.

Munson, B. (2007). Lexical access, lexical representation, and vowel production. In J. S. Cole & J. I. Hualde (Eds.), *Laboratory Phonology 9* (pp. 201-228). New York: Mouton de Gruyter.

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research, 47,* 1048-1058.

Newman, R. S., & German, D. J. (2005). Life span effects of lexical factors on oral naming. *Language and Speech, 48,* 123-156.

Nooteboom, S. G. (1969). The tongue slips into patterns. In A. G. Sciarone, A. J. van Essen, & A. A. van Raad (Eds.), *Leyden studies in linguistics and phonetics* (pp. 114–132). The Hague: Mouton.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). *Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words.* Bloomington, IN: Psychology Department, Indiana University.

Pechmann, T., & Zerbst, D. (2002). The activation of word class information during speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 233-243.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America, 32,* 693-703.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 539–557.

Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology VII* (pp. 101–139). New York: Mouton de Gruyter.

Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E. & Fosler-Lussier, E. (2007). *Buckeye Corpus of Conversational Speech* (2nd release) [www.buckeyecorpus.osu.edu] Columbus, OH: Department of Psychology, Ohio State University (Distributor).

Potter, M. C., Kroll, J. F., & Harris, C. (1980). Comprehension and memory in rapid sequential reading. In R. S. Nickerson (Ed.), *Attention and Performance VIII* (pp. 395-418). Hillsdale, NJ: Erlbaum.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107,* 460-499.

Rapp, D. N., & Samuel, A. G. (2002). A reason to rhyme: Phonological and semantic influences on lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 564-571.

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language, 57,* 348-379.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object set: The role of surface detail in basic-level object recognition. *Perception, 33,* 217-236.

Sadat, J., Martin, C. D., Costa, A., Alario, F.-X. (2014). Reconciling phonological neighborhood effects in speech production through single trial analysis. *Cognitive Psychology, 68,* 33-58.

Scarborough, R. A. (2010). Lexical and contextual predictability: Confluent effects on the production of vowels. In C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.), *Laboratory Phonology 10* (pp. 557–586). New York: Mouton de Gruyter.

Scarborough, R. (2012). Lexical similarity and speech production: Neighborhoods for nonwords. *Lingua, 122,* 164–176.

Scarborough, R. (2013). Neighborhood-conditioned patterns in phonetic detail: Relating coarticulation and hyperarticulation. *Journal of Phonetics, 41,* 491-508.

Scarborough, R., & Zellou, G. (2013). Clarity in communication: “Clear” speech authenticity and lexical neighborhood density effects in speech production and perception. *Journal of the Acoustical Society of America, 134,* 3793-3807.

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language, 54,* 228-264.

Sekerina, I. A. (2008). Gender priming and mapping of referential expressions in Russian. In G. Zybatow, L. Szucsich, U. Junghans, & R. Meyer (Eds.), *Formal Description of Slavic Languages*. *The Fifth Conference, Leipzig 2003*. (pp. 180-191). Frankfurt am Main, Germany: Peter Lang.

Sekerina, I. A., Brooks, P. J., & Kempe, V. (2006). Gender transparency facilitates noun selection in Russian. In J. Levine & L. Babby, (Eds.), *The Proceedings of the 14th Annual Workshop on Formal Approaches to Slavic Linguistics. The Princeton Meeting*. Michigan Slavic Publications.

Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass, 3,* 621–640.

Smith, M., & Wheeldon, L. (2004). Horizontal information flow in spoken sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 675-686.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory, 6,* 174–215.

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging, 14,* 458-972.

Spinelli, E., Meunier, F., & Seigneuric, A. (2006). Spoken word recognition with gender-marked context. *The Mental Lexicon, 1,* 277-297.

Strand, J., Simenstad, A., Cooperman, A., & Rowe, J. (In press). Grammatical context constrains lexical competition in spoken word recognition. *Memory & Cognition.*

Tily, H., Gahl, S., Arnon, I., Snider, N., Kothari, A., & Bresnan, J. (2009). Syntactic probabilities affect pronunciation variation in spontaneous speech. *Language and Cognition, 1-2,* 147-165.

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication, 49,* 2–7.

Vaden, K. I., Halpin, H. R., & Hickok, G. S. (2009). *Irvine Phonotactic Online Dictionary, Version 2.0.* [Data file]. Available from http://www.iphod.com.

Vigliocco, G., Vinson, D. P., & Siri, S. (2005). Semantic similarity and grammatical class in naming actions. *Cognition, 94,* B91-B100.

Vitevitch, M. S. (1997). The neighborhood characteristics of malapropisms. *Language & Speech, 40,* 211-228.

Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 735-747.

Vitevitch, M. S., & Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory & Cognition, 31,* 491-504.

Wright, R. (2004). Factors of lexical competition in vowel articulation. In J. J. Local, R. Ogden, & R. Temple (Eds.), *Laboratory Phonology VI* (pp. 75-87). Cambridge, UK: Cambridge University Press.

APPENDIX A

Supplemental materials for Chapter 2: Grammatical constraints on phonological encoding in speech production

A1 Target word properties

Properties for all 24 target words are listed in Table A1. 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 Appendix A, section 2). 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 phontactic 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 A2.

A2 Picture name agreement

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

A2.1 Participants

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

A2.2 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.

A2.3 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.

Table A1. Details of stimulus properties for high and low noun phonological neighborhood density (ND) groups.

Target	Vowel	Noun ND	Frequency -weighted Noun ND	Total ND	Frequency -weighted Total ND	Freq-uency (per million)	Contextual Diversity	Image-ability	Name Agree-ment	Onset Voice	Coda Voice	Sum Uniphone Phonotactic Probability	Sum Biphone Phonotactic Probability
Low Noun Density Group													
hat†	æ	34	72.16	37	112.92	64.18	1728	6.7	0.77	0	0	0.0728	0.0015
rat	æ	31	68.84	37	98.24	32.61	941	6.8	0.85	1	0	0.0730	0.0031
doll	a	16	32.78	17	44.99	24.76	648	6.4	0.77	1	1	0.0644	0.0015
pot*	a	32	67.91	35	98.36	22.53	763	6.2	0.85	0	0	0.0661	0.0014
head	ɛ	20	44.49	27	81.18	371.51	6294	6.8	0.77	0	1	0.0622	0.0014
bed	ɛ	20	42.13	29	84.20	187.12	4113	6.6	1.00	1	1	0.0664	0.0014
beak	i	24	52.21	29	71.96	2.10	88	5.2	0.96	1	0	0.0689	0.0012
seal	i	30	66.93	35	92.77	14.75	472	5.5	0.88	0	1	0.0678	0.0019
nose	o	14	27.89	17	49.09	69.75	2125	6.4	1.00	1	1	0.0566	0.0010
boat	o	26	59.76	28	81.67	95.78	1457	6.7	0.81	1	0	0.0691	0.0010
goose	u	13	31.12	14	36.06	13.04	376	6.9	0.92	1	0	0.0554	0.0005
moose	u	17	46.89	18	53.03	5.53	126	6.7	0.96	1	0	0.0588	0.0006
High Noun Density Group													
pan*	æ	27	69.09	30	91.11	12.29	366	6.3	0.88	0	1	0.0730	0.0043
tack	æ	35	84.84	36	100.56	2.12	73	5.3	0.31	0	0	0.0689	0.0025
bomb	a	16	37.09	19	51.33	53.65	934	6.4	1.00	1	1	0.0596	0.0013
lock†	a	31	75.78	32	88.88	56.57	1866	5.6	0.92	1	0	0.0634	0.0015
bell	ɛ	26	62.30	28	87.32	39.33	1035	6.6	1.00	1	1	0.0667	0.0019
pen	ɛ	26	52.97	27	76.85	24.73	816	6.6	1.00	0	1	0.0655	0.0045
heel	i	28	63.84	29	85.33	7.39	281	5.8	0.96	0	1	0.0650	0.0008
peel	i	31	70.10	32	85.41	5.35	189	5.4	0.65	0	1	0.0650	0.0013
coat	o	27	62.51	28	75.86	42.08	1272	6.7	0.62	0	0	0.0662	0.0012
bowl	o	26	64.02	28	76.81	21.45	682	6.3	1.00	1	1	0.0661	0.0012
moon	u	21	51.35	23	66.06	49.96	1167	6.8	0.77	1	1	0.0636	0.0006
boot	u	28	69.45	30	84.57	11.14	374	6.9	0.96	1	0	0.0678	0.0009
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 A2. Group summary statistics for stimulus subsets analyzed

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

A2.4 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 A1), name agreement had a value of 1. Results for items selected for use in the current study can be seen in Table A1.

A3 Sentence construction and norming

A3.1 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 A4, below.

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.*

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).

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 A3).

Table A3. RT condition means for prime conditions

	Phonologically Related	Phonologically Unrelated
Noun Prime	564 ms	557 ms
Verb Prime	562 ms	557 ms

Table A4. Sentence frames

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

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

A3.2 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.

A3.2.1 Participants

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

A3.2.2 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.

A3.2.3 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.

A3.2.3 Analysis and results

A summary of results can be seen in Fig. A1. 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 (β = -0.70, SE = 0.15, t = -4.61, p < 0.0001) and target sentences (β = -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 (β = -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.

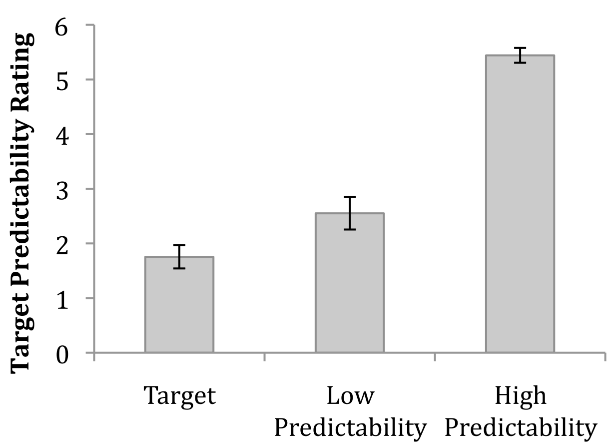

Fig. A1. Mean participant rating of sentence-final word predictability across stimulus group

A4 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.

APPENDIX B

Supplemental materials for Chapter 3: Meaning-related constraints modulate speech planning and phonetic processing but not phonological neighborhood activation during speech production

B1 Target word properties (Experiment 1)

Properties for the 20 target words used in Experiment 1 are listed in Table B1. These are identical to target words analyzed in vowel analyses in Heller and Goldrick (2014). 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. For a description of pictures used and information on picture name norming, see Appendix A, section 2 (Heller & Goldrick, in press). 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 phontactic probabilities were taken

<u>Table B1</u>. Properties for high and low noun phonological neighborhood density (ND) groups used in Experiment 1

Target	Vowel	Noun ND	Frequency -weighted Noun ND	Total ND	Frequency -weighted Total ND	Freq-uency (per million)	Contextual Diversity	Image-ability	Name Agree-ment	Onset Voice	Coda Voice	Sum Uniphone Phonotactic Probability	Sum Biphone Phonotactic Probability
Low Noun Density Group													
rat	æ	31	68.84	37	98.24	32.61	941	6.8	0.85	1	0	0.0730	0.0031
doll	a	16	32.78	17	44.99	24.76	648	6.4	0.77	1	1	0.0644	0.0015
head	ɛ	20	44.49	27	81.18	371.51	6294	6.8	0.77	0	1	0.0622	0.0014
bed	ɛ	20	42.13	29	84.20	187.12	4113	6.6	1.00	1	1	0.0664	0.0014
beak	i	24	52.21	29	71.96	2.10	88	5.2	0.96	1	0	0.0689	0.0012
seal	i	30	66.93	35	92.77	14.75	472	5.5	0.88	0	1	0.0678	0.0019
nose	o	14	27.89	17	49.09	69.75	2125	6.4	1.00	1	1	0.0566	0.0010
boat	o	26	59.76	28	81.67	95.78	1457	6.7	0.81	1	0	0.0691	0.0010
goose	u	13	31.12	14	36.06	13.04	376	6.9	0.92	1	0	0.0554	0.0005
moose	u	17	46.89	18	53.03	5.53	126	6.7	0.96	1	0	0.0588	0.0006
High Noun Density Group													
tack	æ	35	84.84	36	100.56	2.12	73	5.3	0.31	0	0	0.0689	0.0025
bomb	a	16	37.09	19	51.33	53.65	934	6.4	1.00	1	1	0.0596	0.0013
bell	ɛ	26	62.30	28	87.32	39.33	1035	6.6	1.00	1	1	0.0667	0.0019
pen	ɛ	26	52.97	27	76.85	24.73	816	6.6	1.00	0	1	0.0655	0.0045
heel	i	28	63.84	29	85.33	7.39	281	5.8	0.96	0	1	0.0650	0.0008
peel	i	31	70.10	32	85.41	5.35	189	5.4	0.65	0	1	0.0650	0.0013
coat	o	27	62.51	28	75.86	42.08	1272	6.7	0.62	0	0	0.0662	0.0012
bowl	o	26	64.02	28	76.81	21.45	682	6.3	1.00	1	1	0.0661	0.0012
moon	u	21	51.35	23	66.06	49.96	1167	6.8	0.77	1	1	0.0636	0.0006
boot	u	28	69.45	30	84.57	11.14	374	6.9	0.96	1	0	0.0678	0.0009
Low Average		20.00	44.91	23.78	66.10	87.15	1744.33	6.36	0.90	0.78	0.55	0.0633	0.0012
High Average		26.40	61.85	28.00	79.01	25.72	682.30	6.28	0.83	0.50	0.70	0.0654	0.0016
p-value (paired t-test, df = 9)		< 0.01	< 0.01	0.14	0.07	0.15	0.13	0.51	0.45			0.54	0.44

from the Irvine Phonotactic Online Database (IPhOD; Vaden, Halpin, & Hickok, 2009).

B2 Sentence pair norming (Experiment 1)

B2.1 Stimulus construction

Two alternative written context sentences were designed to serve as the first sentence of each stimulus sentence pair. For the Meaning + Grammatical Constraint condition, this sentence was designed to be thematically related to the target word in the second sentence, such that the target word was more predictable in context than its phonological neighbors. For the Grammatical Constraint condition, the sentence was designed to be a plausible preamble to the target sentence, but not to be predictive of the target word in any way, such that the target word's neighbors were also potentially plausible sentence pair completions. Number of words were matched as closely as possible within a pair of alternative context sentences.

These context sentences preceded written target sentences in which the last word in the sentence was the target word. Target sentences were not thematically predictive of the target word, as determined in a previous norming study (see Heller & Goldrick, in press). However, all sentences contained a syntactic cue in the form of a definite determiner preceding the target. Example sentence pairs are shown in 1a-1b, and the full list of stimuli are presented in Table B2.

1a) MEANING + GRAMMATICAL CONSTRAINT

> Annie was shocked that David guessed correctly that she'd had a banana for lunch. He confessed a peep at the *PEEL*.

1b) Grammatical Constraint

Annie was shocked that David guessed correctly what she'd brought to work that day. He confessed a peep at the *PEEL.*

Table B2. Sentence pairs used in Experiment 1

Target	Condition	Sentence
beak	Related	The park was full of modern sculptures of bird parts.
	Unrelated	The park was full of brand new sculptures by local artists.
	Target	He started the bike near the beak.
bed	Related	Reed was pleased to see a lot of promising furniture.
	Unrelated	Reed was pleased to have a lot of exciting ideas.
	Target	He announced a bet on the bed.
bell	Related	Kim guessed that her problems would be solved at the very last minute.
	Unrelated	Kim guessed that the situation was too complicated for her to fix alone.
	Target	She expected the bail at the bell.
boat	Related	James knew he wouldn't have much time when he got off the ferry.
	Unrelated	James knew he wouldn't have much time to eat anything later.
	Target	He hurried a bite on the boat.
bomb	Related	Mary is a master explosives strategist.
	Unrelated	Mary is a master personnel strategist
	Target	She promised a boss near the bomb.
boot	Related	After the avalanche, all Emily could see was a single leg above her.
	Unrelated	After the avalanche, Emily found that she could see almost nothing.
	Target	She attempted a bat at the boot.
bowl	Related	It was a beautiful set of serving platters, but one piece bothered Jim.
	Unrelated	It was an inspired set of surprise plans, but one piece bothered Jim.
	Target	He regretted the bill for the bowl.
coat	Related	Louisa was very protective about the cleanliness of her outerwear.
	Unrelated	Louisa was very protective about some very odd things.
	Target	She refused a comb near her coat.
doll	Related	Courtney felt like all her toys were united against her that afternoon.
	Unrelated	Courtney felt like everything was united against her that afternoon.
	Target	She proposed a deal with the doll.
goose	Related	John had been working on new jokes about farm animals.
	Unrelated	John had been working on new jokes for his routine.
	Target	He tried the goof about the goose.
head	Related	Mike acquired all kinds of new skills while making that turtle Halloween costume.
	Unrelated	Mike acquired all kinds of new skills while completing his latest craft project.
	Target	He learned a hem for the head.
heel	Related	Elizabeth was known for being sassy in fancy shoes.
	Unrelated	Elizabeth was known for being efficient and on-task.
	Target	She ordered the haul with her heel.

moon	Related	Steven often wondered about that night under the stars.
	Unrelated	Steven often wondered about the way the movie ended.
	Target	He remembered a moan at the moon.
moose	Related	While hiking in Alaska, Deb kept spotting forest animals.
	Unrelated	While running the marathon, Deb kept passing the strangest things.
	Target	She continued her move by the moose.
nose	Related	Before he regained the use of his limbs, Chris relied on his face.
	Unrelated	Before he recovered from the accident, Chris acted very strangely.
	Target	He indicated the note with his nose.
peel	Related	Annie was shocked that David guessed correctly that she'd had a banana for lunch.
	Unrelated	Annie was shocked that David guessed correctly what she'd brought to work that day.
	Target	He confessed a peep at the peel.
pen	Related	Henrietta was a chicken with a thing for writing utensils.
	Unrelated	Henrietta was a chicken who wanted a bite of everything.
	Target	She wanted a peck at the pen.
rat	Related	Casey hated the city's new statues of vermin, but plans had to be made.
	Unrelated	Casey hated the city's new public art project, but plans had to be made.
	Target	She approved the route by the rat.
seal	Related	On the weekends, Katie loved visiting her favorite friendly water animal.
	Unrelated	On the weekends, Katie loved visiting her favorite secret places.
	Target	She liked the sail by the seal.
tack	Related	Paul passed by a stranger admiring how things were stuck to the bulletin board.
	Unrelated	Paul passed by a stranger who seemed to be methodically touching everything in sight.
	Target	He noticed a tap on the tack.

B2.2 Sentence pair norming

The goal of sentence pair norming was twofold. First, we wanted to ensure that targets were more predictable in the Meaning + Grammatical Constraint condition than in the Grammatical Constraint condition. Second, we wanted to ensure that targets were more predictable than their phonological neighbors in the Meaning + Grammatical Constraint condition, but not any more predictable than their neighbors in the Grammatical Constraint condition. In order to do this, the plausibility of targets in the sentence pairs listed in Table B2 were compared to the plausibility of a high frequency neighbor of the target in the same contexts. Participants were asked to rate, on a seven-point Likert scale, the likelihood of each target (or neighbor) as a sentence completion given the information available in a particular sentence pair.

B2.2.1 Participants

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

B2.2.2 Stimuli

Stimuli consisted of the sentence pairs described above in section B2.1, which featured target words sentence-finally, as well as an identical set of sentence pairs for which each target's highest frequency neighbor (Foil) that matched in grammatical category (and was not otherwise already represented in the experiment) replaced the target as the sentence completion. Neighbors were selected from the IPhOD database (Vaden et al., 2009), and neighbor frequencies were assessed using the SUBTLEX-US corpus (Brysbaert et al., 2012). Foils are listed in Table B3.

Table B3. Target foils (high frequency same-category neighbors)

Target	Foil	Target	Foil
beak	back	head	hair
bed	bird	heel	deal
bell	ball	moon	man
boat	note	moose	miss
bomb	mom	nose	news
boot	bit	peel	piece
bowl	hole	pen	men
coat	code	rat	rate
doll	dock	seal	seat
goose	gas	tack	talk

B3.2.3 Procedure

Sentences pairs 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 in the second (target) sentence, given the information in the sentences preceding that word. Before they completed the task, they were familiarized with high and low predictability examples. Participants rated each sentence pair once.

B3.2.3 Analysis and results

A summary of results can be seen in Fig. B1. Differences between stimulus groups were assessed using a linear mixed effects regression analysis of participants' predictability ratings, with contrast-coded context condition (Meaning + Grammatical Constraint, Grammatical Constraint), target type (Target, Foil), and their interaction as fixed effects. Random effects included random intercepts for participant and item, as well as random slopes for context condition, target type, and their interaction, for both participant and item. Significance was assessed by model comparison. Predictability was rated significantly higher in the Meaning + Grammatical Constraint condition (β = 0.79, SE = 0.15, $\chi^2(1)$ = 18.30, p < 0.0001), and also significantly higher for targets than for their foils (β = 1.13, SE = 0.20, , $\chi^2(1)$ = 21.65, p < 0.0001). Crucially, the interaction of these factors was also significant, such that predictability ratings were higher in the Meaning + Grammatical Constraint for target items (β = 2.07, SE = 0.30, $\chi^2(1)$ = 25.52, p < 0.0001).

Follow-up regressions collapsing these two factors into a single contrast-coded four-level factor (TargetRelated, TargetUnrelated, FoilRelated, FoilUnrelated) were performed to assess specific comparisons between conditions. Random intercepts were included for item and participant, and a random slope for the single four-level factor was included for both item and participant. Significance of comparisons within this sole four-level fixed effect were assessed by assuming that the distribution of *t*-statistics follows that of *z*-statistics, rather than by model comparison. Compared to the TargetRelated condition, predictability in the TargetUnrelated condition was rated significantly lower (β = -1.83, SE = 0.28, t = -6.45, p < 0.0001), as was FoilRelated (β = -2.17, SE = 0.33, t = -6.59, p < 0.0001). In a releveled analysis, FoilUnrelated predictability scores were not significantly different from those in the TargetUnrelated condition (β = 0.10, SE = 0.13, t = 0.77, p > 0.1). Thus, Meaning + Grammatical Constraint contexts made targets more predictable than when they appeared in Grammatical Constraint contexts, Meaning + Grammatical Constraint contexts made targets more predictable than foils, and in Grammatical Constraint contexts targets and foils did differ reliably in predictability.

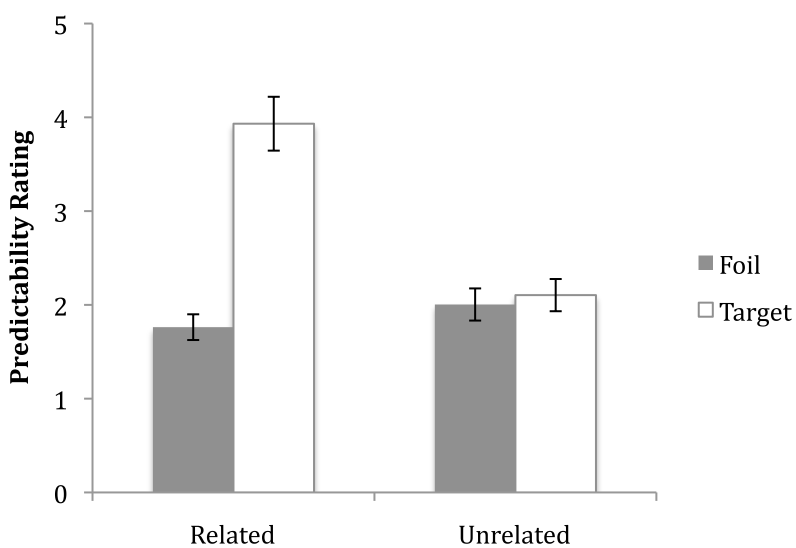

Fig. B1. Mean participant rating of sentence-final word predictability across stimulus conditions. Error bars represent by-participant standard error

B3 Target word properties (Experiment 2)

In order to test whether effects seen could generalize to a larger set of items, target words used in Experiment 2 consisted of the 20 target words used in Experiment 1, with the addition of 14 additional items, for a total of 34 target words. Items were matched on the same factors as Experiment 1 (see section S1 for details). Properties of the total set of items are presented in Table B4 (low noun density group), Table B5 (high noun density group), and Table B6 (group summaries and comparison).

B4 Sentence norming (Experiment 2)

In order to boost meaning-related predictability of the target word in the Meaning + Grammatical Constraint condition, the goal was to increase the strength of meaning-related constraints. This was operationalized in two ways. First, target words were presented in single sentences for which they were highly probable or less probable completions, rather than being presented at the end of sentence pairs, where the source of their predictability was in the previous sentence. Second, using a test of cloze probability (the probability that a word is supplied to complete a sentence), we ensured that targets were the most frequently supplied completion (as opposed to simply being more probably than a high frequency neighbor, as in the previous experiment).

Table B4. Details of low noun phonological neighborhood density (ND) group (Experiment 2)

Target	Vowel	Noun ND	Frequency -weighted Noun ND	Total ND	Frequency -weighted Total ND	Freq-uency (per million)	Contextual Diversity	Image-ability	Name Agree-ment	Onset Voice	Coda Voice	Sum Uniphone Phonotactic Probability	Sum Biphone Phonotactic Probability
fan	æ	21	49.65	25	74.18	35.14	1097	6.0	1.00	0	1	0.0696	0.0040
rat	æ	31	68.84	37	98.24	32.61	941	6.8	0.85	1	0	0.0730	0.0031
hat	æ	34	72.16	37	112.92	64.18	1728	6.7	0.77	0	0	0.0728	0.0015
bomb	a	16	37.09	19	51.33	53.65	934	6.4	1.00	1	1	0.0596	0.0013
car	a	24	57.75	26	76.77	483.06	5154	6.9	0.85	0	1	0.0545	0.0042
knot	a	23	50.48	27	75.67	3.69	151	5.9	0.88	1	0	0.0598	0.0011
net	ɛ	21	48.08	26	78.33	15.55	541	5.9	0.88	1	0	0.0591	0.0015
head	ɛ	20	44.49	27	81.18	371.51	6294	6.8	0.77	0	1	0.0622	0.0014
bed	ɛ	20	42.13	29	84.20	187.12	4113	6.6	1.00	1	1	0.0664	0.0014
sheep	i	18	36.00	21	54.43	13.43	392	6.8	0.59	0	0	0.0473	0.0007
beak	i	24	52.21	29	71.96	2.10	88	5.2	0.96	1	0	0.0689	0.0012
seal	i	30	66.93	35	92.77	14.75	472	5.5	0.88	0	1	0.0678	0.0019
nose	o	14	27.89	17	49.09	69.75	2125	6.4	1.00	1	1	0.0566	0.0010
boat	o	26	59.76	28	81.67	95.78	1457	6.7	0.81	1	0	0.0691	0.0010
bone	o	24	49.72	30	82.39	26.06	855	6.4	1.00	1	0	0.0692	0.0010
goose	u	13	31.12	14	36.06	13.04	376	6.9	0.92	1	0	0.0554	0.0005
moose	u	17	46.89	18	53.03	5.53	126	6.7	0.96	1	0	0.0588	0.0006

Table B5. Details of high noun phonological neighborhood density (ND) group (Experiment 2)

Target	Vowel	Noun ND	Frequency -weighted Noun ND	Total ND	Frequency -weighted Total ND	Freq-uency (per million)	Contextual Diversity	Image-ability	Name Agree-ment	Onset Voice	Coda Voice	Sum Uniphone Phonotactic Probability	Sum Biphone Phonotactic Probability
pan	æ	27	69.09	30	91.11	12.29	366	6.3	0.88	0	1	0.0730	0.0043
cat	æ	31	72.48	33	95.64	66.33	1386	6.5	1.00	0	0	0.0742	0.0021
tack	æ	35	84.84	36	100.56	2.12	73	5.3	0.31	0	0	0.0689	0.0025
doll	a	16	32.78	17	44.99	24.76	648	6.4	0.77	1	1	0.0644	0.0015
sock	a	27	59.10	28	72.85	8.98	302	6.7	0.89	0	0	0.0657	0.0011
lock	a	31	75.78	32	88.88	56.57	1866	5.6	0.92	1	0	0.0634	0.0015
neck	ɛ	17	35.02	17	45.46	59.51	1961	6.5	0.96	1	0	0.0559	0.0026
pen	ɛ	26	52.97	27	76.85	24.73	816	6.6	1.00	0	1	0.0655	0.0045
bell	ɛ	26	62.30	28	87.32	39.33	1035	6.6	1.00	1	1	0.0667	0.0019
heel	i	28	63.84	29	85.33	7.39	281	5.8	0.96	0	1	0.0650	0.0008
wheel	i	27	65.42	30	86.18	27.06	851	6.5	0.94	0	1	0.0605	0.0010
peel	i	31	70.10	32	85.41	5.35	189	5.4	0.65	0	1	0.0650	0.0013
comb	o	24	48.85	24	62.90	6.06	221	6.4	0.60	0	1	0.0556	0.0010
coat	o	27	62.51	28	75.86	42.08	1272	6.7	0.62	0	0	0.0662	0.0012
bowl	o	26	64.02	28	76.81	21.45	682	6.3	1.00	1	1	0.0661	0.0012
moon	u	21	51.35	23	66.06	49.96	1167	6.8	0.77	1	1	0.0636	0.0006
boot	u	28	69.45	30	84.57	11.14	374	6.9	0.96	1	0	0.0678	0.0009

Table B6. Summary of high and low noun phonological neighborhood density (ND) groups (Experiment 2)

	Noun ND	Frequency -weighted Noun ND	Total ND	Frequency -weighted Total ND	Freq-uency (per million)	Contextual Diversity	Image-ability	Name Agree-ment	Onset Voice	Coda Voice	Sum Uniphone Phonotactic Probability	Sum Biphone Phonotactic Probability
Low Average	22.12	44.48	26.18	73.78	87.47	1579.06	6.39	0.89	0.65	0.47	0.0629	0.0016
High Average	26.35	61.17	27.76	78.04	27.36	793.53	6.31	0.84	0.47	0.59	0.0651	0.0018
p-value (paired t-test, df = 16)	< 0.01	< 0.01	0.24	0.33	0.10	0.13	0.59	0.33			0.18	0.63

B4.1 Stimulus construction

Two sentences were created for each target word, for which a target word was either a highly predictable completion (Meaning + Grammatical Constraint) or an unpredictable completion (Grammatical Constraint). The target word always appeared sentence-finally. For the 20 target words that were used in previous experiments, the non-predictive target sentence frame used in those experiments (see, e.g., Table B2) served as its Grammatical Constraint frame. Meaning + Grammatical Constraint frames and remaining Grammatical Constraint frames were either adapted by the first author from previous studies of word predictability in sentence contexts (Block & Baldwin, 2010; Bloom & Fischler, 1980; Fallon, Trehub, & Schneider, 2002; Griffin, 2002; Griffin & Bock, 1998; Hamberger, Friedman, & Rosen, 1996; Kalikow, Stevens, & Elliott, 1977) or created by the first author. In all sentences, the target was preceded by a definite determiner. Number of words was matched as closely as possible across conditions. The full list of sentences can be seen below in Table B7.

B4.2 Sentence norming

The goal of sentence norming was to ensure that targets were the most predictable sentence continuations for Meaning + Grammatical Constraint contexts, and that they were unpredictable sentence completions in Grammatical Constraint contexts. These properties were assessed using a cloze task. Targets were considered the most predictable sentence completion for a given sentence frame if participants completed that frame with the target word more than 50% of the time.

B4.2.1 Participants

Thirty-nine native speakers of American English (aged 18-35) participated. Nineteen were Northwestern University undergraduates who participated in exchange for partial course credit. Twenty were participants recruited through Amazon Mechanical Turk who were paid $1 for participation.

Table B7. Related and unrelated sentences

Target	Condition	Sentence	Cloze Probability
beak	Related	The chicken pecked corn with its beak.	1.00
	Unrelated	He started the bike near the beak.	0.00
bed	Related	She put fresh sheets on the bed.	0.97
	Unrelated	He announced a bet on the bed.	0.00
bell	Related	For service, please ring the bell.	0.82
	Unrelated	She expected the bail at the bell.	0.00
boat	Related	He began to row the boat.	0.92
	Unrelated	He hurried a bite on the boat.	0.00
bomb	Related	The bridge exploded when he detonated the bomb.	0.74
	Unrelated	She promised a boss near the bomb.	0.00
bone	Related	The dog buried the bone.	0.97
	Unrelated	She believed that he often thought about the bone.	0.00
boot	Related	He was a cowboy from his hat to his boot.	0.87
	Unrelated	She attempted a bat at the boot.	0.00
bowl	Related	She ladled soup into her bowl.	0.95
	Unrelated	He regretted the bill for the bowl.	0.00
car	Related	George taught his son to drive the car.	0.85
	Unrelated	His mouth dropped open when he heard about the car.	0.00
cat	Related	Give a saucer of milk to the cat.	0.92
	Unrelated	He thought he heard them whispering about the cat.	0.00
coat	Related	Let me take your hat and coat.	0.90
	Unrelated	She refused a comb near her coat.	0.00
comb	Related	He parted his hair with the comb.	0.95
	Unrelated	The tall man was shown the drawing of the comb.	0.00
doll	Related	The little girl cradled her doll.	0.72
	Unrelated	She proposed a deal with the doll.	0.00
fan	Related	The air was too still, so I turned on the fan.	0.87
	Unrelated	The next topic of discussion will be the fan.	0.00
goose	Related	The wife of the gander is the goose.	0.54
	Unrelated	He tried the goof about the goose.	0.00

hat	Related	The magician pulled the rabbit out of the hat.	0.97
	Unrelated	She suggested a hit through the hat.	0.00
head	Related	The pinecone fell and hit Terry in the head.	0.59
	Unrelated	He learned a hem for the head.	0.00
heel	Related	Measure your foot from the toe to the heel.	0.87
	Unrelated	She ordered the haul with her heel.	0.00
knot	Related	The sailor grabbed the rope and tied the knot.	0.77
	Unrelated	I believe you will soon see a familiar image of the knot.	0.00
lock	Related	She secured her bicycle to the rack with her lock.	0.72
	Unrelated	He chose a look at the lock.	0.00
moon	Related	The astronauts landed on the moon.	1.00
	Unrelated	He remembered a moan at the moon.	0.08
moose	Related	On Rocky and Bullwinkle, Rocky was the squirrel and Bullwinkle was the moose.	0.79
	Unrelated	She continued her move by the moose.	0.00
neck	Related	She loosened the tie around her neck.	1.00
	Unrelated	You must have noticed the mark on his neck.	0.08
net	Related	Nick catches butterflies with the net.	0.92
	Unrelated	Dad liked to read aloud to us about the net.	0.00
nose	Related	He got a tissue and blew his nose.	0.97
	Unrelated	He indicated the note with his nose.	0.00
pan	Related	She preheated the oven and greased the pan.	0.85
	Unrelated	The computer screen showed a bright display of the pan.	0.00
peel	Related	When you eat a banana, don't eat the peel.	0.90
	Unrelated	He confessed a peep at the peel.	0.00
pen	Related	A writing implement filled with ink is called the pen.	0.97
	Unrelated	She wanted a peck at the pen.	0.00
rat	Related	The long-tailed rodent living in the garbage is called the rat.	0.85
	Unrelated	She approved the route by the rat.	0.00
seal	Related	The animal trained to balance a ball on its nose is the seal.	0.74
	Unrelated	She liked the sail by the seal.	0.00
sheep	Related	The shepherd herded the sheep.	0.95
	Unrelated	Please show him your new presentation about the sheep.	0.00
sock	Related	We put the shoe on after the sock.	0.79
	Unrelated	Her mom carefully inspected the image of the sock.	0.00
tack	Related	Because it was sharp, the teacher yelped when she sat on the tack.	0.54
	Unrelated	He noticed a tap on the tack.	0.00

B4.2.2 Stimuli

Stimuli consisted of the 68 Meaning + Grammatical Constraint and Grammatical Constraint sentence frames described above in section S4.1 and listed in Table B7, each of which was missing its final (target) word. Thirty-four filler sentence frames, varying in syntactic structure and the semantic constraint on their missing final words, were interspersed in a pseudorandomized order.

B4.2.3 Procedure

Sentence frames (missing their final word) appeared one at a time on a computer screen. Participants were asked to type the best single-word sentence completion into a text box below the sentence frame. Before they completed the task, they were familiarized with high- and low-constraint examples. Participants completed sentence once.

B4.2.4 Analysis and results

Cloze probability for each item was calculated as the fraction of words provided as sentence completions by participants that matched the intended target word. 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 completed a sentence with the intended target name (as was the case for, e.g., the Meaning + Grammatical Constraint frame for *beak*, Table B7), the cloze probability for that target in that sentence frame had a value of 1. Several typos and variants were corrected in this process (e.g., *heal* was corrected to *heel*, *kneck* was corrected to *neck*).

A summary of results can be seen in Table B7. In Meaning + Grammatical Constraint frames, all targets had a cloze probability of greater than 0.5 (M = 0.84, SD = 0.13). In

Grammatical Constraint frames, all targets had cloze probabilities of less than 0.1 (M = 0.00, SD = 0.02). Thus, targets were the most predictable sentence completions for Meaning + Grammatical Constraint contexts, but were unpredictable in Grammatical Constraint contexts.