Using psychological realism to advance phonological theory

Matthew Goldrick

Department of Linguistics
Northwestern University

Address for Correspondence:
Matthew Goldrick
Department of Linguistics
Northwestern University
2016 Sheridan Rd.
Evanston, IL 60208 USA
Phone: (847) 467-7020
Fax: (847) 491-3770
Email: matt-goldrick@northwestern.edu

Acknowledgements. Thanks to Ann Bradlow and especially the editors for helpful discussions. Many thanks to Lisa Davidson and Susannah Levi for helpful discussions and comments on the manuscript. Preparation of this manuscript was supported in part by National Institutes of Health grant DC 0079772.

0 Overview

Following its introduction by Sapir (1933), the term "psychological reality" has provoked intense reactions from within linguistics as well its neighboring disciplines. Discussions have been particularly heated since the rise of generative grammar, whose proponents made quite strong claims regarding the relationship of theoretical concepts from linguistics to the internal cognitive mechanisms underlying the acquisition and processing of sound patterns. For example, *The Sound Pattern of English* is asserted to be "a hypothesis concerning the actual internalized grammar of the speaker-hearer" where grammar refers to "a system which is used in the production and interpretation of utterances (Chomsky and Halle, 1968: 4)." Although this perspective is by no means universally adopted by phonologists, its dominance in linguistics since the mid-20th century reflects a major conceptual shift from previous perspectives on the study of language. As noted by Anderson (1985:6; emphasis in original):

> "Traditionally, linguists have assumed that their concern was the study of *languages*, taken as (potentially unlimited) sets of possible sentences (or utterances, etc.) forming unitary and coherent systems. Gradually, however, the emphasis in research has shifted…to the properties of *grammars*, in the sense of systems…which specify the properties of the (well-formed) sentences in such a system."

This shift, coupled with the claim that linguistic systems define a capacity (or competence) possessed by *individual* speakers of a language, has focused attention on the correspondence between constructs from linguistic theories and the cognitive systems of individual speaker-hearers. *Psychological realism* views such correspondences as cornerstones of linguistic research—both in terms of empirical practice and theory development. This chapter considers the content and import of this approach in competent, adult individuals (for discussion of language acquisition, see Demuth, this volume). Three core issues are considered:

- What is psychological realism? Psychological realism adopts a cognitive psychological perspective to explain human linguistic behavior. This offers a functional-level account of how different components of the human cognitive system interact to yield particular behaviors.
- Why is psychological realism critical for linguistic research? Human behavior always reflects the interaction of multiple cognitive components. Without making explicit (and empirically justified) assumptions about the nature of these interactions, we cannot correctly draw inferences about the structure of the cognitive system. The perils of failing to specify these assumptions will be illustrated using well-formedness judgments.
- How can psychological realism help resolve theoretical issues in linguistics? If we take seriously the need to articulate the functional architecture underlying specific tasks, we can better understand the import of behavioral data. This can help resolve outstanding theoretical questions such as the nature of the relationship between lexical and grammatical knowledge.

1 What is psychological realism?

1.1 The structure of psychological theories

Psychological realism adopts the theoretical perspective of cognitive psychology to understand language-related behavior. (n.b.: This is by no means the only perspective on human psychological capacities; see, e.g., van Gelder, 1998.) It explains language behavior as the coordinated interaction, in real time, of a set of more primitive capacities or *functions* that map between inputs and outputs.

1.1.1 Functional explanation

Cognitive psychological theories aim to specify the psychological capacities of individuals. Capacities are the regularities that govern the behavior of the cognitive system. These regularities are *lawlike* in that given certain precipitating conditions (environmental or internal), the system will exhibit certain (behavioral or internal) manifestations (Cummins, 1983). The term "function" is used to evoke the idea of a precisely specified relation or mapping between starting configurations of the system (i.e., precipitating conditions) and ending configurations (manifestations). When describing the cognitive system at this level of description, there is no specification of *how* this mapping is accomplished. Our theory simply specifies that given certain inputs, a certain distribution of outputs will be produced. (Note this mapping could be deterministic or stochastic.) This corresponds to Marr's (1982) *computational* level of analysis of cognitive systems. The discussion here follows Smolensky (2006) in referring to this as the *functional* level of description.

To illustrate this level of analysis, many psycholinguistic theories of speech perception assume two broad functional stages are involved in the perception of single spoken words (McClelland, Mirman & Holt, 2006). The first stage, *pre-lexical processing*, takes as input a relatively fine-grained representation of acoustic information (e.g., acoustic features) and produces as output a pre-lexical representation elaborating the linguistic structure of the acoustic input (e.g., by specifying segmental and prosodic structure). The second stage, *lexical processing*, uses this pre-lexical representation to retrieve a lexical representation of the utterance (e.g., a unitary whole-word representation <DOG>; this is used to access semantic and syntactic information).

Although this level of description does not specify *how* a mapping is computed, claims stated at this level are contentful statements about the psychological organization of speakers. First, as anyone who has attempted to construct a generative grammatical analysis can tell you, it is no trivial matter to precisely specify a function that maps a large set of inputs to the correct set of outputs. Second, functional analyses represent a critical initial step in the pursuit of reductionist accounts of behavior. Functional theories take a complex capacity such as "the ability to perceive the meaning of single spoken words" and decompose it into simpler capacities—for example, "mapping acoustic signals onto phonemes," "retrieving words matching the perceived phonemes," and "retrieving the meaning of words." (The hope is that this reductionist procedure will terminate in simple capacities which can be realized as neural computations; see below). That such claims are taken to be contentful is clear from many psycholinguistic studies aiming to distinguish theories with contrasting decompositions of complex capacities. For example, with respect to the capacity of perceiving single spoken words, other psycholinguistic theories of perception have proposed that there is no explicit pre-lexical stage intervening between acoustic signal processing and meaning retrieval (e.g., Gaskell and Marslen-Wilson, 1997).

1.2 The role of capacities in real-time linguistic behavior

Cognitive psychological theories aim to account for linguistic behavior. An essential component of such theories is therefore specifying how functions are utilized in the performance of specific behaviors or tasks. For example, what functions are utilized in an auditory lexical decision task (where a participant must decide if a string of sounds is a lexical item or not)?

Within the general framework outlined above, it is generally assumed that performance in this task is related in part to the outcome of lexical processes; participant responses reflect at least in part whether a word representation is or is not successfully retrieved for the input. Within the architecture above, this means that pre-lexical processes are also engaged; sensory input cannot influence lexical processing without the mediation of pre-lexical processes. Subsequent to lexical processing, there must also be decision processes that allow the hearer to make a word/nonword response (see Ratcliff, Gomez, and McKoon, 2004, for a recent review of such models in the context of lexical decision tasks using printed words).

It is critical to note that within psychological theories functions, like human behavior, exist in real time; they have temporal extent. For example, Palmer and Kimchi (1986: 40) define psychological capacities as "informational events" consisting of "the *input information* (what it starts with), the *operation* performed on the input (what gets done to the input) and the *output information* (what it ends up with) [emphasis original]." Rather than speaking broadly of a function as an a-temporal specification of a relation or function between input and output, this approach makes the stronger assumption that the capacity literally starts with the input at some time and after some distinct period actively *produces* the output. Likewise, the interaction of these capacities is inherently temporal. According to Palmer and Kimchi decomposition of a complex capacity is specification of a set of informational events plus "the temporal ordering relations among them that specify how the information 'flows' through the system of components (p. 47)." That is, the primary task of structuring the interaction of simpler capacities is specifying their temporal relationships.

1.2 Psychological vs. algorithmic or neural accounts

Psychological theories, stated at the functional level of description, do not offer complete accounts of the cognitive system. Ultimately a complete theory must specify not only what functions are but (1) how they are computed and (2) how they are realized physically by the nervous system (Marr, 1982). Accounts that address these two issues are sketched below.

1.2.1 The algorithmic level

The algorithmic level of description is an abstract, computational characterization of the process which satisfies the function specified at the higher level of description. That is, if the cognitive system is placed in the appropriate initial configuration, the algorithm will place the system in the desired ending configuration.

For example, the lexical function described above has been instantiated within spreading activation networks (e.g., McClelland & Elman, 1986; Norris, McQueen, & Cutler; 2000). In these networks, one set of processing units instantiates pre-lexical speech sound representations (e.g., there is a unit corresponding to initial /d/, another for initial /t/, etc.). Another set of units instantiates the lexical representations (e.g., there is a unit for DIG, another for DOT, etc.). The function specified above is realized by connection weights that allow activation to flow between these two levels of representation. When the word-initial input /d/ is provided to the network by imposing a certain pattern of activation on the pre-lexical units (e.g., activating /d/ but not /t/), activation will automatically flow along these connections to the appropriate output units (e.g., activating DIP but not TIP). This process is entirely mechanical; given an input, the network will automatically produce (via spreading activation) the output that satisfies the function specified above.

1.2.2. Neural level accounts

Of course, a complete account of the cognitive system cannot stop at the algorithmic level. Human cognitive systems are ultimately realized by neurobiological structures and processes. Therefore, at the lowest level of description—the neural level—the algorithms specified at the preceding level are implemented in terms of neural systems. For example, the process of accessing lexical representations from acoustic input has been argued to be instantiated by brain structures in the vicinity of the temporal-parietal-occipital junction (see Hickok and Poeppel, 2000, for a recent review). A neural specification of the spreading activation networks above would have to detail how the algorithm specified above (i.e., abstract pre-lexical as well as word-sized processing units; activation flow among these units) is instantiated in these neuronal assemblies.

1.2.3 Psychological accounts are not algorithmic or neural accounts

Cognitive psychological explanations are, in general, stated in terms of capacities; that is, they are *functional* level explanations. They do not typically address the algorithmic (much less neural) realization of capacities characterized at the functional level. For example, following the general reductionist strategy of functional accounts, cognitive psychological theories account for complex behavior in terms of the interaction of (relatively) simple capacities (Cummins, 1983; Palmer and Kimchi, 1986). These simple component capacities are assumed to be physically embodied (i.e., algorithmically and neurally realized) but the details of how this occurs are typically not spelled out. "In reality, most IP [information-processing or cognitive psychological] theorists give, at best, a rather vague, verbal description of the input-output characteristics of the components…unfortunately, simulations [algorithmic implementations] are seldom actually done… (Palmer and Kimchi, 1986:53-54)."

That is not to say that psychological (or linguistic) theorizing categorically avoids other levels of description. In particular, many theories are at least partially specified at both the functional and algorithmic levels. For example, as discussed above the processing of monosyllabic monomorphemic forms has been computationally implemented by a spreading activation network (note, however, the limited range of inputs this specific algorithm can process). However, such work is the exception rather than the rule. In most psycholinguistic theories, many processing components are wholly unanalyzed algorithmically (as noted by Palmer and Kimchi above). The situation is far worse with respect to the neural level of description. As far as I am aware, no theories of language-related capacities have attempted to specify the physically instantiated neurobiological processes that realize cognitive functions (although many have investigated the neurobiological structures associated with linguistic capacities). For example, although connectionist research has attempted to specify algorithms that are broadly compatible with neuronal computational principles, it is still extremely abstract relative to actual neurobiological mechanisms (Smolensky, 2006).

Psychological realism is therefore like most research in linguistics; it adopts a functional level approach to understanding human behavior. A critical issue for any functional theory is *realizability*: how is the functional level description instantiated algorithmically and ultimately neurally? These issues are critical, as they address the *physical* reality of theoretical constructs. If no algorithm can be specified that instantiates a hypothesized function, or if there is no way to realize that algorithm neurobiologically, the functional level description becomes significantly less plausible. But (contra authors such as Linell, 1979) these issues are distinct from specifying a psychological (functional level) account of linguistic behavior.

1.3 Linking linguistic and psychological theories

As noted above, linguistic theories are also, in general, functional level theories. They decompose complex linguistic knowledge into a set of simpler functions (e.g., syntactic vs. phonological components of the grammar). Although research in computational phonology aims to specify algorithms that compute grammatical functions (see Coleman, this volume, for further discussion), the typical linguistic analysis does not consider how it is computationally (much less neurally) implemented.

Psychological theory can enrich such theories by providing a framework for thinking about how these linguistic functions are deployed during behavioral. Like linguistic theories, cognitive psychological theories explain language behavior as the coordinated interaction of a set of more primitive functions that map between inputs and outputs. Unlike many linguistic theories, psychological accounts are situated within specific behavioral tasks and in real time. This enables psychological theories to make predictions (that can be confirmed or refuted) for behavioral experiments. Linking a linguistic theory with a psychological theory allows the linguistic theorist to draw upon this rich body of evidence to inform their theory.

Making such connections is facilitated by the use of functional level descriptions in each tradition. However, difficulties can arise due to contrasting assumptions regarding the specificity of linguistic knowledge. Most linguistic theories aim to characterize capacities common to all linguistic behaviors; in contrast, many psychological theories aim to characterize the capacities involved in particular sets of behaviors[1]. For example, the psychological theory discussed above

[1] In practice, of course, theories often lie in between these two extremes. For example, some recent work in the framework of Optimality Theory has pursued the hypothesis that distinct grammars (with distinct rankings and/or constraints) are used in perception and production (e.g., Boersma, 1999; Kenstowicz, 2001). Similarly, psycholinguistic theories have postulated that a single system is used in both modalities (see Martin and Saffran, 2002, for discussion).

concerns the relationship of various levels of sound structure in speech perception; it makes no claims regarding speech production. In contrast, a typical linguistic theory would attempt to characterize the general relationship between levels of sound structure representation—a relationship that subserves perception, production, acquisition, well-formedness judgments, etc. For example, Jakobson (1941: 92) claims "the *same laws* of solidarity [emphasis mine]" underlie child language, aphasia, and typological sound structure patterns. Chomsky and Halle (1968) assume the (singular) grammar is "*a* system used in the production *and* interpretation of utterances (p. 4; emphasis mine)."

A critical issue in psychological realism is therefore establishing how components of linguistic theories link up to components of psychological theories. Without such links, it is impossible for linguistic theories to use psychological theories to help draw inferences from behavior. The nature of such links cannot be established a priori. Some theories have assumed relatively direct connections between components of linguistic theories and psychological mechanisms (e.g., Goldrick & Daland, 2009). However, linguistic theories making differing assumptions regarding the specificity of knowledge may necessitate more complex relationships with psychological theories. For example, the phonological component of the grammar may be distributed over many distinct psychological capacities (specific to memory, language production, perception, etc.).

In spite of such complexities, the establishment of these links is imperative for linguistic theorists that wish to make use of behavioral data. If a linguistic theory is not situated within specific tasks that occur in real time it cannot be informed by data from on-line behavioral tasks. Psychological theories provide an appropriate set of linking hypotheses licensing such inferences.

2 Why is psychological realism critical for linguistic research?

As discussed above, psychological theories account for behavior in any given task through the complex interaction of *many* simple capacities. This point has been noted by many authors; for example, Chomsky (1980: 188) writes "the system of language is only one of a number of cognitive systems that interact in the most intimate way in the actual use of language." Such interactions clearly complicate the interpretation of behavioral data. When assessing behavior, it is not sufficient that we richly articulate a theory of the cognitive component of interest (e.g., the "phonological grammar"). We must also specify how this component interacts with other relevant cognitive process to produce the behavior(s) of interest. As discussed by Caramazza (1986: 47):

> "observations do not carry on their sleeves signs indicating whether or not they constitute relevant evidence in some domain of investigation. An especially important point is that a specific set of observations…will assume evidential status with respect to some model only if we are able to provide adequate arguments…to explicitly link the type of observations in question to the component or components of processing being investigated."

Using behavioral data to inform theories therefore *requires* "a sufficiently detailed model of the cognitive systems of interest to guide the search for richly articulated patterns of performance (Caramazza, 1986: 66)." Until this has been specified to at least some level of detail, we cannot establish (in Caramazza's term) the "evidential status" of behavioral data. These concerns are by no means unique to cognitive psychologists. For example, Fodor (1981: 200) notes that "[a]ny science is under the obligation to explain why *what it takes to be* data relevant to the confirmation of its theories *are* data relevant to the confirmation of its theories [emphasis original].”

In this section, we examine a critical domain of behavioral data—wordlikeness judgments—where the inference from data to the structure of the cognitive system has been impaired by the failure to consider how multiple cognitive components interact to produce behavior. The particular domain we focus on here is phonotactic knowledge—our knowledge that certain combinations of phonological structures are dispreferred relative to others. First, the role of such knowledge in generative theories is briefly reviewed. We then critically review the data from wordlikeness judgments.

2.1 Generative models of phonotactic knowledge

2.1.1 Phonotactic knowledge

Evidence from a wide array of linguistic behaviors suggests that our cognitive systems are structured in such a way as to disfavor certain combinations of phonological structures relative to others (n.b. these preferences are to a certain degree language specific, but general patterns are found across languages). For example, although English words contain both /p/ and /l/, a phonological string with an initial cluster /lp/ will be disfavored over a string with an initial /pl/. The dispreference for certain structures may manifest itself in a number of different ways, including: native speaker judgments of acceptability (a string with initial /lp/ is judged to be a poor English word); statistical under-representation (or absence) in corpora (there are no words in English with initial /lp/); difficulties in memory, perception, and production (for English speakers, it is difficult to recall, perceive or produce initial /lp/ clusters). *Phonotactic knowledge* concerns how cognitive functions are structured so as to yield these behaviors. The discussion here assumes that phonotactic knowledge distinguishes among forms in terms of their degree of well-formedness. Favored structures are well-formed; disfavored structures are ill-formed.

Note that phonotactic knowledge may allow speakers to disfavor structures to varying degrees. For example, consider the fricatives /f, v, h/ in word final position. In English sequences with /h/ in this position are completely absent, while /f/- and /v/-final sequences (laugh, live) are attested—suggesting that word-final /h/ may be strongly disfavored by English speakers' phonotactic knowledge. However, although both are attested, /v/ is much less frequent than /f/ in this position; /v/ is found in fewer words and occurs with a lower frequency in running speech. This suggests that relative to /f/, word-final /v/ may be disfavored to some degree. Phonotactic knowledge may therefore also make gradient distinctions in well-formedness.

2.1.2 Generative phonological models

There is a long history in linguistic theory of seeking to develop theories of phonotactic knowledge. "Whereas traditional phonology generally gives rules for articulating all sounds… and stops there," Saussure (1916: 51) writes, "combinatory phonology limits the possibilities and defines the constant relations of interdependent phonemes." The dominant theoretical framework since the time of Chomsky & Halle (1968) has been generative grammars. A generative grammar specifies a relation mapping a set of underlying phonological structures to a set of surface phonological structures (Smolensky, Legendre, & Tesar, 2006; n.b. the discussion here assumes a probabilistic formulation of this relation). Such grammars model phonotactic knowledge by specifying probability distributions over the set of surface phonological structures; well-formed structures are assigned higher probability than ill-formed structures.

Most work in this tradition has focused on a binary distinction between legal vs. illegal strings, characterizing the latter as categorically ill-formed and the former as categorically well-formed (see Goldsmith, 1995, for a review)[2]. Categorically well-formed structures are all generated by the grammar with equal probability (i.e., they are all equally well formed); categorically ill-formed structures are never generated by the grammar (i.e., they are all equally ill-formed). For example, if we believe that the phonotactic knowledge of an English speaker specifies that /h/ is ill-formed in word final position, we can represent this within a generative grammar in two ways. First, we can alter the probability distribution over underlying structures. For example, we could ban underlying representations containing /h/ in this position (e.g., morpheme structure constraints). A second mechanism involves altering the structure of the grammatical function. For example, we can ban any mapping (as defined by a set of rules or a constraint ranking) that allows /h/ to be generated in this position.

More recently, interest has grown in modeling gradient distinctions in phonotactic knowledge. This has been most frequently modeled in terms of generation probability. Less well-formed structures have lower generation probability than more well-formed structures. For example, if we believe that the phonotactic knowledge of an English speaker specifies that /v/ is disfavored in word final position relative to /f/, we can represent this within a generative grammar by (a) assigning underlying representations containing /v/ in this position lower probability than comparable representations containing /f/ (e.g., Frisch, Pierrehumbert, and Broe, 2004, assume a gradient constraint on roots in Arabic) and/or (b) assigning a lower probability to any mapping resulting in /v/ rather than /f/ in this position. A variety of formal mechanisms have been specified to assign probabilities to mappings: within derivational theories by associating probabilities to rule applications (e.g,. Coleman and Pierrehumbert, 1997); or within Optimality

[2] Categorical distinctions between possible and impossible structures were also the primary focus of a good deal of pre-generative research (e.g., Bloomfield, 1933; Harris, 1951; Trubetzkoy, 1939; Whorf. 1940).

Theoretic (Prince & Smolensky, 1993) approaches by assigning probabilities to constraint rankings (e.g., Boersma and Hayes, 2001) or output candidates (e.g., Coetzee, 2006; Hayes & Wilson, 2008). Further extensions to these formal mechanisms can allow generative grammars to represent gradient distinctions in well-formedness among unattested structures (e.g., for an English speaker, assigning varying probabilities to the generation of unattested initial clusters /fn/ and /zg/; Davidson, 2006b).

The discussion below examines the influence all of these various types of well-formedness distinctions on behavior: categorical distinction between attested and unattested strings (e.g., in English , final /f/ is possible, but final /h/ is not); gradient distinctions among attested strings (e.g., final /v/ is disfavored relative to final /f/); and gradient distinctions among unattested strings (e.g., for an English speakers, varying preferences for initial /fn/ vs. /zg/). In each case, these well-formedness distinctions are part of the (hypothesized) mental knowledge that underlies behavior. Note that certain theories use objective measures (e.g,. relative frequency or probability of structures in corpora) to estimate these mentally represented distinctions in well-formedness. It is important not to confuse these two notions; although the objective measures are used to *estimate* phonotactic knowledge, it is the mentally represented distinctions in well-formedness that are casually involved in producing behavior.

2.2 Inferring phonotactic knowledge from well-formedness judgments

To develop models within the various formal frameworks discussed above, researchers have drawn inferences concerning our knowledge of well-formedness distinctions based on various types of empirical data. Much of this work relies on categorical judgments of acceptability or possibility of various forms (e.g., "Is /zɑh/ a possible English word?"). (These judgments may be systematically organized and codified in a written grammar.) The patterns identified within this set of judgments (e.g., the tendency to judge forms with word-final /h/ as unacceptable) then inform the construction of generative grammatical models (e.g., the postulation of morpheme structure constraints banning /h/ in this position). Judgments can also inform models by providing a test of their predictions. For example, suppose a model classifies a form as categorically well-formed. If the form is judged to be unacceptable, this provides some evidence against the model; if it is acceptable, the evidence is consistent with the model.

Below, we briefly review three issues with using this type of data to inform grammatical theories. The first is purely methodological; more quantitative methods are required to accurately assess judgments. We then turn to two more substantive issues with this work. These issues reflect the failure of this type of research to consider psychological realism—how the cognitive system is structured so as to yield judgments. As discussed below, these issues are likely to lead to errors in inferring well-formedness distinctions from behavior.

2.2.1 Issue 1: Quantitative analysis of behavior

Although the collection of judgments and identification of patterns within them is done with great care and precision, it often does not do complete justice to the complexity of the underlying behavioral data from judgments. Frequently (as in many written grammars) acceptability judgments are codified as binary distinctions in well-formedness. (This reflects the assumption that acceptability judgments are a more or less direct reflection of phonotactic knowledge—a conflation of behavior with the mental representation of degrees of well-formedness[3].) Critically, this binary categorization obscures a fair amount of variation and gradience in participant responses (see also Bard, Gurman, Robertson and Sorace, 1996, for discussion).

[3] See Culbertson & Gross (in press) for a discussion of similar issues in grammaticality judgments.

The problem of limiting judgments to binary distinctions is addressed by studies that utilize more quantitative assessments of participant's judgments. For example, Greenberg and Jenkins (1964) asked participants to provide a numerical estimate of how far a stimulus word was from English (a technique called free magnitude estimation). Other studies have used rating scales: for example, Vitevitch, Luce, Charles-Luce, and Kemmerer (1997) asked participants to rate stimuli on a scale from 1 (bad example of an English word) to 10 (good example of an English word). The above scales ask participants to judge how wordlike a given stimulus is; other studies focus more on "acceptability." For example, Berent and Shimron (2003) asked participants to rate, from best to worst on a 5 point scale, how a stimulus sounds. Bailey and Hahn (2001) asked participants to rate how typical a stimulus sounds (on a scale from 1 to 9).

Studies using these types of tasks report that participants reliably judge structures that are classified by grammatical models[4] as ill-formed as being less acceptable or wordlike than structures classified as well-formed (e.g., Arabic: Frisch and Zawaydeh, 2001; Cantonese: Kirby and Yu 2007; English: Greenberg and Jenkins, 1964; Pierrehumbert, 1994; Hebrew: Berent and Shimron, 1997; Hindi: Ohala, 1983; Tagalog: Zuraw, 2007; Turkish: Zimmer, 1969).

Cross-linguistic research has also documented that participants' ratings of attested items correlate with relative degrees of well-formedness (as predicted by grammatical models incorporating gradient distinctions). Work using judgments to test gradient models of phonotactic well-formedness has a long history; it was a key motivation behind Greenberg and Jenkins' (1964) seminal wordlikeness study. They in fact found that English speakers' judgments were graded, consistent with the mental representation of gradient distinctions in well-formedness. Subsequent studies utilizing multiple methodologies in many languages have

[4] These models may be generative (as described above) or based on simpler n-gram statistics (e.g., biphone frequency).

also documented gradient distributions in participants' judgments. In studies examining categorical judgments by English speakers (e.g., "is this a possible English word?"), mean ratings across participants or items are gradient (e.g., Coetzee, 2008; Coleman and Pierrehumbert, 1997; Dankovicova, West, Coleman, and Slater, 1998; Scholes, 1966). Gradient distinctions are also found when individual English speakers make use of scales on each item (e.g., Bailey and Hahn, 2001; Dankovicova et al., 1998; Frisch et al., 2000; Ohala and Ohala, 1986; Shademan, 2006, 2007; Vitevitch et al. 1997). Similar results have been reported in a diverse set of other languages (Arabic: Frisch and Zawaydeh, 2001; French: Perruchet and Peerman, 2004; Korean: Lee, 2006; Tagalog: Zuraw, 2007).

Although quantified observations allow for a more nuanced picture of the distributional properties of behavioral responses, the precise connection of these observations to the underlying phonotactic knowledge is not entirely clear. Scant attention has been paid to how phonotactic knowledge is deployed in real time to yield judgments of well-formedness (see Schütze, 1996, for detailed discussion of similar issues in grammaticality judgments). Without clarifying, to some degree of approximation, how this task is performed, it is difficult to determine the precise implications of these results. The sections below consider two areas where this lack of clarity could lead to incorrect inferences regarding phonotactic knowledge.

2.2.2 Issue 2: Dynamic weighting of multiple factors in judgments

It is likely that as in other decision tasks word-likeness or acceptability judgments are not a pure reflection of well-formedness; judgments most likely reflect a combination of factors. Following this latter assumption, most research assumes judgments reflect both phonotactic well-formedness as well as similarity to existing lexical items (e.g., Bailey and Hahn, 2001;

Shademan, 2007)[5]. These two factors are conceptually distinct. There are forms that are unattested but well formed (e.g., *hing*). Measures of similarity to existing lexical items will be sensitive to such absences, whereas a pure measure of phonotactic well-formedness would not. Empirical results suggest that these two factors independently contribute to judgments. When similarity to existing items is controlled, forms that are classified as ill-formed are still judged less wordlike than those classified as well-formed (e.g., Arabic: Frisch and Zawaydeh, 2001). For English nonwords classified as relatively ill-formed, wordlikeness judgments correlate with degrees of well-formedness, not similarity to existing lexical items (e.g., Frisch, Large, and Pisoni, 2000; Coetzee, 2008). Finally, regression analyses can examine the degree to which judgments are influenced solely by degrees of well-formedness, similarity to existing items, or some combination of the two factors. Such analyses show well-formedness exerts an independent effect on judgments (e.g., Cantonese: Kirby and Yu, 2007; English: Albright, 2009; Bailey and Hahn, 2001; Shademan, 2006, 2007). (Note that these two factors do interact in judgments; see Shademan, 2007, for discussion).

Although these studies have taken into account the influence of multiple factors on participant judgments, little work has addressed the possibility that these influences are not static. A basic finding in psychology and psychophysics across many paradigms and domains is that decision processes are dynamic—i.e., sensitive to the context in which they are presented (see Vickers & Lee, 1998, for a general review). It is therefore unlikely that a simple static function maps judgments to internally represented well-formedness distinctions. There are

[5] The potential role for multiple factors in judgments has long been recognized. Greenberg and Jenkins (1964) concluded judgments reflect similarity to existing words. Writing about the same results a few years later Jenkins (1966) instead appealed to “systematic relations or bodies of rules” that speakers have internalized. Chomsky & Halle (1968: 416-418) discussed a function for determining acceptability that is sensitive both to phonological rules as well as the particular structure of the lexicon.

number of examples of such contextual effects in psycholinguistics. Consider the well-studied task of lexical decision (where participants judge whether an auditory or visual stimulus corresponds to a word in one of their languages). Judgments are influenced by the composition of nonword fillers (e.g., more wordlike fillers tend to slow responses), the proportion of high vs. low frequency targets, and repetition of target items (see Ratcliff et al. 2004 for a recent review). In same-different judgments with nonword auditory stimuli (e.g., is "bep" different from "mep"?), Vitevitch (2003) found the degree to which filler items are composed of real lexical items leads to shifts in the relative weighting of different factors. The use of mostly word fillers lead to a greater influence of similarity to existing lexical items, whereas mostly nonword fillers leads to a greater weighting of well-formedness. Critically, recent results suggest that wordlikeness judgments are similarly sensitive to properties of the experimental context in which the judgment is given. Shademan (2006, 2007) reported changes to the relative weighting of similarity to lexical items and well-formedness in judgments depending on whether the stimulus set contains both words and nonwords or is composed of nonwords only. Although her results are not clearly consistent across analysis methods, some results suggest that when words are excluded from the stimulus set participants' judgments tend to more strongly reflect well-formedness. Finally, Shademan (2007) reported that relative to those of young adults, the judgments of healthy older individuals are more sensitive to similarity to existing lexical items. These findings provide some preliminary support for the claim that wordlikeness judgments are context dependent—just like decisions in many other cognitive domains.

If factors other than well-formedness exert a *variable* influence on wordlikeness judgments, utilizing such judgments to inform linguistic theories becomes less straightforward. For example, suppose including real words as well as nonwords within in an experiment causes individuals' judgments to more strongly reflect similarity to existing lexical items. This may cause judgments to be less sensitive to gradient distinctions in well-formedness. Without awareness of such effects, we might mistakenly conclude that well-formedness distinctions are mentally represented in a more categorical fashion than they actually are. It is imperative for researchers using wordlikeness judgments to investigate how contextual factors influence judgment performance. This will allow for development of more complete models of the judgment process involved in this task (in much the same way that extensive investigation of context has enriched theories of lexical decision performance).

2.2.3 Issue 3: The interface of judgment processes with other cognitive processes

As repeatedly emphasized above, any behavior reflects the complex interaction of multiple psychological capacities. Wordlikeness judgments are no different; in order to make a wordlikeness judgment, one must perceive the acoustic structure of the form, assign a phonological parse to it, etc. . However, such relations have been left largely unspecified in the literature. The failure to articulate how judgments are situated within the cognitive system could lead to misinterpretations of behavioral data.

Suppose that in perception the phonological grammar corresponds to some specific cognitive process. Judgment processes respond to the output of this grammatical component of the cognitive system. However, this does not mean that judgments are influenced solely by the grammar; this component is but the last link in a causal chain of processes mediating the stimulus and the wordlikeness judgment.

Suppose we present an English speaker with two stimuli: one containing a word-final /h/, the other a word-final /s/. We then ask them to characterize the relative acceptability of the two stimuli. One possible result is that they will judge the form with /h/ as being less acceptable.

Does this imply that their grammatical processes encode word-final /h/ as less well-formed than word-final /s/? Not necessarily. It is possible that some well-formedness effects emerge in earlier stages of processing. For example, ill-formed sequences may be corrupted or distorted by relatively early perceptual processes (e.g., when exposed to an /h/ in word-final position, an English listener may have difficulty perceiving the intended sound). There is some empirical support for such a possibility; studies in several languages have shown that in fairly basic perceptual tasks hearers have difficulty perceiving categorically ill-formed stimuli—such that their perception is distorted towards a well-formed percept (e.g., Dupoux, Kakehi, Hirose, Pallier, and Mehler, 1999; see below for further discussion). This "repair" of the perceptual input may resulted in a corrupted or distorted representation (i.e., it may be an incomplete repair—a poor exemplar of the well-formed percept). This distorted input to grammatical processes might then cause hearers to judge a stimulus as less word-like. The distorted quality of the input—rather than the mental representation of well-formedness between /h/ and /s/—would give rise to the judgment of word-final /h/ as less acceptable than word-final /s/.

Alternatively, suppose the speaker judges word-final /s/ and /h/ to be equally acceptable. Does this imply that their grammatical processes encode word-final /h/ and /s/ as being equally well formed? Not necessarily. Suppose that perceptual repairs do not resulted in distorted inputs but instead produce completely well-formed representations. Under this scenario, the repair converts a categorically ill-formed representation to a categorically well-formed representation (see below for further discussion). A word-final /h/ stimulus is therefore transformed by perceptual processes into a completely well-formed word-final /s/. Even if the participant's grammar encodes word-final /h/ as less well formed, this perceptual transformation prevents grammatical processes from influencing judgment behavior.

We cannot draw a simple, direct connection from judgments to the mental representation of degrees of well-formedness. One means of attempting to circumvent this issue is to assume that the phonological component of the grammar is in fact distributed across multiple psychological capacities. Thus, the influence of well-formedness on more basic perceptual processes still has implications for the structure of the grammar. This may be a promising approach, but it would require a great deal of specification to become a plausible hypothesis. For example, returning to the issue raised above, if the multiple psychological processes encoding the grammar are each subject to independent contextual variation, the problem of relating judgments back to grammatical structure becomes many orders more complicated. Assuming the grammar is distributed over multiple psychological processes does not eliminate the need to consider the functional architecture of these processes; if anything, it makes such issues even more critical.

2.3 The perils of avoiding psychological realism

The common use of binary classifications to characterize judgments is a clear methodological issue in many studies. But the exclusive use of well-formedness judgments, disconnected from any theory of human language processing, has much deeper flaws that are likely to yield incorrect inferences regarding the nature of our knowledge of well-formedness.

In order to correctly draw inferences from on-line behavioral data, linguistic theories must be situated within specific tasks that occur in real time. Current research into well-formedness judgments has not adequately addressed this problem; neither the judgment process itself nor its interface with other cognitive systems have been well articulated or empirically explored. As discussed above, this lack of clarity is likely to lead to not only imprecise but also faulty conclusions regarding the nature of phonotactic knowledge. Because all linguistic data (including distributional data in corpora or dictionaries) ultimately arises from on-line behavior, interpretation of such data requires serious consideration of the symphony of coordinated cognitive functions that give rise to that behavior. Of course, it is highly likely that current theories will be incomplete in many respects; the cognitive functions will be less than completely specified, their interactions only partially spell-out. But even a partial account will provide purchase on many of the issues identified above. It will provide *some* justification for linking behavioral observations to the cognitive component(s) of interest.

This is not to say that well-formedness judgments bear *no* relation to phonotactic knowledge. It's highly likely that phonotactic knowledge does contribute in some way to well-formedness judgments. The critical issue is that without specifying the nature of this contribution in greater detail, we cannot draw inferences from behavioral data. We have no basis for arguing that variation in judgments reflects the function(s) of interest rather than some other function involved in the judgment task.

3 How can psychological realism help resolve theoretical issues in linguistics?

3.1 Utilizing psychological realism

To establish links between behavioral data and the structure of particular cognitive functions, psychological research adopts two basic strategies. One is through detailed examination of particular tasks. Researchers adopt theories of a particular function and its interaction with other functions; empirical research is then used to test and refine these theories (both in terms of the structure of functions as well as their interactions). A second approach is to seek converging evidence from a wide variety of tasks. Many cognitive functions are typically assumed to be utilized in a wide variety of tasks. For example, psycholinguistic theories of speech perception typically assume that there is a common set of cognitive functions engaged in all perceptual tasks. At some level this is obviously true; for any perceptual task, the hearer must at a minimum perform some analysis of the speech sounds (otherwise there would be no stimulus to perform the task on!). Researchers can rely on these functional commonalities to draw on performance across a wide range of tasks to constrain theories of specific cognitive functions.

The sections below examine how these strategies have been used to examine the nature of phonotactic knowledge in both speech perception and production. The first step is to establish that phonotactic knowledge is in fact utilized in both perception and production. To support this claim, the functional architecture of core processes involved in perception and production is outlined. These architectures characterize how phonotactic knowledge is deployed on-line in these behavioral tasks. Converging evidence from a variety of behavioral tasks that engage these core processes is then reviewed. This body of work provides support for the hypothesized functional architecture's claim that phonotactic knowledge plays a critical role in these tasks.

Armed with some validation of our linking assumptions, we can then turn to more detailed questions regarding the nature of phonotactic knowledge. In the course of reviewing data from perception and production, we touch on two types of results that bear on two key issues in linguistic theory: the relationship between grammar and lexicon and the gradient vs. categorical nature of phonotactic knowledge.

Within the generative tradition systematic aspects of phonological knowledge (e.g., phonotactics) have typically been reflected by the structure of the grammar. This is seen as a distinct component of the cognitive system from that of the (phonological) lexicon (Chomsky & Halle, 1968). This specifies those phonological structures which correspond to lexical items of the language (i.e., words). However, this distinction has not been universally adopted (even with the generative tradition. Recent work in linguistic theory has strongly emphasized the view that lexicon and grammar are highly intertwined, inseparable components of the cognitive system (e.g., Burzio, 1996; Bybee, 2001).

The second issue addressed below concerns the nature of phonotactic well-formedness. As discussed above, many generative grammatical models have focused on categorical distinctions in well-formedness. However, more recently there has been a growing interest in modeling gradient distinctions in grammatical knowledge. There has been considerable disagreement in the literature regarding whether the former or latter perspective best characterizes the nature of our linguistic knowledge (see, e.g., Newmeyer, 2003, and associated commentaries in *Language* 81(1) for a recent discussion).

As discussed below, behavioral data—interpreted within functional frameworks for speech perception and production—can help resolve these questions. With regard to the first issue, cognitive functions encoding phonotactic knowledge are distinct from (but interact with) those encoding word-specific knowledge; this is consistent with the claim that lexicon and grammar are independent components of linguistic knowledge. Second, gradient distinctions in phonotactic well-formedness influence perception and production—suggesting phonotactic knowledge is not limited to categorical distinctions.

3.1 Phonotactic knowledge in speech perception

3.1.1 Phonotactic knowledge within the functional architecture of speech perception

As discussed above, many psycholinguistic theories of speech perception assume two broad functional stages in speech perception tasks involving stimuli up to the size of single spoken words (e.g., McClelland et al., 2006). The first stage takes as input a relatively fine-grained representation of acoustic information (e.g., acoustic features) and produces as output a pre-lexical representation elaborating the linguistic structure of the acoustic input (e.g., by specifying segmental and prosodic structure). The second stage uses this pre-lexical representation to retrieve a lexical representation of the utterance (e.g., a unitary whole-word representation <DOG> which is used to access semantic and syntactic information). In this discussion the first stage will be referred to as *pre-lexical processing* and the second as *lexical processing*. (Note: these two stages of processing may overlap and interact; see McClelland et al. 2006; Norris et al., 2000, for discussion). This basic decomposition of the system into two distinct lexical and pre-lexical representations has received support from a variety of perceptual tasks (for recent reviews, see Gaskell, Quinlan, Tamminen, & Cleland, 2008; McQueen, Cutler, & Norris, 2006).

In this two-stage framework, phonotactic knowledge is assumed to be encoded within pre-lexical processes; the speed and accuracy with which pre-lexical representations are activated reflects their phonotactic well-formedness. For example, with respect to categorical well-formedness Dupoux et al. (1999) assume that categorically ill-formed sequences lack stored suprasegmental representations possessed by categorically well-formed sequences (specifically, demisyllables). The lack of such stored representations leads to greater errors and slowed processing times on ill-formed sequences. With respect to gradient distinctions in phonotactic well-formedness, Luce, Goldinger, Auer and Vitevitch (2000; see also Vitevitch and Luce, 1999) assume that representational units within relatively well-formed sequences facilitate one another—allowing for more rapid and accurate retrieval for more vs. less well-formed structures (for discussion of similar mechanisms, see Newman, Sawusch and Luce, 1999; Norris et al., 2000).

This functional architecture predicts a number of behavioral manifestations of phonotactic well-formedness. This section focuses on sub-word perceptual tasks in which participants identify, categorize, or discriminate auditory stimuli without being required to explicitly attend to properties of whole words. (Note that similar effects are found in tasks that do require attention to whole words; e.g., lexical decision: Berent, Everett, and Shimron 2001; word segmentation: McQueen, 1998; Van der Lugt, 2001.) Under some accounts (e.g., McClelland et al., 2006), performance in sub-word tasks directly reflects pre-lexical representations; phonotactic constraints therefore directly influence sub-word task performance. In contrast, other accounts (e.g., Norris et al., 2000) augment the two-stage architecture above with an additional set of representations that receive input from both pre-lexical and lexical representations. These *decision* representations support performance in sub-word level tasks (e.g., allowing participants to judge whether a phoneme was present in a stimulus). Note, however, that these representations directly receive input from pre-lexical representations; this account therefore also predicts phonotactic well-formedness should influence behavior in sub-word tasks.

Although sub-word tasks do not explicitly invoke lexical representations, most current accounts predict that lexical factors can exert an indirect influence on performance (this interaction between lexical and phonotactic knowledge will be discussed in greater detail below). In some theories (e.g., McClelland et al., 2006), lexical and pre-lexical processes interact with one another. Under other accounts, sub-lexical decision representations receive input both from pre-lexical and lexical processes (e.g. Norris et al., 2000). Therefore, under both architectures, both lexical and sub-lexical properties exert an influence on sub-word task performance.

3.1.2 Evidence that pre-lexical processes disprefer categorically ill-formed structures

Categorical distinctions in phonotactic well-formedness clearly exert an influence on speech perception. Many researchers have noted that relative to categorically well-formed phonological structures (e.g., attested English clusters such as /pl/), categorically ill-formed structures (e.g,. unattested English clusters such as /dl/) are more difficult to perceive. This has been noted for quite some time; for example, Trubetzkoy (1939: 62-64) discusses how the constraints of one's native language impede accurate identification of sound sequences from foreign languages. Converging evidence from a variety of perceptual tasks has provided more systematic confirmation of these observations. All of these studies assume that categorical ill- vs. well-formedness of a phonological structure is roughly indexed by its absence vs. presence in lexical items of a language. For example, no English word ends in /h/; the phonotactic knowledge of English speakers therefore specifies that word-final /h/ is categorically ill-formed.

3.1.2.1 Identification tasks. The seminal study of Brown and Hildum (1956) found English listeners made more transcription errors on categorically ill-formed sequences than they did on well-formed sequences. More recent studies in other languages have confirmed this result, both in terms of accuracy (e.g., French: Hallé, Segui, Frauenfelder and Meunier, 1998; Japanese: Dupoux et al., 1999) as well as reaction times (i.e., in phoneme monitoring tasks; Segui, Frauenfelder, and Hallé, 2001). Furthermore, errors on these categorically ill-formed sequences are not random; they tend to result in categorically well-formed sequences (e.g., French speakers mistranscribe ill-formed /dl/ as well-formed /gl/: Hallé et al., 1998). This suggests that ill-formed stimuli activate well-formed pre-lexical representations.

Identification of perceptually ambiguous stimuli is biased towards categorically well-formed sequences (e.g., English: Massaro and Cohen, 1983; Japansese: Dupoux et al., 2001). For example, when /r/ but not /l/ forms a categorically well-formed cluster (e.g., /tri/ vs. */tli/), English listeners' categorization of stimuli on a synthesized /r/-/l/ continuum is biased towards /r/. The opposite pattern is found when the well-formed cluster is composed of /l/ (e.g., /sli/ vs.

*/sri/). (See Moreton, 2002, for recent related results.) Similar results have been found in dichotic fusion. When conflicting stimuli are presented to each ear, their perceptual fusion into a single stimulus tends to result in categorically well-formed sequences (e.g., in English, /b/+/l/ is resolved as /bl/ not /lb/; Day, 1968, 1976; for similar results in Portuguese, see Morais, Castro, Scliar-Cabral, Kolinsky, and Content, 1987).

<u>3.1.2.2 Discrimination tasks</u>. Participants have difficulty discriminating (at least some) ill-formed from well-formed sequences. Dupoux et al. (1999) documented Japanese speakers' difficulty (slower reaction times, greater errors) in discriminating categorically ill-formed consonant-consonant sequences from the corresponding categorically well-formed consonant-vowel-consonant sequence (e.g., *ebzo* is difficult to discriminate from *ebuzo*). Davidson (2007) reported similar perceptual difficulties for English speakers. Kabak and Idsardi (2007) found Korean speakers have difficulties in discrimination tasks with sequences that are categorically ill-formed with respect to syllable-conditioned patterns (e.g., */g/ in coda)[6]. For Japanese speakers, Dehaene-Lambertz, Dupoux and Gout (2000) reported similar results using an implicit electrophysiological measure of discrimination; furthermore, Jacquemot, Pallier, LeBhian, Dehaene and Dupoux (2003) found distinct patterns of neural activity for categorically well- vs. ill-formed sequences during discrimination tasks.

<u>3.1.2.3 Perceptual difficulties with categorically ill-formed sequences arise in pre-lexical processing</u>. To establish the evidential status of the observations from identification and discrimination tasks, it is crucial to show the effects derive from pre-lexical processes. With respect to the involvement of more basic auditory processing, it is unlikely that these identification results reflect inherent uncertainties in the acoustic signal. Listener groups with

---

[6] Note that Kabak and Idsardi (2007) failed to find similar discrimination deficits for heterosyllabic sequences that are categorically ill-formed in Korean (e.g., *[k.m]).

varying language backgrounds exhibit different patterns of performance (e.g., Dupoux et al., 1999, showed that French listeners have no difficulty perceiving sequences that are difficult for Japanese participants). Acoustically identical ambiguous tokens (e.g., stimuli on an /r/-/l/ continuum) are processed differently depending on the context in which they appear (Massaro and Cohen, 1983), again suggesting that it is not intrinsic properties of the acoustic signal that gives rise to these effects. Finally, with respect to lexical effects, Dupoux, Pallier and Kakehi (2001) found comparable patterns of transcription errors regardless of whether the error results in a word or nonword—suggesting these phonotactic effects are not a simple reflection of a bias to report existing lexical items (see Day, 1968, for similar results in dichotic fusion tasks).

3.1.2.4 Implications for theories of phonotactic knowledge. Converging results from a variety of perceptual tasks and languages provide support for the functional architecture outlined above. They are broadly consistent with the claim that categorical distinctions in phonotactic well-formedness are encoded by pre-lexical processes in perception. Beyond simply validating our linking assumptions, they also provide insight into the relationship between cognitive functions reflecting grammatical (e.g., phonotactic well-formedness) and lexical (e.g., word-specific) knowledge. The architecture above assumes that these are encoded by two distinct functions (pre-lexical and lexical); consistent with this, empirical studies show that phonotactic well-formedness exerts an influence on perceptual performance independent of lexical knowledge. This supports the claim that lexicon and grammar form distinct components of our phonological knowledge.

3.1.3 Evidence that pre-lexical processes disprefer relatively ill-formed structures

3.1.3.1 Converging evidence for perceptual sensitivity to gradient variations in well-formedness. More recent research suggests that degrees of well-formedness within attested sequences of a language lead to relative ease of perceptual processing. These studies have indexed degrees of well-formedness using *phonotactic probability*—the probability of a given phonological structure within a language. This can be indexed by various n-gram measures (e.g., relative probability of single phones such as /s/ or biphones such as /st/, estimated using either a corpus of utterances or a lexical database). Pitt and Samuel (1995) argued that stimuli composed of phonological structures with high phonotactic probability are detected more quickly in monitoring tasks. Pitt and McQueen (1998) reported that identification of ambiguous sounds is biased towards structures with higher phonotactic probability. Vitevitch and Luce (1999) examined English speakers' speeded same/different judgments. For nonwords, the reaction time for correct "same" responses is faster for syllables with high vs. low phonotactic probability (see Vitevitch, Pisoni, Kirk, Hay-McCutcheon, and Yount, 2002, for similar results for cochlear implant patients with good word recognition abilities). Finally, Coetzee (2008) found that in the identification of perceptually ambiguous sounds English listeners show a dispreference for attested forms that violate phonotactic constraints on consonant co-occurrence (i.e., the Obligatory Contour Principle).

These effects can be plausibly attributed to pre-lexical processing. Research by Vitevitch and colleagues suggest that effects in discrimination tasks cannot be attributed to more basic auditory or lexical processes. Two critical observations argue for this point. First, behavioral effects change as a function of task. For tasks that emphasize lexical properties (e.g., lexical decision), performance primarily reflects similarity to lexical items—in contrast to the phonotactic probability effects observed in discrimination tasks. This is found even when identical stimuli are used in both tasks (Vitevitch and Luce, 1999). Second, within the same task and items, the degree to which filler items are composed of real lexical items leads to shifts in the influence of lexical vs. phonotactic factors (i.e., the use of mostly real word fillers lead to a greater influence of lexical factors; Vitevitch, 2003). Since identical stimuli yield contrasting behavioral effects across tasks, it is unlikely that the effects derive from purely acoustic properties of the stimuli. With respect to lexical effects, the ability of participants to shift from phonotactically- to lexically-driven behavior within and across tasks is consistent with independent contributions of both pre-lexical and lexical representations to discrimination.

Distinct electrophysiological responses have been documented for high vs. low probability sequences (Dutch: Bonte, Mitterer, Zellagui, Poelmans and Blomert, 2005). However, it is unclear if these reflect pre-lexical processes. Although Pylkkänen, Stringfellow, and Marantz (2002) reported one electrophysiological measure is influenced only by phonotactic probability, Stockall, Stringfellow and Marantz (2004) found that the same measure is also sensitive to similarity to existing lexical items.

Inconsistent results with respect to the role of lexical vs. pre-lexical processes have also been reported in studies of identification of ambiguous stimuli. Newman et al. (1999) reported that lexical effects on identification of perceptually ambiguous stimuli are suppressed for stimuli with very high phonotactic probability. Pitt and McQueen (1998) reported phonotactic probability effects on identification when similarity to existing lexical items is equated. These two studies suggest an independent contribution of phonotactic probability on perceptual processing. However, Magnuson, McMurray, Tanenhaus, and Aslin (2003a) found that when phonotactic probability effects conflict with similarity to lexical items, the latter dominates performance (see Magnuson et al., 2003b; McQueen, 2003, for further discussion).

As made clear by these latter studies, sub-word tasks are clearly influenced by properties of both lexical and pre-lexical representations. However, finding such as those of Vitevitch

(2003) reviewed above suggest that pre-lexical representations make an independent contribution to sub-word task performance. This is consistent with an influence of phonotactic well-formedness on perception independent of that of the lexicon.

A final source of evidence for perceptual sensitivity to gradient distinctions in well-formedness comes from studies examining the perception of sequences which are unattested in one's native language. As noted above, it has been proposed that grammars can distinguish degrees of well-formedness among unattested forms (e.g., Coetzee, 2008; Davidson, 2006b). Berent, Steriade, Lennertz, and Vaknin (2007) examined English hearers' identification of clusters that are absent from English. These clusters vary in the degree to which they respect cross-linguistic markedness generalizations (e.g., /bn/, with a sonority rise, is less marked than /lb/, with a sonority fall). They found greater rates of misidentification for clusters which are cross-linguistically marked (see Berent & Lennertz, 2007; Peperkamp, 2007, for additional discussion of these findings; Berent, Lennertz, Jun, Moreno & Smolensky, 2008 for similar results with Korean listeners; and Berent, Lennertz, Smolensky, & Vaknin-Nusbaum, 2009 for extension of these results to nasal-initial clusters). This is consistent with the encoding of gradient distinctions in well-formedness (as indexed by cross-linguistic markedness) among unattested forms. Unlike the other perceptual studies reviewed above, these experiments are not simply limited to categorical contrasts between relatively ill-formed and well-formed groups of items. Berent and colleagues found that accuracy rates vary across the various ill-formed clusters—suggesting perceptual processes are sensitive to degrees of well-formedness

3.1.3.2 Implications for theories of phonotactic knowledge. Studies of gradient well-formedness provide further evidence that lexicon and grammar form distinct components of our phonological knowledge. Additionally, they provide support for the claim that our knowledge of phonotactics can do more than distinguish categorically well- from ill-formed structures. Among both attested and unattested form, our knowledge of phonotactics encodes gradient distinctions in well-formedness.

3.2 Phonotactic knowledge in speech production

3.2.1 Phonotactic knowledge within the functional architecture of speech production

Like psycholinguistic theories of speech perception, production accounts assume the presence of whole-word lexical representations that mediate between phonological and syntactic/semantic information. In production, these whole-word representations serve as input to cognitive processes that manipulate sound structure. Generally, theories assume two broad stages of processing (for a recent review of theoretical perspectives and supporting evidence, see Goldrick and Rapp, 2007). The first (*phonological spell-out*) takes as input whole word representations and yields relatively abstract, coarse-grained phonological information as output (e.g., unprosodified segments, unspecified for featural content). The second stage of processing (*phonetic encoding*) takes these abstract representations as input and yields (more) fully specified representations of sound structure that drive subsequent articulatory planning and execution processes.

Phonotactic constraints are assumed to exert an influence on phonetic encoding processes with the result that structures that are more phonotactically well-formed are retrieved more rapidly and accurately than ill-formed structures. For example, Wheeler and Touretzky (1997) proposed that phonetic encoding assigns segments (retrieved during lexical phonological processing) to prosodic positions via a constraint satisfaction procedure. When segments are misordered in speech errors, licensing constraints tend to block ill-formed sequences from being produced (e.g., blocking the error "too blue" -> */tlu blu/). Goldrick and Larson (in press) accounted for effects of gradient distinctions in well-formedness by extending Warker and Dell's (2006) theory of phonetic encoding. Under this account, the strength of connections between lexical phonological and phonetic representations varies with phonotactic probability. Since error probability is related to relative activation levels (Warker and Dell, 2006), the more strongly activated high probability forms are more likely to occur as errors than low probability forms.

The sensitivity of phonetic encoding processes to phonotactic well-formedness predicts a number of effects in speech production. The discussion below focuses on tasks that do not involve a substantial perceptual component (e.g., immediate repetition/shadowing; Munson, 2001; Onishi, Chambers, and Fisher, 2002) or metalinguistic tasks (e.g., word blending; Treiman, Kessler, Knewasser, Tincoff, and Bowman, 2000). However, it should be noted that similar effects are observed in these studies.

3.2.2. Evidence that phonetic encoding processes disprefer categorically ill-formed structures

Production difficulties for phonotactically ill-formed structures have long been recognized by researchers in linguistics (e.g., Whorf, 1940). Subsequent behavioral studies have more systematically confirmed these observations. As in perception studies, categorical well-formedness distinctions are indexed by presence vs. absence in a language's lexical items. In elicited production studies, targets that are categorically ill-formed targets yield higher error rates than well-formed targets (Amharic and Chaha: Rose and King, 2007; English: Davidson, 2006a). The architecture above also predicts that when errors are made, it is likely that targets will be replaced by categorically well-formed structures (as was observed in perceptual errors; e.g., Hallé et al., 1998). Consistent with this, numerous transcription-based studies in various languages report that spontaneous speech errors rarely result in structures which are categorically ill-formed in a speaker's native language (e.g., English speakers rarely produced errors such as "miff" -> "mih"; Arabic: Abd-El-Jawad and Abu-Salim, 1987; English: Vousden, Brown and Harley, 2000; German: MacKay, 1972; Mandarin: Wan and Jaeger, 1998). Although higher rates of phonotactically ill-formed productions have been observed in experimentally elicited tongue twister productions[7], transcription analysis still reports a bias towards categorically well-formed structures (e.g., English: Butterworth and Whittaker, 1980).

Complementing work on native language phonotactic constraints, recent transcription studies have examined the effect of implicitly acquired categorical distinctions in well-formedness. Dell, Reed, Adams and Meyer (2000) exposed English speakers to sound distributions that suggest some structure is categorically ill-formed within the experiment. For example in one condition, /f/ is confined to onset, suggesting /f/ is categorically ill-formed in coda. For another set of participants, /f/ is confined to coda, suggesting /f/ is categorically ill-formed in onset. They find that virtually all speech errors result in forms that are categorically well-formed with respect to the experimental condition (e.g., in the former experimental condition, virtually all /f/ error outcomes appear in onset as in "ned"->"fed"; in the latter, /f/ errors appear in coda).

As with the perceptual data reviewed above, it is critical to establish that such effects reflect the cognitive component of interest—here, phonetic encoding processes. Dell et al.'s (2000) results strongly support this. The fact that participant behavior reflects relatively arbitrary well-formedness distinctions specific to each condition suggests that these effects do not reflect

---

[7] At a subsegmental level, recent instrumental studies suggest that speech errors can result in ill-formed combinations of articulatory gestures (e.g., simultaneous tongue tip and tongue dorsum raising during production of a stop; Goldstein, Pouplier, Chen, Saltzman and Byrd, 2007; McMillan, Corley, & Lickley, 2008; Pouplier, 2007, 2008). Findings such as these underscore the importance of using more quantitative measures of participant's behavior—not only in well-formedness judgments but in more prototypical 'laboratory' studies.

intrinsic motoric properties of ill-formed structures. With respect to lexical effects in the same paradigm, Goldrick (2004) finds the effects of implicitly acquired well-formedness distinctions are not eliminated by the exclusion of word error outcomes. Participants' error outcomes still respected experiment-specific well-formedness when responses like "fed" were excluded from the analysis (leaving only nonwords like "fep"). This suggests a simple lexical bias cannot account for the observed error patterns. (Furthermore, a simple transcriber bias cannot account for contrasting patterns across conditions; in both studies, the same transcribers analyzed each condition.)

3.2.2 Evidence that phonetic encoding processes disprefer relatively ill-formed structures

Among attested sequences, relatively well-formed structures are more accurately and rapidly processed than ill-formed structures. As in studies of speech perception, this has been indexed via phonotactic probability. Vitevitch, Armbrüster and Chu (2004) found pictures with high probability names have shorter naming latencies in English than pictures with low probability names. Laganaro and Alario (2006) found that pictures whose names are composed of high vs. low frequency syllables have shorter naming latencies in French. With respect to accuracy, elicited production studies (e.g., English: Kupin, 1982) and studies of individuals with aphasia (e.g., English: Blumstein, 1973) report greater accuracy for high vs. low probability structures.

With respect to error outcomes, the evidence is somewhat mixed. Some speech error analyses find errors are more likely to result in high vs. low probability sequences (e.g., English: Levitt and Healy, 1985; Motley & Baars, 1975). Similar results have been found in cases of aphasia (English: Buchwald, 2009; Buchwald, Rapp, & Stone, 2007; Goldrick & Rapp, 2007; French: Béland & Paradis, 1997; Italian: Romani & Calabrese, 1998). In a recent study utilizing the Dell et al. (2000) paradigm discussed above, Goldrick and Larson (2008) found that the probability of error outcomes is influenced by phonotactic probability. However, many studies have failed to find any asymmetries in the probability of contrasting error outcomes (e.g., Arabic: Abd-El-Jawad and Abu-Salim, 1987; English: Shattuck-Hufnagel and Klatt, 1979; Swedish: Söderpalm, 1979). Finally, in a series of studies Stemberger argued that although there is a general preference for higher frequency error outcomes, there is also a preference for lower frequency forms over (very high frequency) default structures (see Stemberger, 2004, for a recent review).

When phonotactic probability effects have been observed, they arguably reflect the operation of the cognitive component of interest—phonetic encoding. With respect to lexical effects, Vitevitch et al. (2003) reported phonotactic probability effects on latency even when items are matched in terms of similarity to existing lexical items. With respect to lower level articulatory processes, at least two studies have shown that performance on the same structures varies across task conditions (Goldrick and Larson, 2008; Laganaro and Alario, 2006) suggesting that intrinsic articulatory properties cannot account for these behavioral effects. Finally, two additional results provide positive evidence for a phonetic encoding locus for phonotactic probability effects. Laganaro and Alario (2006) reported effects of syllable frequency are eliminated when a secondary task disrupted phonetic encoding—suggesting this process is necessary for producing the behavioral effect. Goldrick and Rapp (2007) contrasted the performance of two aphasic individuals with deficits to phonological spell-out processes vs. phonetic encoding; effects of phonotactic probability are found only when errors arose subsequent to a phonetic encoding deficit. Thus, although both lexical properties and phonotactics well-formedness influence speech production, there is ample evidence that phonotactic well-formedness independently influences speech production.

3.2.3 Implications for theories of phonotactic knowledge

Results from speech production converge with those from speech perception. There is clear evidence that phonotactic and lexical knowledge are encoded by distinct cognitive functions—supporting the claim that grammatical and lexical knowledge form distinct components of our knowledge of phonological structure. Several studies also suggest that phonotactic knowledge encodes gradient distinctions in well-formedness. Specifically, a number of studies show not only differences between groups of relatively well- vs. ill-formed structures, but varying degrees of performance related to degrees of well-formedness. Davidson (2006a) reported that English speakers produce distinct error rates for different clusters that are unattested in English. (She attributes to these varying error rates to varying degrees of well-formedness reflecting the articulatory and perceptual properties of the targets; see Davidson, 2006b, for further discussion). In an elicited speech error task in two Ethiopian Semitic languages (Amharic and Chaha) Rose and King (2007) found distinct error rates for forms with different degrees of OCP violations. Finally, Goldrick and Larson (2008) found that the probability of speech errors varied according to experiment-specific phonotactic probability.

3.3 Phonotactic knowledge in other behavioral domains

3.3.1 Memory

Current theories postulate three core functional components for memory tasks (for recent reviews see Baddeley, 2003; Vallar, 2006). A *phonological analysis* component takes acoustic input of the to-be-remembered stimuli and outputs a phonological representation (similar to pre-lexical processing in theories of speech perception). The phonological representation output by this system serves as input to the *phonological short term store*, a limited capacity memory system. This interfaces with a *phonological output buffer*. This maintains the activation levels of phonological representations within the short term store and supports recall by interfacing between the short term store and other processes supporting spoken production (for discussion of how the buffer interacts with other production processes, see Goldrick and Rapp, 2007; Jacquemot and Scott, 2006). Many theories assume a variety of other speech perception and production processes also contribute to memory processing (e.g., lexical processes involved in perception and production; see Martin and Gupta, 2004, for a recent review).

The wide variety of processes involved in short-term memory tasks creates several opportunities for phonotactic well-formedness to influence processing. Focusing specifically on short-term recall tasks, Gathercole, Frankish, Pickering and Peaker (1999) outlined two broad hypotheses to account for such effects. Since phonotactic well-formedness influence pre-lexical perceptual processes, less well-formed structures may be less robustly stored than more well-formed structures. Alternatively, phonotactic well-formedness effects could arise in the context of either rehearsal or recall, as information from other (perceptual or production) processes sensitive to phonotactic well-formedness is used to restore or reconstruct decaying memory representations (a process termed 'redintegration'; Schweickert, 1993).

Many studies have shown that performance in memory tasks is correlated with measures of phonotactic well-formedness. Lists composed of structures with relatively high vs. low phonotactic probability have higher recall accuracy (English: Nimmo and Roodenrys, 2002; French: Majerus and Van der Linden, 2003; experiment-specific probability: Majerus, Van der Linden, Mulder, Meulemans, and Peters, 2004). Participants show better recognition memory for English nonwords with relatively high vs. low phonotactic probability (Frisch, Large, and

Pisoni, 2000). Finally, high probability have a greater likelihood of appearing intact in short term memory errors in English and Korean (Lee and Goldrick, 2008). These effects are not likely to derive from lexical factors, as phonotactic probability effects on accuracy are found when similarity to existing lexical items is controlled (although both influence accuracy; Thorn and Frankish, 2005; Lee and Goldrick, 2008; but see Roodenrys and Hinton, 2002).

These findings provide further evidence for both the independence of grammatical and lexical knowledge, as well as the representation of gradient distinctions in well-formedness. However, the precise functional locus (or loci) of these effects are unclear—they may derive from stimulus coding and/or redintegration processes (see Baddeley, 2003; Thorn and Frankish, 2005, for discussion). The evidential status of these results is thus less clear than studies of production and perception, where behavioral effects could be more clearly linked to a specific cognitive function.

3.3.2 Naturalistic corpora

Ultimately the goal of a psychological theory of real-time linguistic behavior is to move beyond highly controlled experimental tasks (such as those discussed above) to fully natural communication contexts. Unfortunately, given our limited understanding of many of the cognitive functions involved in natural communication, it is extremely difficult to clearly establish the evidential status of such data. For example, differences in relative frequency with which different phonological structures appear in corpora may not only reflect properties of a wide array of phonological processes (involved in perception, production, or memory processing) but also a range of *non*-phonological processes[8] (involving other dimensions of linguistic structure as well as more general cognitive processes). More generally, it is critical to

[8] Newmeyer (2003) raises similar points, although he draws quite different conclusions.

note that psychological realism is an issue for those that would commit sins of commission as well as omission. Just as it is incorrect to assume that some type of data is categorically irrelevant (e.g., experimental data from perception/production has no bearing on linguistic theory), it is also incorrect to assume all behavioral observations are relevant for evaluating theoretical claims.

3.4 Utilizing psychological realism: Progress and prospects

Adopting the perspective of psychological realism has allowed us to draw principled inferences concerning phonotactic knowledge from speech perception and production data. In contrast to word-likeness / acceptability judgments, in these behavioral domains the functional architecture of the core processes encoding phonotactic knowledge has been articulated. Although not reviewed in detail above, the basic structure of these architectures has received broad empirical support. With respect to phonotactic knowledge specifically, the sections above reviewed data from a range of behavioral tasks that make use of these core processes. The results provide converging evidence to support the assumption that phonotactic well-formedness influences processing in both modalities. Building on this empirical support for our linking assumptions, we have used data from these tasks to address two key issues in linguistic theory. Across modalities and tasks we find broad support for two claims: the phonological grammar and lexicon are distinct components of linguistic knowledge; and our knowledge of phonotactic well-formedness can encompass gradient as well as categorical distinctions. While this represents important progress on key issues, a great many theoretical questions are clearly still unresolved. The following sections discuss two such areas for future work.

3.4.1 Interactions between lexical and grammatical knowledge

As discussed in detail in the sections above, more recent work in each of these empirical domains has established that phonotactic constraints exert an effect independent of the lexicon. In perception, effects of phonotactic well-formedness derive from pre-lexical, not lexical processes; similarly, in production, effects arise specifically within phonetic encoding. This is incompatible with a strong claim that phonotactic and lexical knowledge are completely functionally indistinct. However, it should also be clear from the review above that both factors contribute to performance in a variety of domains; furthermore, in many cases their contributions are not simply additive. This suggests the independent knowledge systems representing these two factors must interact during processing. An interesting area for future research will be examining the degree and nature of interaction required to account for the observed patterns of behavior across domains. In developing this line of inquiry, linguistic research should capitalize on the extensive empirical and theoretical work in psycholinguistics that has examined interaction between lexical and sub-lexical phonological processes (for a recent review in perception, see McClelland et al., 2006; production, Goldrick, 2006).

3.4.2 Is phonotactic knowledge distributed across distinct psychological functions?

As discussed in section 1, a critical issue in psychological realism is establishing how components of linguistic theories link up to components of psychological theories. The discussion above has identified links *within* perception and production between phonotactic knowledge and specific cognitive functions. However, the link between these functions has not been discussed. Does this work suggest that phonotactic knowledge is distributed (i.e., across distinct pre-lexical processes in perception and phonetic encoding processes in production)? Or, alternatively, is there a one-to-one relation between phonotactic knowledge and a single cognitive function (i.e., a sublexical process shared across perception and production)?

The evidence on this issue is decidedly mixed. The fact that phonotactic well-formedness influences both perception and production (as well as memory, well-formedness judgments, and other tasks) certainly suggests strong similarities between the capacities engaged in these various processing domains. Furthermore, across domains phonotactic constraints appear to be similarly structured (e.g., reflecting gradient as well as categorical well-formedness distinctions). However, similarity does not necessarily imply identity, making it unclear if these data support a common capacity. In favor of the specialized capacity view, there is ample evidence that distinct 'repair strategies' (i.e., mappings from phonotactically ill- to well-formed structures) are invoked in perception and production (e.g., Kabak and Idsardi, 2007). However, these differing strategies could reflect the differential utilization of a single function specifying well-formedness (e.g., Smolensky, 1996). Existing empirical evidence is therefore unable to decide whether effects of phonotactic well-formedness reflect a truly general capacity or distinct but similar capacities specific to particular behavioral domains. This represents a critical area for future work exploring the relationship between concepts from linguistic theories and the psychological mechanisms involved in perception, production and memory.

4 Discussion: The Necessity and Poverty of Psychological Realism

This chapter has argued that psychological realism should critically inform not just the methodological but also theoretical facets of linguistic research. A psychological framework specifies the functional architecture of language processing—how linguistic knowledge is deployed in real time in specific behavioral tasks. Specifying this organization to some level of detail is a *necessary* step in drawing inferences from linguistic data (as all such data ultimately reflects real-time behavior). When this is not done, as in the case of well-formedness judgments, inferential errors are likely to occur. Stronger inferences can be drawn in behavioral domains such as speech perception and production where core aspects of the functional architecture underlying an array of behavioral tasks have been articulated and empirically justified via converging evidence from these tasks.

Although psychological realism is a critical component of linguistic research, it should not be the only component. An exclusive focus on the cognitive capacities of individual speakers offers an incomplete account of the structure of phonological systems. For example, a number of theorists have provided explanations of phonological phenomena that explicitly reject the idea that all aspects of the linguistic system reflect the capacities of individual language users. Two broad types of such proposals have been advanced. One set attributes some systematic aspects of phonological systems to the properties of linguistic populations (for a review, see Pierrehumbert, 2006). For example, Pierrehumbert (2001) attributes the coarse-grained nature of phonological patterns to their stability across individual differences in vocabulary. Another type of account focuses on the role that historical change plays in determining the synchronic structure of languages (e.g., Blevins, 2004). Such contemporary accounts of language structure have deep roots within linguistic thought. Historical explanations have long played a role in explanations of language structure (see Blevins, 2004: Chapter 3 for a review). With respect to the properties of linguistic populations, the tension between the role of the individual and collective has been apparent from the outset of modern linguistics. As noted by Saussure: "Language exists in each individual, yet it is common to all (1916: 19)." This leads him to speak of language both as something "psychological" which is "deposited in the brain (p. 19)" but also as an entity that is "outside the individual…; it exists only by virtue of a sort of contract signed by the members of a community (p. 14)." Adopting the perspective of psychological realism should not therefore lead one to ignore the role of linguistic *communities* in shaping language structure; our theories must strive for a balance between the group and the individual.

Claims of "psychological reality" and the use of psychological data in the development of linguistic theories have provoked strong negative reactions in both the linguistic and psychological research communities. Such reactions are at least partially attributable to a variety of misconceptions. Some are shared by practitioners of both disciplines; for example, conflating the issue of algorithmic/neural realizability with developing psychological accounts of human linguistic behavior. Other misunderstandings are primarily cross-disciplinary. On the one hand, psychologists have sometimes assumed that because linguists tend to posit generalized capacities (subserving many tasks) the constructs of linguistic theory are of a different kind than those posited by psychologists. In fact, both theories focus on highly abstract, (incomplete) characterizations of the cognitive system at the functional level of description. At most, cognitive psychological and linguistic theories merely differ in the *degree* of abstraction they adopt. This is unsurprising, as characterizing more generalized capacities may require greater abstraction. On the other hand, linguists have sometimes assumed that using psychological data requires a indiscriminate assimilation of any and all facts about behavior. In fact, cognitive psychologists are acutely aware of the difficulty of inferring cognitive structure from observed behavior.

Addressing these misunderstandings therefore serves not only to illuminate the connections between concepts in psychological and linguistic theories. It also enriches the empirical base for linguistic theories and the theoretical vocabulary of psychological theories.

With respect to the case study examined here—knowledge of phonotactic well-formedness—the evidence suggests that concepts from linguistic theories are part of the psychological capacities of individuals, deployed in on-line behavioral tasks. Establishing such correspondences has allowed us to draw on experimental data to inform theories of phonotactic knowledge. Although not directly acknowledged above, these experimental studies have drawn (and will continue to draw) heavily on theoretical concepts from linguistics. Addressing questions of psychological realism is therefore not simply an exercise in facilitating interdisciplinary relationships. It holds enormous practical importance for the development of more comprehensive theories of the structure, acquisition and use of language.

References

Abd-El-Jawad, Hassan R. and Issam M. Abu-Salim (1987). Slips of the tongue in Arabic and their theoretical implications. *Language Sciences* 9: 145-171.

Albright, Adam (2009). Feature-based generalisation as a source of gradient acceptability. *Phonology* 26: 9-41.

Anderson, Stephen (1985). *Phonology in the Twentieth Century*. Chicago: University of Chicago Press.

Baddeley, Alan (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience* 4: 829-839.

Bailey, Todd M. and Ulrike Hahn (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language* 44: 568-591.

Bard, Ellen Gurman, Dan Robertson and Antonella Sorace (1996). Magnitude estimation of linguistic acceptability. *Language* 72:32-68.

Béland, Renée and Carole Paradis (1997). Principled syllabic dissolution in a primary progressive aphasia case. *Aphasiology* 11: 1171-1196.

Berent, Iris, Daniel L. Everett and Joseph Shimron (2001). Do phonological representations specify variables? Evidence from the obligatory contour principle. *Cognitive Psychology* 42: 1-60.

Berent, Iris, and Joseph Shimron (2003). Co-occurrence restrictions on identical consonants in the Hebrew lexicon: Are they due to similarity? *Journal of Linguistics* 39: 31-55.

Berent, Iris, and Tracy Lennertz (2007). What we know about what we have never heard: Beyond Phonetics. Reply to Peperkamp. *Cognition* 104: 638-643.

Berent, Iris, Tracy Lennertz, Jongho Jun, Miguel A. Moreno, & Paul Smolensky (2008). Language universals in human brains. *PNAS* 105: 5321-5325.

Berent, Iris, Tracy Lennertz, Paul Smolensky, and Vered Vaknin-Nusbaum (2009). Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology* 26: 75-108.

Berent, Iris, Donca Steriade, Tracy Lennertz, and Vered Vaknin (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104: 591-630.

Bloomfield, Leonard (1933). *Language*. New York: Holt, Rinehart and Winston.

Blumstein, Sheila (1973). *A Phonological Investigation of Aphasic Speech*. The Hague: Mouton.

Boersma, Paul (1999). On the need for a separate perception grammar. Unpublished manuscript, University of Amsterdam. Rutgers Optimality Archive ROA–358, http://roa.rutgers.edu

Boersma, Paul and Bruce Hayes (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32, 45–86.

Bonte, Milene L., Holger Mitterer, Nadia Zellagui, Hane Poelmans, and Leo Blomert (2005). Auditory cortical tuning to statistical regularities in phonology. *Clinical Neurophysiology* 116: 2765-2774.

Brown, Roger W. and Donald C. Hildum (1956). Expectancy and the perception of syllables. *Language* 32: 411-419.

Buchwald, Adam B. (2009). Minimizing and optimizing structure in phonology: Evidence from aphasia. *Lingua* 119: 1380-1395.

Buchwald, Adam B., Brenda Rapp, & Maureen Stone (2007). Insertion of discrete phonological units: An articulatory and acoustic investigation of aphasic speech. *Language and Cognitive Processes* 22: 910-948.

Burzio, Luigi (1996). Surface constraints versus underlying representation. In Jacques Durand and Bernard Laks (eds.) *Current Trends in Phonology: Models and Methods* 97-122. Salford: University of Salford Publications.

Butterworth, Brian and Steven Whittaker (1980). Peggy Babcock's relatives. In George E. Stelmach and Jean Requin (Eds.) *Tutorials in Motor Behavior* 647-656. Amsterdam: North Holland.

Bybee, Joan (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

Caramazza, Alfonso (1986). On drawing inferences about the structural of normal cognitive systems from the analysis of patterns of impaired performance: The case for single-patient studies. *Brain and Cognition* 5: 41-66.

Chomsky, Noam (1980). *Rules and Representations*. New York: Columbia University Press.

Chomsky, Noam and Morris Halle (1968). *The Sound Pattern of English*. New York: Harper and Row.

Coetzee, Andries W. (2006). Variation as accessing 'non-optimal' candidates. *Phonology* 23: 337-385.

Coetzee, Andries W. (2008). Grammaticality and ungrammaticality in phonology. *Language* 84: 218-257.

Coetzee, Andries W. & Joe Pater (2008). Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory* 26: 289-337.

Coleman, John S. and Janet Pierrehumbert (1997). Stochastic phonological grammars and acceptability. In *Computational phonology: Third meeting of the ACL special interest group in computational phonology* 49-56. Somerset, NJ: Association for Computational Linguistics.

Culbertson, Jennifer and Steven Gross (in press). Are linguists better subjects? *British Journal for the Philosophy of Science*.

Cummins, Robert (1983). *The Nature of Psychological Explanation.* Cambridge, MA: MIT Press.

Dankovicová, Jana, Paula West, John Coleman, and Andrew Slater (1998). Phonotactic grammaticality is gradient. Poster presented at the 6th Laboratory Phonology Conference, University of York, York, UK. Retrieved June 1st, 2007, from http://www.phon.ox.ac.uk/~jcoleman/LabPhon-poster-paper.ps

Davidson, Lisa (2006a). Phonology, phonetics, or frequency: Influence on the production of non-native sequences. *Journal of Phonetics* 34: 104-137.

Davidson, Lisa (2006b). Phonotactics and articulatory coordination interact in phonology: Evidence from nonnative production. *Cognitive Science* 30:837-862.

Davidson, Lisa (2007). The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology* 24:261-286.

Day, Ruth S. (1968). *Fusion in Dichotic Listening.* Unpublished doctoral dissertation, Stanford University. Palo Alto, CA.

Day, Ruth S. (1976). Temporal order judgments in speech: Are individuals language-bound or stimulus-bound? *Status Report on Speech Research* 21/22: 71-87. New Haven, CT: Haskins Laboratories.

Dehaene-Lambertz, Ghislaine, Emmanuel Dupoux and Ariel Gout (2000). Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience* 12: 635-647.

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

Dell, Gary. S., Kristopher D. Reed, David R. Adams and Antje S. Meyer (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26: 1355–1367.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier, and Jacque Mehler (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25:1568-1578.

Dupoux, Emmanuel, Christophe Pallier, and Kazuhiko Kakehi (2001). New evidence for prelexical phonological processing in spoken word recognition. *Language and Cognitive Processes* 16: 491-505.

Fodor, Jerry A. (1981). Introduction: Some notes on what linguistics is about. In Ned Block (ed.) *Readings in philosophy of psychology, vol. 2* 197-202. Cambridge, MA: Harvard University Press.

Frisch, Stefan A., Nathan R. Large, and David B. Pisoni (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language* 42:481-496.

Frisch, Stefan A., Janet B. Pierrehumbert, and Michael B. Broe (2004). Similarity avoidance and the OCP. *Natural Language & Linguistic Theory* 22: 179-228.

Frisch, Stefan A., and Bushra Adnan Zawaydeh (2001). The psychological reality of OCP-place in Arabic. *Language* 77: 91-106.

Gaskell, M. Gareth, and William D. Marslen-Wilson (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes* 12: 613-656.

Gaskell, G. Gareth, Philip T. Quinlan, Jakke Tamminen, & Alexandra A. Cleland (2008). The nature of phoneme representations in spoken word recognition. *Journal of Experimental Psychology: General* 137: 282-302.

Gathercole, Susan E., Clive R. Frankish, Susan J. Pickering, and Sarah Peaker (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 25: 84-95.

Gathercole, Susan E. and Amanda J. Martin (1996). Interaction processes in phonological memory. In Susan E. Gathercole (ed.) *Models of Short Term Memory* 73-100. Hove, UK: Psychology Press.

Goldsmith, John A. (1995). Phonological theory. In John A. Goldsmith (ed.) *The Handbook of Phonological Theory* 1-23. Cambridge, MA: Blackwell.

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

Goldrick, Matthew and Robert Daland (2009). Linking speech errors and phonological grammars: Insights from Harmonic Grammar networks. *Phonology* 26: 147-185.

Goldrick, Matthew and Meredith Larson (2008). Phonotactic probability influences speech production. *Cognition* 107: 1155-1164.

Goldrick, Matthew and Meredith Larson (in press). Constraints on the acquisition of variation. In Cécile Fougeron and Mariapaola D'Imperio (Eds.) *Papers in Laboratory Phonology 10: Variation, Detail and Representation*. Berlin: Mouton de Gruyter.

Goldrick, Matthew, and Brenda Rapp (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition* 102: 219-260.

Goldstein, Louis, Marianne Pouplier, Larissa Chen, Elliot Saltzman, and Dani Byrd (2007). Dynamic action units slip in speech production errors. *Cognition* 103: 386-412.

Greenberg, Joseph H. and James J. Jenkins (1964). Studies in the psychological correlates of the sound system of American English. *Word* 20:157-177.

Hale, Mark and Charles Reiss (2000). Phonology as cognition. In Noel Burton-Roberts, Philip Carr and Gerard Docherty (eds.) *Phonological Knowledge: Conceptual and Empirical Issues* 161-184. Oxford: Oxford University Press.

Hallé, Pierre A., Juan Segui, Uli Frauenfelder and Christine Meunier (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24: 592-608.

Harris, Zellig S. (1951). *Methods in Structural Linguistics*. Chicago: University of Chicago Press.

Hayes, B. & C. Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39: 379-440. ROA 858.

Hickok, Gregory and David Poeppel (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences* 4: 131-138.

Jacquemot, Charlotte, Christophe Pallier, Denis LeBihan, Stanislas Dehaene, and Emmanuel Dupoux (2003). Phonological grammar shapes the auditory cortex: A functional magnetic resonance imaging study. *The Journal of Neuroscience* 23: 9541-9546.

Jacquemot, Charlotte, and Sophie K. Scott (2006). What is the relationship between phonological short-term memory and speech processing? *Trends in Cognitive Sciences* 10:480-486.

Jakobson, Roman (1941/1968). *Kindersprache, aphasie, and allgemeine lautgesetze*. Trans. Allan R. Keiler. The Hauge: Mouton.

Jenkins, James J. (1966). Meaningfulness and concepts: Concepts and meaningfulness. In Herbert J. Klausmeier and Chester W. Harris (eds.) *Analyses of Concept Learning* 65-79. New York: Academic Press.

Kabak, Baris, and William J. Idsardi (2007). Perceptual distortions in the adaptation of English consonant clusters: Syllable structure of consonantal contact constraints? *Language and Speech* 50: 23-52.

Kenstowicz, Michael (2001/2003). The role of perception in loanword phonology. *Linguistique Africaine* 20. Published in *Studies in African Linguistics* 32: 95-112.

Kirby, James and Alan C. L. Yu (2007) Lexical and phonotactic effects on wordlikeness judgements in Cantonese. In *Proceedings of the International Congress of the Phonetic Sciences XVI* 1161-1164,

Kupin, Joseph J. (1982). *Tongue Twisters as a Source of Information about Speech Production*. Bloomington: Indiana University Linguistics Club.

Laganaro, Marina and F.-Xavier Alario (2006). On the locus of the syllable frequency effect in speech production. *Journal of Memory and Language* 55: 178-196.

Lee, Yongeun (2006). *Sub-Syllabic Constituency in Korean and English*. Unpublished doctoral dissertation, Northwestern University. Evanston, IL.

Lee, Y., & Goldrick, M. (2008). The emergence of sub-syllabic representations. *Journal of Memory and Language* 59: 155-168.

Levelt, Willem J. M. (1974). *Formal Grammars in Linguistics and Psycholinguistics, Vol. III: Psycholinguistic Applications*. The Hauge: Mouton.

Levitt, Andrea G., and Alice F. Healy (1985). The roles of phoneme frequency, similarity, and availability in the experimental elicitation of speech errors. *Journal of Memory and Language* 24: 717-733.

Linell, Per (1979). *Psychological Reality in Phonology: A Theoretical Study*. Cambridge: Cambridge University Press.

Luce, Paul A., Stephen D. Goldinger, Edward T. Auer, Jr., and Michael S. Vitevitch (2000). Phonetic priming, neighborhood activation and PARSYN. *Perception & Psychophysics* 62: 615-625.

MacKay, Donald G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology* 3: 210-227.

Magnuson, James S., Bob McMurray, Michael K. Tanenhaus and Richard N. Aslin (2003a). Lexical effects on compensation for coarticulation: The ghost of Christmash past. *Cognitive Science* 27: 285-298.

Magnuson, James S., Bob McMurray, Michael K. Tanenhaus and Richard N. Aslin (2003b). Lexical effects on compensation for coarticulation: A tale of two systems? *Cognitive Science* 27: 801-805.

Majerus, Steve and Martial Van der Linden (2003). Long-term memory effects on verbal short-term memory: A replication study. *British Journal of Developmental Psychology* 21: 303-310.

Majerus, Steve, Martial Van der Linden, Ludivine Mulder, Thierry Meulemans, and Frédéric Peters (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language* 51: 297-306.

Marr, David (1982). *Vision*. San Francisco: W. H. Freeman and Company.

Martin, Nadine and Prahlad Gupta (2004). Exploring the relationship between word processing and verbal short-term memory: Evidence from associations and dissociations. *Cognitive Neuropsychology* 21: 213-228.

Martin, Nadine and Eleanor M. Saffran (2002). The relationship of input and output phonological processing: An evaluation of models and evidence to support them. *Aphasiology* 16:107-150.

Massaro, Dominic W. and Michael M. Cohen (1983). Phonological context in speech perception. *Perception & Psychophysics* 34:338-348.

McClelland, James L. and Jeffrey L. Elman (1986). The TRACE model of speech perception. *Cognitive Psychology* 18: 1-86.

McClelland, James L., Daniel Mirman, and Lori L. Holt (2006). Are there interactive processes in speech perception? *Trends in Cognitive Sciences* 10: 363-369.

McMillan, Corey T., Martin Corley, & Robin J. Lickley (2009). Articulatory evidence for feedback and competition in speech production. *Language and Cognitive Processes* 24: 44-66.

McQueen, James (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language* 39: 21–46.

McQueen, James (2003). The ghost of Christmas future: Didn't Scrooge learn to be good? Commentary on Magnuson, McMurray, Tanenhaus, and Aslin (2003). *Cognitive Science* 27: 795-799.

McQueen, James, Anne Cutler, and Dennis Norris (2006). Phonological abstraction in the mental lexicon. *Cognitive Science* 30: 1113-126.

Morais, José, Sao Luis Castro, Leonor Scliar-Cabral, Régine Kolinsky, and Alain Content (1987). The effects of literacy on the recognition of dichotic words. *The Quarterly Journal of Experimental Psychology Section A* 39: 451 – 465.

Moreton, Elliott (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84: 55-71.

Munson, Benjamin (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research* 44: 778-792.

Newman, Rochelle S., James R. Sawusch and Paul A. Luce (1999). Underspecification and phoneme frequency in speech perception. In Michael B. Broe and Janet B. Pierrehumbert (eds.) *Papers in Laboratory Phonology V: Acquistion and the Lexicon* 298-311. Cambridge: Cambridge University Press.

Newmeyer, Frederick (2003). Grammar is grammar and usage is usage. *Language* 79: 682–707.

Nimmo, Lisa M. and Steven Roodenrys (2002). Syllable frequency effects on phonological short-term memory tasks. *Applied Psycholinguistics* 23: 643-659.

Norris, Dennis, James M. McQueen and Anne Cutler (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences* 23: 299-370.

Ohala, John J. and Manjari Ohala (1986). Testing hypotheses regarding the psychological manifestation of morpheme structure constraints. In John J. Ohala and Jeri J. Jaeger (eds.) *Experimental Phonology* 239-252. San Diego, CA: Academic Press.

Ohala, Manjari (1983). *Aspects of Hindi phonology*. Delhi: Motilal Banarsidass.

Onishi, Kristine H., Kyle E. Chambers, and Cynthia Fisher (2002). Learning phonotactic constraints from brief auditory experience. *Cognition* 83: B13-B23.

Palmer, Stephen E. and Ruth Kimchi (1986). The information processing approach to cognition. In Terry J. Knapp and Lynn C. Roberts (eds.) *Approaches to Cognition: Contrasts and Controversies* 37-77. Hillsdale, NJ: Lawrence Erlbaum.

Peperkamp, Sharon (2007). Do we have innate knowledge about phonological markedness? Comments on Berent, Steriade, Lennertz, and Vaknin. *Cognition* 104: 631-637.

Perruchet, Pierre and Ronald Peereman (2004). The exploitation of distributional information in syllable processing. *Journal of Neurolinguistics* 17: 97-119.

Pierrehumbert, Janet B. (1994). Syllable structure and word structure: A study of triconsonantal clusters in English. In Patricia A. Keating (ed.) *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III* 168-188. Cambridge: Cambridge University Press.

Pierrehumbert, Janet B. (2001). Why phonological constraints are so coarse-grained. *Language and Cognitive Processes* 16: 691-698.

Pierrehumbert, Janet B. (2006). The next toolkit. *Journal of Phonetics* 34: 516-530.

Pitt, Mark A., and James M. McQueen (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language* 39: 347-370.

Pitt, Mark A., and Arthur G. Samuel (1995). Lexical and sublexical feedback in auditory word recognition. *Cognitive Psychology* 29: 149-188.

Pouplier, M. (2007). Tongue kinematics during utterances elicited with the SLIP technique. *Language and Speech* 50:311-341.

Pouplier, M. (2008). The role of a coda consonant as error trigger in repetition tasks. *Journal of Phonetics* 36: 114-140.

Prince, Alan and Paul Smolensky (1993/2002/2004). *Optimality Theory: Constraint interaction in generative grammar*. Technical report TR–2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ. Technical report CU–CS–696–93, Department of Computer Science, University of Colorado, Boulder. Revised version, 2002: Rutgers Optimality Archive ROA–537, http://roa.rutgers.edu/. Published 2004, Oxford: Blackwell.

Pylkkänen, Liina, Andrew Stringfellow, and Alec Marantz (2002). Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language* 81: 666-678.

Ratcliff, Roger, Pablo Gomez, and Gail McKoon (2004). A diffusion model account of the lexical decision task. *Psychological Review* 111: 159-182.

Roodenrys, Steven and Melina Hinton (2002). Sublexical or lexical effects on serial recall of nonwords? *Journal of Experimental Psychology: Learning, Memory and Cognition* 28: 29-33.

Romani, Cristina and Andrea Calabrese (1998). Syllabic constraints in the phonological errors of an aphasic patient. *Brain and Language* 64: 83-121.

Rose, Sharon and Lisa King (2007). Speech error elicitation and co-occurrence restrictions in two Ethiopian Semitic languages. *Language and Speech* 50:451-504.

Sapir, Edward (1933). La Réalité psychologique des phonèmes. *Journal de Psychologie Normale et Pathologuique* 30: 247-265. English translation reprinted in David G. Mandelbaum (ed.) (1949), *Selected Writings of Edward Sapir in Language, Culture and Personality* 46-60. Berkeley, CA: University of California Press.

Saussure, Ferdinand de (1916/1959). *Course in General Linguistics*. trans. Wade Baskin. New York: McGraw-Hill.

Scholes, Robert J. (1966). *Phonotactic Grammaticality*. The Hauge: Mouton.

Schütze, Carson T. (1996). *The Empirical Base of Linguistics*. Chicago: University of Chicago Press.

Schweickert, Richard (1993). A multinomial processing tree model of degradation and redintegration in immediate recall. *Memory and Cognition* 21: 168-175.

Segui, Juan, Ulricht Frauenfelder and Pierre Hallé (2001). Phonotactic constraints shape speech perception: Implications for sublexical and lexical processing. In Emmanuel Dupoux (ed.) *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler* 195-208. Cambridge, MA: MIT Press.

Shademan, Shabnam (2006). Is phonotactic knowledge grammatical knowledge? In Donald Baumer, David Montero, and Michael Scanlon (eds.) *Proceedings of the 25th West Coast Conference on Formal Linguistics* 371-379. Somerville, MA: Cascadilla Press.

Shademan, Shabnam (2007). *Grammar and Analogy in Phonotactic Well-Formedness*. Unpublished doctoral dissertation, University of California, Los Angeles. Los Angeles, CA.

Shattuck-Hufnagel, Stephanie, and Dennis H. Klatt (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior* 18: 41-55.

Smolensky, Paul (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27:720-731.

Smolensky, Paul (2006). Computational levels and integrated connectionist/symbolic explanation. In Paul Smolensky and Géraldine Legendre *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar (Vol. 2, Linguistic and Philosophical Implications)* 503-592. Cambridge, MA: MIT Press.

Smolensky, Paul, Geraldine Legendre, & Bruce Tesar (2006). Optimality theory: The structure, use, and acquisition of grammatical knowledge. In Paul Smolensky and Géraldine Legendre *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar (Vol. 1: Cognitive architecture)* 453-544. Cambridge, MA: MIT Press.

Söderpalm, Ewa (1979). *Speech errors in normal and pathological speech* (Travaux de l'Institute de Linguistique de Lund 14). Lund: CWK Gleerup.

Stemberger, Joseph Paul (2004). Neighbourhood effects on error rates in speech production. *Brain and Language* 90: 413-422.

Thorn, Annabel S. C. and Clive R. Frankish (2005). Long-term knowledge effects on serial recall of nonwords are not exclusively lexical. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31: 729-735.

Treiman, Rebecca, Brett Kessler, Stephanie Knewasser, Ruth Tincoff and Margo Bowman (2000). English speakers' sensitivity to phonotactic patterns. In Michael B. Broe and

Janet B. Pierrehumbert (eds.) *Papers in Laboratory Phonology V: Acquistion and the Lexicon* 269-282. Cambridge: Cambridge University Press.

Trubetzkoy, Nikolai S. (1939/1969). *Grundzüge der Phonologie*. Trans. Christine A. M. Baltaxe. Berkeley, CA: University of California Press,

van Gelder, Timothy J. (1998) The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences* 21: 1-14.

Vallar, Giuseppe (2006). Memory systems: The case of phonological short-term memory. A festschrift for *Cognitive Neuropsychology*. *Cognitive Neuropsychology* 23: 135-155.

Van der Lugt, Arie H. (2001). The use of sequential probabilities in the segmentation of speech. *Perception & Psychophysics* 63: 811-823.

Vickers, Douglas and Michael D. Lee (1998). Dynamic models of simple judgments: I. Properties of a self-regulating accumulator module. *Nonlinear Dynamics, Psychology, and Life Sciences* 2: 169-194.

Vitevitch, Michael S. (2003). The influence of sublexical and lexical representations in the processing of spoken words in English. *Clinical Linguistics & Phonetics* 17: 487-499.

Vitevitch, Michael S., Jonna Armbrüster and Shinying Chu (2004). Sublexical and lexical representations in speech production: Effects of phonotactic probability and onset density. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30: 514-529.

Vitevitch, Michael S., Paul A. Luce (1999). Probabilistic phonotactics and neighborhood density in spoken word recognition. *Journal of Memory and Language* 40: 374-408.

Vitevitch, Michael S., Paul A. Luce, Jan Charles-Luce, and David Kemmerer (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech* 40:47-62.

Vitevitch, Michael S., David B. Pisoni, Karen Iler Kirk, Marcia Hay-Cutcheon and Stacey L. Yount (2002). Effects of phonotactic probabilities on the processing of spoken words and nonwrds by adults with cochlear implants who were postlingually deafened. *The Volta Review* 102: 283-302.

Vousden, Janet I., Gordon D. A. Brown, and Trevor A. Harley (2000). Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology* 41: 101-175.

Wan, I-Ping, and Jeri Jaeger (1998). Speech errors and the representation of tone in Mandarin Chinese. *Phonology* 15: 417-461.

Warker, Jill A. and Gary S. Dell (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory and Cognition* 32: 387-398.

Wheeler, Deirdre W. and David S. Touretzky (1997). A parallel licensing model of normal slips and phonemic paraphasias. *Brain and Language* 59: 147-201.

Whorf, Benjamin Lee (1940). Linguistics as an exact science. *Technology Review* 43: 61-63, 80-83. Reprinted in John B. Carroll (ed.) (1956) *Language, Thought, and Reality* 220-232. Cambridge, MA: Technology Press.

Zimmer, Karl E. (1969). Psychological correlates of some Turkish morpheme structure conditions. *Language* 45: 309-321.

Zuraw, Kie (2007). The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language* 83: 277-316.