Published in final edited form as:

Linguist Vanguard. 2015 December 1; 1(1): 215–225. doi:10.1515/lingvan-2015-1003.

The Influence of Word Retrieval and Planning on Phonetic Variation: Implications for Exemplar Models

Angela Fink and

Department of Linguistics, Northwestern University

Matthew Goldrick

Department of Linguistics, Northwestern University

Abstract

Over the past several decades, an increasing number of empirical studies have documented the interaction of information across the traditional linguistic modules of phonetics, phonology, and lexicon. For example, the frequency with which a word occurs influences its phonetic properties of its sounds; high frequency words tend to be reduced relative to low frequency words. Exemplar Models have been successful in accounting for this body of results through a single mechanism, *exemplars*—memory representations that integrate lexical, phonological, and phonetic information into a single structure. We review recent studies that suggest there are critical limitations to assuming that phonetic variation solely reflects the storage of word labels and sound structure in exemplars. Specifically, these studies show that factors related to the on-line retrieval and planning of lexical items also influence phonetic variation. The implications of these findings for exemplar models are discussed; the relationship of exemplar storage to the broader cognitive system is examined, as well as alternative theoretical frameworks incorporating gradience at all levels of linguistic representation.

Keywords

exemplars; speech production; reduction; neighborhood density; accent; speech errors; gradient representations

1 Introduction

Modularity—the discrete separation of different levels of language representation and processing—was a core principle of foundational proposals in both generative theories (Chomsky & Halle, 1968) as well as psycholinguistic theories of speech production (Garrett, 1975). An increasingly large body of empirical results has challenged this assumption, documenting interactive effects—empirical patterns that reflect the simultaneous influence of multiple levels of representation and processing. For example, as reviewed by Gahl (2008; see also Bell, Gregory, Girand, & Jurafsky, 2009), a wide array of data suggests that lexical frequency (a word-level property) influences the phonetic properties of sounds within

words. Sounds are more likely to be reduced in high vs. low frequency words (see Sadat, Martin, Alario, & Costa, 2011, for related evidence from bilingual speakers). Under a strictly modular approach, this should not occur. The phonetic (articulatory/acoustic) properties of a speech sound should be determined solely by the phonological and phonetic context in which the sound occurs; they should not be influenced by lexical properties of the word in which they occur.

Observations such as these have fueled the development of exemplar models of speech production (e.g., Goldinger, 1998; Kirchner, Moore, & Chen, 2010; Pierrehumbert, 2002, 2006; Walsh, Möbius, Wade, & Schütze, 2010; Wedel, 2007; see Pisoni & Levi, 2007, for discussion of similar developments in speech perception and word recognition). In such models, speakers store a large set of rich memory representations, each of which encodes several dimensions of linguistic structure. These representations specify not only detailed information about the phonological and phonetic structure of the intended word but also information about the word itself. We refer to this type of proposal as a *Lexicalist Exemplar Model* to emphasize the inclusion of word-specific information¹ along with sound structure.

By storing word-specific information alongside phonetic/phonological structure, a Lexicalist Model provides a natural account of interactive effects like those reviewed by Gahl (2008). For example, speakers encounter many examples of high frequency words that are reduced, but do not encounter such examples for low frequency words. Storage of these phonetic details for each word influences subsequent productions, causing speakers to be more likely to produce reduced high vs. low frequency words.

The success of this approach naturally suggests a strong null hypothesis: storage of exemplars, specifying word labels and sound structure, provides a full account of the range of phonetic variation. In other words, if speech production can be reduced to the retrieval of finely detailed, lexicalist exemplar representations and their faithful motor execution, then phonetic variation is predicted to derive only from the linguistic structures stored in memory. We review five sets of findings that challenge this strong claim, suggesting other aspects of language production processing also affect phonetic variation. We then consider several areas of active theoretical development that may offer possible solutions to these challenges. These include: incorporating a wider array of information into Lexicalist exemplars; articulating how the simple Lexicalist Model might interface with other cognitive systems; and/or adopting more dynamic models where ease of processing can directly affect the representations generated during production.

2 Challenges for the Lexicalist Exemplar Model

2.1 Multiple sources of variation in reduction

When produced in contexts in which they are more predictable, words have less extreme phonetic properties (e.g., shorter word durations, smaller vowel space sizes) relative to contexts in which they are unpredictable (see Bell et al, 2009, for a review). These patterns

¹Although not discussed here, many such models also include social/indexical information in exemplars (see Pierrehumbert, 2006, for discussion).

go beyond the context-independent effects of lexical properties such as word frequency. Indeed, the same word may have different attributes depending on its context (e.g., *sugar* will tend to be reduced in “I prefer coffee with both cream and ____.” compared to “I went to the store and bought ____.”).

The Lexicalist Model does not account for contextual variation in reduction; it assumes that exemplars contain no information above the lexical level. The model might accommodate such evidence by enriching exemplars to include some notion of context (see the section on *Exemplar expansion* below). However, recent studies suggest that incorporating informational context into lexicalist exemplars will not account for the full range of reduction patterns. Ernestus (2014) provides an extensive review and discussion of such evidence; here, we briefly discuss two additional sets of studies.

Kahn and Arnold (2012; see Kahn & Arnold, 2013, for related work) examined phonetic variation in a communicative production task. Speakers watched objects move on a computer screen and then named them while describing the scene to an interlocutor. Some of these objects had been presented visually on previous trials, and were thus given (with respect to the discourse context) for both speaker and interlocutor. These object names were expected to be reduced based on this information status (Fowler & Housum, 1987). Critically, a subset of these given names had also been produced by the speaker. If selection of context-appropriate exemplars was the sole source of reduction, names from this previously-produced subset should be no different than any other given item; all of these names have the same status relative to the discourse. This prediction was not borne out: Kahn and Arnold found increased reduction for previously named items.

Results from Lam and Watson (2014) provide further evidence that mere repetition (independent of discourse status) can give rise to reduction. Participants described events presented on a computer screen. These involved actions by characters (e.g., man 1) that had certain occupations (e.g., chef). Since the same occupation could be occupied by multiple characters (e.g., man 1 and man 2 could both be chefs), participants could repeat the same word without repeating the same referent (e.g., “The chef (man 1) is leaving job A. The *chef* (man 2) is leaving job B.”). The repeated production (*chef*) was compared to the same word in a non-repeated context (e.g., “The cop (man 1) is leaving job A. The chef (man 2) is leaving job B.”) Lexical repetition in the absence of referent repetition caused reduction in both intensity and word durations, suggesting that reduction does not simply reflect linguistic factors or the informational context in which a word is produced.

2.2 Task-driven variation in the effects of neighborhood density

Turning from effects of informational context and word retrieval on phonetic outcomes, we consider evidence that task parameters also influence the realization of word-specific properties. The seminal work of Wright (2004) and Scarborough (2004) suggested that neighborhood density—the number of lexical items phonologically similar to a target form—influences phonetic variation. Wright found that vowels in words with many neighbors are realized with more extreme acoustic properties than those with few neighbors (i.e., vowels in high density words are more dispersed in first/second formant space). Scarborough found enhanced anticipatory nasalization for vowels in words with many vs.

few neighbors. Such results have been replicated in a number of laboratory studies (e.g., Munson & Solomon, 2004; see Scarborough & Zellou, 2013, for a review); furthermore, (Baese-Berk and Goldrick 2009) extended Wright's observation of enhanced phonetic properties for vowels to word-initial consonants (but see Goldrick, Vaughn, & Murphy, 2013, for discussion of different patterns for word-final consonants).

If neighborhood density is a word-specific property stored in lexicalist exemplars, then we would expect its effects on phonetic variation to remain stable across different processing contexts. Some results are consistent with this prediction. Scarborough (2010) finds that neighborhood density effects are the same when participants read words in predictable vs. unpredictable sentence frames, and Scarborough and Zellou (2013) find no differences across clear vs. conversational speech. However, other results suggest neighborhood effects might vary across different speech production tasks. Gahl, Yao, and Johnson (2012) analyzed spontaneous speech and found that increasing neighborhood density leads to reduction rather than hyperarticulation, suggesting the effects of neighbors on phonetic variation might vary across read vs. spontaneous speech.

While the empirical picture is not yet clear, there is some evidence that the effect of lexical neighbors on phonetic processing varies across tasks. This contradicts the Lexicalist Model's prediction that word-level effects should remain constant, given that only lexical and phonetic/phonological information are stored in exemplars.

2.3 Phonetic consequences of disruptions to lexical access

Next, we consider two data sets that strongly challenge the idea that all phonetic variation can be captured through lexicalist exemplar storage. In each case, phonetic effects reflect unexpected and unintended disruptions to cognitive processing. Kello, Plaut, and MacWhinney (2000) used the Stroop interference paradigm (see MacLeod, 1991, for a review) to disrupt speech planning and examined the conditions under which such disruptions would influence the phonetic properties of speech. They presented participants with written color words and asked them to name the color of the text (e.g., if RED is written in blue text, say "blue"). Classic Stroop interference was observed in response times and error rates, with slower responses and more errors emerging on incongruent trials (e.g., RED written in blue text) than congruent ones (e.g., RED written in red text). When participants were forced to respond quickly, Stroop disruptions to speech planning also interfered with phonetic processing (producing longer word durations on incongruent trials). Consistent with these results, Kello (2004) finds that under increasing time pressure word durations show a greater influence of word-level variables that influence retrieval and planning.

However, Damian (2003) failed to replicate these effects. He not only performed a direct replication of Kello et al. (2000), but also manipulated time pressure during two other tasks that disrupt lexical access (picture-word interference with semantically related distractor words; semantically blocked picture naming). Across the board, interference during lexical access yielded slower reaction times (RTs) and more errors, but it had no effect on word durations. It is unclear what gave rise to the disagreements across studies by Damian vs. Kello and colleagues. Given that the effects are relatively small (in Kello et al., less than 40

msec differences across contexts, with total word durations exceeding 300 msec), it may be difficult to consistently observe them. This is especially likely if there is substantial variation across speakers in sensitivity to disruptions to lexical access.

Clearly, more work is needed to clarify when disrupted lexical access does or does not have phonetic consequences. What we emphasize here is that the observation of such effects under some circumstances requires further explanation. The Lexicalist Model's appeal to storage of linguistic structure as the source of phonetic variation does not provide a ready account of the effects of processing disruptions.

2.4 Phonetic consequences of language switching

The phenomenon of language switching provides an ideal testing ground for examining the phonetic effects of on-line disruptions to word retrieval and planning. The behavioral cost of alternating languages has been well-documented. This is often studied in laboratory settings using the cued language switching paradigm, where language cues (e.g., colored squares or national flags) prompt bilinguals to alternate naming targets (e.g., digits or pictures) in their L1 and L2 (Meuter & Allport, 1999). Parallel to other domains (Monsell, 2003), comparison of RTs and error rates across these trial types reliably shows a *switch cost*, such that trials where the response language is different from the previous trial have longer RTs and higher error rates than trials where the language does not switch (see Bobb & Wodniecka, 2013, for a recent review). Similar effects are found when participants voluntarily switch languages (e.g., Gollan & Ferreira, 2009), suggesting this processing cost is pervasive.

Recent evidence from Spanish-English bilinguals indicates that on-line difficulty at the point of switching languages can disrupt subsequent articulatory processing². Voice-onset time (VOT), an important acoustic cue to the distinction between voiced and voiceless stops, is utilized in different ways in Spanish vs. English (Spanish contrasts short lag VOT vs. prevoiced stops, while English contrasts long lag vs. short lag VOT). The conflicting realizations of this contrast are reflected in the accented productions of non-native speakers. For example, in English, native Spanish speakers tend to produce shorter and more prevoiced VOTs than monolingual English speakers (Flege, 1991). Using cued language switching, Olson (2013) demonstrated more accented speech on voiceless stops during switch trials compared to stay trials. This effect emerged only during production of the dominant language (Spanish). Using a similar paradigm, Goldrick, Runnqvist, and Costa (2014) reported an effect of language switching on both voiceless and voiced productions in Spanish-English bilinguals' non-dominant language (English) but not in their dominant language. Finally, in a study of voiceless stop production in spontaneous speech, Balukas and Koops (in press) also found effects of switching in the non-dominant language. Like the monolingual data presented in the previous section, these bilingual data necessitate revision or elaboration of the Lexicalist Model, so that difficulties in retrieving and planning words can influence phonetic processing.

²It should be noted that other paradigms (e.g., reading aloud sentences containing a code switch) have yielded inconsistent results (for review and discussion, see Balukas & Koops, in press).

2.5 Phonetic effects in speech errors

In this final section, we turn to a source of phonetic variation that is truly unlikely to emerge from stored information: phonological speech errors during novel tongue twisters. When uttering a tongue twister sequence such as *pin bin bin pin*, speakers commonly produce onset errors like *pin* → *bin*. A number of studies have shown such errors reflect a blend of properties of the intended (/p/) and intruding error (/b/) sounds (e.g., Frisch & Wright, 2002; Goldrick & Blumstein, 2006; Goldstein, Pouplier, Chen, Saltzman, & Byrd, 2007; McMillan, Corley, & Lickley, 2009). For example, a /b/ produced in error (*pin*→*bin*) has a longer VOT (i.e., a more /p/-like VOT) than a correctly produced /b/ (*bin*→*bin*; Goldrick, Baker, Murphy, & Baese-Berk, 2011).

These phonetic blends reflect, in part, an interaction between lexical and phonetic levels of representation. Goldrick et al. (2011) examined phonetic blends for high and low frequency word targets and error outcomes, finding that low frequency words tend to exert more of an influence on the phonetic properties of errors than high frequency words. They attribute this pattern to the more variable realization of high vs. low frequency words (as noted in the introduction, high frequency words are often phonetically reduced). The association of low frequency words with a more precise set of phonetic targets allows them to exert a stronger influence on these blends than high frequency words. The Lexicalist Exemplar Model naturally captures this pattern of results. If errors reflect a blend of exemplars from the target word and the error word, we would predict that the different properties of exemplars associated with high vs. low frequency words would influence the phonetic properties of the resulting productions.

However, a second finding is problematic for this Lexicalist account. If the blended phonetic properties of speech errors are attributed solely to the activation of word-specific exemplars, we would predict that such blends would be limited (or absent) when producing strings that lack such exemplars—i.e., non-words like *geff*. However, phonetic blends are often observed—in fact, they are magnified—when speakers produce non-word twister sequences. Goldrick and Blumstein (2006) and McMillan et al. (2009) report that phonetic properties of the target sound exert more of an influence on nonword vs. word error outcomes (e.g., the [g] in *keff*→*geff* has more evidence of /k/ than the [g] in *keese*→*geese*). There appears to be no simple storage-based explanation for these data, because these blended productions arise during twisters composed of *non-words*, which by definition are not associated with a specific stored lexical representation

3 Exemplar expansion and its limitations

The Lexicalist Exemplar Model makes a strong prediction that word labels and sound structure are the only type of information that influence phonetic variation. This prediction arises from the assumptions that exemplar representations integrate dimensions of linguistic structure only (i.e., no broader context is included), and that storage of exemplars provides a full and complete account of phonetic variation. However, the data reviewed above demonstrate that factors related to word retrieval and planning (mere repetition, speech task, and disruptions to retrieval/planning in mono- and multilingual speech) also influence

phonetic variation. This section explores the idea of capturing these data by augmenting lexicalist exemplars to include other types of information relevant to language processing.

This solution most straightforwardly captures the finding that phonetic reduction effects vary depending on informational context (recall that *sugar* will tend to be reduced in “I prefer coffee with both cream and ____.” compared to “I went to the store and bought ____.”). If speakers index the informational context in which an exemplar occurs, they can select the most context-appropriate exemplars to determine the phonetic properties of their productions. However, inclusion of simple informational context in exemplars does not explain why repetition reduction occurs even in the absence of referent repetition (Kahn & Arnold 2012; Lam & Watson, 2014). It remains possible that a more precise formulation of the contextual information stored in exemplars could capture these data. For example, what if part of the context stored in an exemplar refers to the recency with which a phonological form was produced? The primary problem with this account is the lack of justification for storing such information. If speakers are already maintaining contextual information about the message they wish to convey, it is unclear what principle would predict that they also track information about how recently a specific form was produced.

Similar concerns arise when we consider expanding exemplars to include other types of contextual information. Following the proposal above, the Lexicalist Model might accommodate task-dependent phonetic variation (Gahl et al., 2012) by enriching exemplars to index whether an exemplar was produced during a particular type of task. However, in contrast to informational context (which reflects the speaker’s intended message), it is unclear why exemplars would incorporate the task in which a speaker produced an utterance.

We might also attempt to explain how disruptions to monolingual and bilingual lexical access cause phonetic disruptions (Balukas & Koops, 2014; Goldrick et al., 2014; Kello et al., 2000; Kello, 2004; Olson, 2013) by storing processing difficulty as a part of context. When monolingual speakers encounter high processing demands, they could select exemplars labeled for “difficult/dysfluent” processing conditions. The resulting increase in durations could allow more time for processing to be completed (Bell et al., 2009; Kello et al., 2000), or it might signal to the listener that the speaker is experiencing problems. When bilinguals are language switching, they could draw on exemplars utilized in mixed language contexts, which are known to be intermediate between the bilingual’s languages (e.g., Antoniou, Best, Tyler, & Kroos, 2011). However, these accounts seem rather contradictory: strategic selection of context-appropriate exemplars would require cognitive control, while disruptions to lexical access entail the loss of such control. Instead of speakers actively choosing representations that are specific to difficult processing contexts, it seems more likely that phonetic disruptions are an uncontrolled by-product of processing difficulty.

These issues suggest that the challenge currently facing the field is to acknowledge the Lexicalist Model’s success in accommodating the interactive effects that originally motivated it, while also recognizing that the data reviewed here may require a solution that goes beyond simply expanding what information is stored in exemplars.

4 Exemplar Processing

Instead of expanding the information that is stored in exemplars, we might improve the Lexicalist Exemplar Model by better articulating the nature of exemplar processing. For example, more detailed specification of how exemplars are accessed from memory might offer an account of effects reflecting word retrieval and planning. Clopper and Pierrehumbert (2008) outline a mechanism in which exemplars associated with dialect-specific variants take more time to become activated than standard variants. This predicts dynamic, context-specific variation in the expression of dialect features without forcing exemplars to explicitly label easy vs. difficult processing contexts. For example, given that speech planning is easier in more vs. less predictable sentences, more predictable sentences will result in the retrieval of a greater number of exemplars specifying dialect-specific variants. As Clopper and Pierrehumbert observe, this predicts greater expression of dialect features in more vs. less predictable sentence contexts.

The Lexicalist Exemplar Model could also be enriched by specifying how exemplar processing interfaces with the larger cognitive processing system. For instance, several exemplar models have proposed that the phonetic form produced by a speaker reflects not only the exemplars retrieved from memory but also general phonetic implementation processes (e.g., gestural reduction processes sensitive to the sound structure context; e.g., Pierrehumbert, 2002; Wedel, 2007). Effects of mere repetition could reflect the role of such phonetic processes, assuming a reduction process that is sensitive to the recency with which a given form was produced. However, as noted by Ernestus' (2014) review, the full range of reduction processes have not been incorporated into existing models. Phonetic implementation might also account for blends observed in phonological speech errors during non-word tongue twisters, which could arise through blends of the speech motor plans in phonetic processing (e.g., Goldstein et al., 2007). Specifying the mechanism(s) underlying possible interactions between exemplars and phonetic processes would be critical here, as the influence of lexical properties on speech errors (e.g., Goldrick et al., 2011) suggests that not all blended productions arise purely within phonetic encoding.

Articulating how exemplar storage interfaces with cognitive control systems might also provide an account of phonetic variation that emerges during disrupted lexical access. Cognitive control refers to processes that distribute attention and cognitive resources depending on the task(s) at hand (Baddeley & Hitch, 1974; Norman & Shallice, 1986). Such processes are likely in high demand in speeded paradigms where production processes must be tightly coordinated, as well as in difficult-to-process experimental conditions where dominant responses must be suppressed in favor of the target. Assuming that exemplar processing and phonetic processing draw upon a common, limited pool of control resources, taxing control during exemplar retrieval might deprive phonetic processing of the resources it requires. This could lead to more dysfluent phonetic outcomes during disrupted lexical access.

Bilingual language production might present a particular challenge to control processes. Several theories propose that bilingual speakers must engage some form of cognitive control in order to select only one language for production (e.g., Green, 1998). If the common set of

resources for control is allocated (at least in part) towards accessing exemplars in the target language, fewer resources will be available for phonetic processing. The resulting difficulties might manifest as inappropriate retrieval of well-practiced native language phonetic representations, enhancing accents.

5 Gradient Representations

Rather than relying solely on stored exemplars to produce interactive effects, an alternative (but mutually compatible) mechanism assumes that variation in lexical access influences the nature and structure of representations retrieved from memory. While many theories have assumed that phonetic representations specify continuous articulatory and/or acoustic properties of word forms, the proposals reviewed here claim that gradience extends to phonological and lexical representations as well. If the gradient aspects of these representations can be influenced by the structure and efficiency of lexical access processes, changes to lexical access will result in distinct representational structures as output. These changes to the input to phonetic processes could produce interactive effects.

Consider effects related to lexical neighbors (reviewed above). Several studies have suggested that vowels in words with many vs. few neighbors are realized with more extreme acoustic properties. How could this basic interactive effect (which helped motivate the Lexicalist Exemplar Model) arise under a gradient representation account? If, during lexical access, representations of lexical neighbors become active, they could serve to reinforce the activation of the phonological representation of the target word's sounds. This would produce a (gradient) difference between the phonological representation of words with many vs. few neighbors; the former representations would be more active. Transmitting this representational difference to phonetic processes would result in different realizations of these two types of words (Baese-Berk & Goldrick, 2009).

Assuming such accounts can capture the full range of interactive effects modeled by the basic Lexicalist Exemplar Model, can they also account for the data reviewed here? Like the Lexicalist Model, such proposals must clearly articulate how phonetic processes are incorporated into the production system to account for the dynamics of phonetic variation. Below, we consider how two gradient representational proposals might be extended to model effects related to word retrieval and planning.

Articulatory Phonology (AP; see Browman & Goldstein, 1992, for an overview) is a long-standing framework that incorporates gradience. Phonological representations are hypothesized to consist of coordinated sets of *gestures*, abstract articulatory goals. AP representations incorporate gradient specifications of coordination (e.g., relative phase at which gestures are coordinated) as well as aspects of gestures themselves (e.g., gestural activation, the degree to which a gesture contributes to the overall articulatory plan). Recent work (e.g., Goldstein, Byrd, & Saltzman, 2006; Kirov & Gafos, 2007; Tilsen, 2011, 2013) has proposed that continuously activated, dynamic planning representations underlie the retrieval and planning of these gestural representations. These provide a natural link between lexical retrieval/planning and phonetic processes. If the gradient properties of these planning representations are (in part) determined by the ease vs. difficulty of lexical access, variation

in these planning representations can produce changes in the phonetic properties of utterances. This framework may offer a means to account for effects related to word retrieval and planning.

Connectionist processing mechanisms (Rumelhart, Hinton, & McClelland, 1986) also allow for dynamic specification of representational structure. In such accounts, mental representations are realized through distributed patterns of activation among simple processing units. Such distributed, quantitative patterns of activation need not correspond to a single representation, meaning that the system can easily represent gradient blends of multiple representational states (e.g., [k]: 0.9, [g]: 0.1). Critically, because the content of such representations reflects the spreading of activation, the degree to which different representations are present in such blends is dynamically specified³—providing another possible mechanism to account for the interactive effects reviewed above.

For example, Goldrick and Chu (2014) developed an account of phonetic blends in speech errors using the Gradient Symbolic Computation (GSC) framework (Smolensky, Goldrick, & Mathis, 2014). GSC assumes that phonological representations are abstract, symbolic representations (as in more traditional generative frameworks). To ensure that connectionist representations respect symbolic structure, GSC incorporates a mechanism (quantization) that pulls the system away from blends, towards pure symbolic states. However, this mechanism does not impose an inviolable constraint on processing. Goldrick and Chu provided simulation results showing disrupting processing can modulate the degree of blending between competing representations. Specifically, when processing resources are taxed—as in conditions producing speech errors—quantization does not have sufficient time to pull the system away from blend states. This allows the partially activated target to exert a significant influence on error processing, producing a representation in which target and error are co-activated. Phonetic implementation of this blended state results in the blended productions observed in speech errors. Dynamic connectionist processes—where the structure of representations can be influenced by variation in the ease vs. difficulty of lexical access—thus provide another possible mechanism to account for the data reviewed above.

6 Challenges and prospects

While a range of data suggests that the strong modularity assumptions of traditional generative grammars and psycholinguistic theories are inadequate, the structure of the appropriate alternative account is still far from clear. This partially reflects uncertainty regarding the empirical status of a variety of interactive effects. For example, studies using different tasks have reported the opposite effect of lexical neighborhood density; in monolinguals, effects of disruptions to lexical access on articulation have not been replicated. Another contributing factor is the imprecision of current theoretical proposals. There are several frameworks that allow for the possibility of interactive effects. However, the strong predictions of specific accounts within each framework have not been adequately explored (e.g., how much interaction is predicted to be present under particular processing

³While this discussion focuses on speech production, there has been extensive work on dynamically specified connectionist representations in perception and memory, including the influential TRACE model (McClelland & Elman, 1986) and the extensive body of work in Adaptive Resonance Theory (see, Grossberg, 2003, for a review).

circumstances?). The rich quantitative and dynamic nature of the empirical patterns reviewed above require us to move beyond very general theoretical statements to develop more specific, testable proposals.

In spite of these challenges, there are many reasons to be optimistic regarding the prospects for progress in understanding interactive effects. These empirical issues have attracted a growing number of researchers at the intersection of laboratory phonology and psycholinguistics. This increased attention is likely to produce a much richer body of data for theoretical development. Similarly, the range of theoretical proposals reflects the depth and diversity of researchers interested in this topic; this is likely to fuel the development of more precisely specified proposals that can make strong, testable predictions.

Acknowledgments

Preparation of this manuscript was supported by NSF Grants BCS-1344269, BCS-1420820, and NIH-NICHD grant HD077140. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF or the NIH. Thanks to the Northwestern SoundLab, Abby Cohn, and a reviewer for helpful discussion and comments.

References

- Antoniou M, Best CT, Tyler MD, Kroos C. Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics*. 2011; 39(4): 558–570. [PubMed: 22787285]
- Baese-Berk M, Goldrick M. Mechanisms of interaction in speech production. *Language and Cognitive Processes*. 2009; 24(4):527–554. [PubMed: 19946622]
- Baddeley, AD.; Hitch, GJ. Working memory. In: Bower, GH., editor. *The psychology of learning and motivation*. Vol. 8. New York: Academic Press; 1974. p. 47-89.
- Balukas C, Koops C. Spanish-English bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism*. (in press).
- Bell A, Brenier JM, Gregory M, Girand C, Jurafsky D. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*. 2009; 60(1):92–111.
- Bobb SC, Wodniecka Z. Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. *Journal of Cognitive Psychology*. 2013; 25(5):568–585.
- Browman CP, Goldstein L. Articulatory phonology: An overview. *Phonetica*. 1992; 49(3–4):155–180. [PubMed: 1488456]
- Chomsky, N.; Halle, M. *The sound pattern of English*. New York: Harper and Row; 1968.
- Clopper CG, Pierrehumbert JB. Effects of semantic predictability and regional dialect on vowel space reduction. *The Journal of the Acoustical Society of America*. 2008; 124(3):1682–1688. [PubMed: 19045658]
- Damian MF. Articulatory duration in single-word speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2003; 29(3):416–431.
- Ernestus M. Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua*. 2014; 142:27–41.
- Flege JE. Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America*. 1991; 89(1): 395–411. [PubMed: 2002177]
- Fowler CA, Housum J. Talkers' signaling of “new” and “old” words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*. 1987; 26(5):489–504.

- Frisch SA, Wright R. The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. *Journal of Phonetics*. 2002; 30(2):139–162.
- Gahl S. Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*. 2008; 84(3):474–496.
- Gahl S, Yao Y, Johnson K. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of memory and language*. 2012; 66(4):789–806.
- Garrett, MF. The analysis of sentence production. In: Bower, G., editor. *Psychology of learning and motivation*. Vol. 9. New York: Academic Press; 1975. p. 133–175.
- Goldinger SD. Echoes of echoes? An episodic theory of lexical access. *Psychological review*. 1998; 105(2):251. [PubMed: 9577239]
- Goldrick M, Blumstein SE. Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes*. 2006; 21(6):649–683.
- Goldrick M, Baker HR, Murphy A, Baese-Berk M. Interaction and representational integration: Evidence from speech errors. *Cognition*. 2011; 121(1):58–72. [PubMed: 21669409]
- Goldrick M, Chu K. Gradient co-activation and speech error articulation: comment on Pouplier and Goldstein (2010). *Language, Cognition, and Neuroscience*. 2014; 29(4):452–458.
- Goldrick M, Runnqvist E, Costa A. Language Switching Makes Pronunciation Less Nativelike. *Psychological science*. 2014; 25(4):1031–1036. [PubMed: 24503870]
- Goldrick M, Vaughn C, Murphy A. The effects of lexical neighbors on stop consonant articulation. *The Journal of the Acoustical Society of America*. 2013; 134(2):EL172–EL177. [PubMed: 23927221]
- Goldstein L, Byrd D, Saltzman E. The role of vocal tract gestural action units in understanding the evolution of phonology. *From action to language: The mirror neuron system*. 2006:215–249.
- Goldstein L, Pouplier M, Chen L, Saltzman E, Byrd D. Dynamic action units slip in speech production errors. *Cognition*. 2007; 103(3):386–412. [PubMed: 16822494]
- Gollan TH, Ferreira VS. Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2009; 35(3):640.
- Green DW. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*. 1998; 1(02):67–81.
- Grossberg S. Resonant neural dynamics of speech perception. *Journal of Phonetics*. 2003; 31:423–445.
- Kahn JM, Arnold JE. Articulatory and lexical repetition effects on durational reduction: speaker experience vs. common ground. *Language and Cognitive Processes*. 2013
- Kahn JM, Arnold JE. A processing-centered look at the contribution of givenness to durational reduction. *Journal of Memory and Language*. 2012; 67(3):311–325.
- Kello CT. Control over the time course of cognition in the tempo-naming task. *Journal of Experimental Psychology: Human Perception and Performance*. 2004; 30(5):942. [PubMed: 15462632]
- Kello CT, Plaut DC, MacWhinney B. The task dependence of staged versus cascaded processing: An empirical and computational study of Stroop interference in speech perception. *Journal of Experimental Psychology: General*. 2000; 129(3):340. [PubMed: 11006904]
- Kirchner R, Moore RK, Chen TY. Computing phonological generalization over real speech exemplars. *Journal of Phonetics*. 2010; 38(4):540–547.
- Kirov C, Gafos A. Dynamic phonetic detail in lexical representations. *Proceedings of the 16th international congress of phonetic sciences*. 2007:637–640.
- Lam TQ, Watson DG. Repetition reduction: Lexical repetition in the absence of referent repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2014; 40(3):829.
- MacLeod CM. Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin*. 1991; 109(2):163. [PubMed: 2034749]
- McClelland JL, Elman JL. The TRACE model of speech perception. *Cognitive Psychology*. 1986; 18:1–86. [PubMed: 3753912]
- McMillan CT, Corley M, Lickley RJ. Articulatory evidence for feedback and competition in speech production. *Language and Cognitive Processes*. 2009; 24(1):44–66.

- Meuter RF, Allport A. Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of memory and language*. 1999; 40(1):25–40.
- Monsell S. Task switching. *Trends in cognitive sciences*. 2003; 7(3):134–140. [PubMed: 12639695]
- Munson B, Solomon NP. The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research*. 2004; 47(5):1048.
- Norman, DA.; Shallice, T. Attention to action: Willed and automatic control of behavior. In: Davidson, RJ.; Schwartz, GE.; Shapiro, D., editors. *Consciousness and self-regulation: Advances in research and theory*. Vol. 4. New York: Plenum; 1986. p. 1-18.
- Olson DJ. Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*. 2013; 41(6):407–420.
- Pierrehumbert J. Word-specific phonetics. *Laboratory Phonology*. 2002; 7:101–139.
- Pierrehumbert JB. The next toolkit. *Journal of Phonetics*. 2006; 34(4):516–530.
- Pisoni, DB.; Levi, SV. Representations and representational specificity in speech perception and spoken word recognition. In: Gaskell, G., editor. *Oxford Handbook of Psycholinguistics*. Oxford: Oxford University Press; 2007. p. 3-18.
- Rumelhart, DE.; Hinton, GE.; McClelland, JL. A general framework for parallel distributed processing. In: Rumelhart, DE.; McClelland, JL.; the PDP Research Group. , editors. *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol. 1: Foundations. Cambridge, MA: MIT Press; 1986. p. 110-146.
- Sadat J, Martin CD, Alario FX, Costa A. Characterizing the bilingual disadvantage in noun phrase production. *Journal of Psycholinguistic Research*. 2012; 41(3):159–179.
- Scarborough, R. Doctoral dissertation. University of California Los Angeles; 2004. Coarticulation and the structure of the lexicon.
- Scarborough, R. Lexical and contextual predictability: Confluent effects on the production of vowels. In: Fougerson, C.; Kuhnert, B.; D’Imperio, M.; Vallee, N., editors. *Laboratory Phonology*. Vol. 10. Berlin: Mouton de Gruyter; 2010. p. 557-586.
- Scarborough R, Zellou G. Clarity in communication: “Clear” speech authenticity and lexical neighborhood density effects in speech production and perception. *The Journal of the Acoustical Society of America*. 2013; 134(5):3793–3807. [PubMed: 24180789]
- Smolensky P, Goldrick M, Mathis D. Optimization and quantization in gradient symbol systems: a framework for integrating the continuous and the discrete in cognition. *Cognitive Science*. 2014; 38:1102–1138. [PubMed: 23802807]
- Tilsen S. Metrical regularity facilitates speech planning and production. *Laboratory Phonology*. 2011; 2(1):185–218.
- Tilsen S. A Dynamical Model of Hierarchical Selection and Coordination in Speech Planning. *PloS one*. 2013; 8(4):e62800. [PubMed: 23638147]
- Walsh M, Möbius B, Wade T, Schütze H. Multilevel exemplar theory. *Cognitive Science*. 2010; 34(4): 537–582. [PubMed: 21564224]
- Wedel AB. Feedback and regularity in the lexicon. *Phonology*. 2007; 24(1):147.
- Wright, R. Factors of lexical competition in vowel articulation. In: Local, JJ.; Ogden, R.; Temple, R., editors. *Laboratory Phonology VI*. Cambridge: Cambridge University Press; 2004. p. 75-87.