Running Head: LIMITED INTERACTION

Limited Interaction in Speech Production:

Chronometric, Speech Error, and Neuropsychological Evidence

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) 467-7092
Fax: (847) 491-3770
Email: goldrick@ling.northwestern.edu

Abstract

Results from chronometric and speech errors studies provide convergent evidence for both lower and upper bounds on interaction within the speech production system. Some degree of cascading activation is required to account for patterns of speech errors in neurologically intact and impaired speakers as well as the results of recent chronometric studies. However, the strength of this form of interaction must be limited to account for the occurrence of selective deficits in the production system and restrictions on the conditions under which interactive effects influence reaction times. Similarly, some amount of feedback from phonological to word-level representations is necessary to account for patterns of speech errors in neurologically intact and impaired individuals as well as the influence of phonological neighbors on response latency. This interactive mechanism must also be limited to account for restrictions on the types of speech errors produced following selective deficits within the production system. Results from a variety of empirical traditions therefore converge on the same conclusion: interaction is present, but it must be crucially limited.

Limited Interaction in Speech Production:

Chronometric, Speech Error, and Neuropsychological Evidence

Historically, research into single word production has been conducted within two distinct empirical traditions (Levelt, 1999)[1]. One has focused on speech errors from a variety of sources. Within this tradition, early theories (e.g., Garrett, 1980) claimed that spoken production processes were highly discrete. The various sub-processes underlying speech production were assumed to be highly independent and non-overlapping. During the 1980s, it became clear that such theories were unable to account for the full range of speech error data, leading to the introduction of (localist connectionist) interactive-activation theories (e.g., Dell, 1986; Stemberger, 1985). These theories incorporated the mechanisms of cascading activation and feedback to allow spoken production processes to interact. A parallel empirical tradition focuses not on errors but on reaction time—the latency of speech production behavior. In contrast to research in errors, extensive work by Levelt, Roelofs, Meyer, and their colleagues (see Levelt, Roelofs, & Meyer, 1999, for a review) has found that theories assuming a discrete relationship between semantic and phonological processes provide an impressive qualitative and quantitative fit to much of the existing latency data.

Why is there disagreement between these two approaches? If speech production is an interactive process, why is a system with substantial discreteness so successful at modeling chronometric data? If it is discrete, why do speech errors require the presence of interaction? This review argues that the first step towards resolving this conflict is recognizing the presence of limited interactivity—that is, the presence of not just lower but also upper bounds on the strength of both cascading activation and feedback.

After laying out a generic processing framework for spoken word production, evidence supporting limited interaction is reviewed. A broad array of results—from speech errors in neurologically intact and impaired speakers, as well as chronometric and accuracy studies of production—converges on the conclusion that although interactive mechanisms are present, they must be crucially limited. Specifically, the results support 3 features for the adult speech production system:

1. Significant cascading activation from word-level to phonological representations is present, but its strength must be limited.
2. Significant feedback from phonological to word-level representations is present, but its strength must be limited.
3. Feedback from word-level to lexical semantic representations is either absent or functionally insignificant

New simulation data are provided to reinforce these claims. The discussion concludes with open questions and future directions for research.

Representational and processing assumptions

Scope of the review

This review concerns the question of interaction among processes that map from lexical semantic representations (e.g., the concept of a furry four-legged feline {CAT}) to phonological representations stored in long term memory (e.g., the phonemes /k/ /ae/ /t/). Two broad clarifications are in order. First, since the review concerns speech production, interactions from other perceptual domains (e.g., object recognition) will not be discussed. Second, this review focuses on lexical access of single, morphologically simplex words. Since full discussion of the influence of syntactic information would require broadening the discussion to include sentence production and other connected speech tasks, the role of syntactic features is also omitted (see Vigliocco & Hartsuiker, 2002, for a review).

Representational assumptions

Lexical semantic representations. Lexical semantic representations are amodal representations (common to all sensory and production modalities) of the meaning of lexical items in a particular language (e.g., the feature set {furry four-legged feline} for lexical concept {CAT}; see the top of Figure 1A-D). These representations codify the message the speaker intends to communicate in a language-specific way, allowing more general (non-linguistic) conceptual processes to interface with those that specify linguistic form. For the purposes of this discussion, lexical semantic representations are depicted as sets of distributed semantic features (following, e.g., Caramazza, 1997; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997), although the findings are also consistent with theories assuming unitary semantic representations (e.g., Levelt et al., 1999).

Phonological representations. Phonological representations are stored, sub-lexical representations of the spoken form of lexical items (e.g., /k/ /ae/ /t/; see the bottom of Figure 1A-D). For the purposes of discussion, these have been depicted as segments (following Dell et al., 1997), although the results are also consistent with the storage of additional aspects of phonological structure (e.g., metrical structure; Levelt et al., 1999).

L-level representations. As shown in Figure 1, theories of speech production generally assume a third level of representation mediating the mapping between semantic and phonological representations—a lexical, or word-level, level of representation (e.g., <CAT>). There is widespread disagreement as to whether a single lexical level is sufficient (see Caramazza, 1997; Levelt et al., 1999, for discussion) and whether lexical representations are amodal or modality specific (see Caramazza, 1997; Dell et al., 1997; Levelt et al., 1999, for discussion). To emphasize the common properties of this theories, this discussion follows Rapp & Goldrick (2000) by assuming a single 'L-level' of lexical representation (which may or may not be modality specific). With the exception of homophonic primes[2] and the influence of feedback across multiple representational levels (see the general discussion for further details), adopting this assumption does not impact the interpretation of the majority of the findings reviewed here.

Processing stages

Most theories of speech production assume that the mapping from lexical semantic to L-level then phonological representations is accomplished through a series of processing stages (stemming from Garrett's (1980) "two-stage" theory of speech production). Here, three broad stages of processing—corresponding to a selection of a representation at each of the three levels discussed above—are assumed to be involved in lexical access in speech production.

Stage 0: Lexical semantic processing. Processing begins with the selection or generation of a message that the speaker wishes to express, and ends with the selection of a lexical semantic representation specifying a lexicalized concept within the speaker's language. This corresponds to selecting a concept from one's language to express a pre-verbal message.

Stage 1: L-level processing. Processing begins with the selection of a lexical semantic representation corresponding to a particular concept, and ends with the selection of an L-level unit. This corresponds to the selection of a particular word or lexical item to express a concept.

Stage 2: Phonological processing. Processing begins with the selection of an L-level unit and ends with the selection of a phonological representation. This corresponds to retrieving the phonological components of a word from long term memory.

Note that during Stages 1 and 2, the units of representation are words; errors at these levels are therefore restricted to words. During Stage 3, sub-lexical representations are selected; at this level, both word and nonword errors can occur.

As seen above, each stage begins with the selection of a representation at one level (e.g., lexical semantics) and ends with the selection of a representation at the next level in the processing stream (e.g., the L-level). In this context, selection refers to those mechanisms that allow one or more representations[3] to dominate subsequent processing. For example, many authors (e.g., Roelofs, 1992) have argued that during L-level selection multiple semantically related words are activated (e.g., in addition to target <CAT>, <DOG> may be partially activated). In spite of this partial activation of other words, on the vast majority of productions the target is correctly produced. Selection mechanisms are what allow the target, rather than its competitors, to determine the form that is ultimately produced.

In speech production theories, a variety of mechanisms accomplish this by singling out the most active contextually appropriate representation and enhance its activation relative to that of competitors: either by boosting the selected representation's activation (e.g., Dell, 1986; Dell et al., 1997; Rapp & Goldrick, 2000); inhibiting less active competitors (Berg & Schade, 1992; Cutting & Ferreira, 1999; Dell & O'Seaghdha, 1994; Harley, 1993; Meyer & Gordon, 1985; Schade & Berg, 1992; Stemberger, 1985); or restricting the activation flow from competitors (Laine, Tikkala, & Juhola, 1998; Levelt et al., 1999). The term "selection strength" will be used to refer to the degree to which selected representations are enhanced. In systems with strong selection points, selected representations will be much more active than competitors (due to very large boosts in activation, very strong inhibition, or very strong restrictions on competitors' activation flow). Weakening selection points will decrease the activation advantage of the selected representation (due to smaller boosts in activation, weaker inhibition, or weakened restrictions on competitors' activation flow).

Mechanisms for increasing interaction

A discrete system is defined by three characteristics: one, processing involving one representational level is not initiated until selection has occurred at the previous level; two, selection strength is extremely high, such that only selected representations are allowed to pass on activation to other processing levels; and three, activation flows in a strictly feed-forward manner. These properties have several notable consequences. First, due to the strength of selection in discrete systems, subsequent processing stages only receive information about the representations selected at prior levels (e.g., phonological processing occurs for the selected L-level representation alone). Second, due to the staged, feed-forward nature of processing, later stages cannot influence processing at higher levels in the system (e.g., phonological processing cannot influence L-level processing). There are two widely used mechanisms for increasing interaction beyond this discrete endpoint: cascading activation and feedback.

Cascading activation. In contrast to discrete systems, those with cascading activation reject the assumption that representations can pass on activation only after they have been selected. This claim leads to the weakening of the first two characteristics of discrete systems. Representations can pass on activation before selection has occurred, such that processing at subsequent stages is initiated prior to selection (e.g., L-level processing can begin prior to the completion of lexical semantic processing). De-linking activation flow from selection points also leads to a decrease in selection strength. Typically, cascading activation systems allow non-selected representations to pass on activation to subsequent processing stages both prior to and after selection has occurred (e.g., in processing target “cat,” the partially activated lexical concept {DOG} activates its corresponding L-level representation <DOG>).

All current theories of speech production assume that cascading activation is present between lexical semantic and L-level representations (Caramazza, 1997; Dell et al., 1997; Levelt et al., 1999; Rapp & Goldrick, 2000). During lexical semantic processing, multiple (semantically related) concepts are activated (via connections linking related concepts (Levelt et al., 1999) or overlapping distributed representations (e.g., Dell et al., 1997)). Cascading activation allows these multiple concepts to activate the L-level representation of words semantically related to the target.

Feedback. Systems with feedback relax the third assumption of discrete systems and allow later stages to influence processing at higher levels in the system. This creates a number of different effects, discussed in more detail below. The discussion here remains agnostic as to whether these effects are driven by mechanisms entirely internal to the production system (e.g., reciprocal connections between representational levels; Dell, 1986) or indirectly, via speech perception mechanisms (e.g., monitoring systems; Levelt et al., 1999; see Roelofs, 2004a, b, Rapp & Goldrick, 2004, for further discussion).

What degree of interaction do speech production theories require?

The following sections consider the consequences of incremental increases in interaction between adjacent representational levels in the production system. As interaction is increased, two questions are examined:

1. Does this increased amount of interaction explain results that cannot be accounted for by a system lacking this type of interaction? This provides a lower bound on the strength of this form of interaction.

2. Do the data suggest any limitations to the strength of this type of interaction? This provides an upper bound on the strength of interaction.

Figure 1A depicts the most discrete system considered here. The consequences of increasing cascading activation by allowing non-selected L-level units to activate phonological representations (Figure 1B) are examined first. Feedback interactions are then examined, moving upward through the processing system (considering feedback from phonological to L-level units, as shown in Figure 1C, then from the L-level to lexical semantics, as in 1D).

(Figure 1 about here)

Evidence Supporting Limited Cascading Activation

Limited cascading activation: Speech error evidence

The mixed error effect. Probably the most widely cited piece of evidence for interaction in the spoken production system is the mixed error effect. This refers to the observation that in a variety of contexts mixed errors (e.g., “cat”-> “rat”) occur more often than predicted based on the rates of purely semantic (e.g., “cat”-> “dog”) or purely phonological (e.g., “cat”-> “cab”) errors (for more detailed discussion of how predicted rates have been calculated, see Goldrick & Rapp, 2002). System with cascading activation predict that such an effect can be generated at the phonological level. For example, as shown in Figure 1B, cascading activation from <RAT> activates /r/. In contrast, /b/ receives no support from the purely phonologically related <CAB>. This activation difference makes “rat” a more likely phonological error than “cab,” meaning that in a cascading activation system the rate of mixed errors will exceed the simple sum of purely semantic and purely phonological errors. Such an effect cannot be produced in a discrete system. As shown in Figure 1A, in such systems /r/ is no more active than /b/—“rat” is therefore no more likely to occur as a phonological error than “cab.”

Consistent with cascading activation, many empirical reports of mixed error effects can plausibly be attributed to the phonological level. Many studies of spontaneous speech errors have shown mixed error effects (Dell & Reich, 1981; Harley, 1984; Harley & MacAndrew, 1995, 2001; Martin, Gagnon, Schwartz, Dell, & Saffran, 1996; but see del Viso, Igoa, & García-Albea, 1991[4] and Igoa, 1996, for null results). Given that it is unclear at what level(s) of the processing system spontaneous speech errors arise, these effects can be plausibly attributed to the phonological level. Similarly, mixed error effects in experimentally induced speech errors (Brédart & Valentine, 1992; Martin, Weisberg, & Saffran, 1989; but see Levelt, 1983, for a null result) can be plausibly attributed to the phonological level. Finally, the mixed error effects observed in the production errors of many aphasic individuals (Blanken, 1998; Dell et al., 1997; Kulke & Blanken, 2001; Martin et al., 1996; Rapp & Goldrick, 2000; but see Best, 1996 and Nickels, 1995, for null results) can also be attributed to phonological processing.

Two caveats are warranted. As discussed in more detail below, mixed error effects are also predicted to occur at the L-level in systems incorporating feedback. However, the data discussed in this section do not require this increase in interactivity; they can be parsimoniously explained by a minimal increase in interaction over a discrete system (i.e., by the incremental addition of cascading activation). Second, some authors have attributed the mixed error bias to the influence of perceptual monitoring systems. We return to this alternative account following the discussion of chronometric evidence supporting cascade.

<u>Homophone priming of semantically-related word substitution errors</u>. Ferreira and Griffin (2003; see Burke, Locantore, Austin, & Chae, 2004, for related speed and accuracy findings) had participants name pictures following visual presentation of cloze sentence fragment. These cloze sentences fragments primed participants to make a large number of

semantic word substitution errors relative to conditions when an unrelated word was primed[5] (e.g., "The woman went to a convent to become a…", primed participants to misname a picture of a priest "nun.") With respect to interaction, a significant numbers of errors were also induced when cloze sentences primed homophones of semantic neighbors (e.g., "I thought that there would be some cookies left, but there were…" primed participants to produce "none" for a picture of a priest). The errors in the homophone priming condition can be plausibly accounted for by cascading activation. Homophone primes serve to pre-activate the phonological representation of semantic neighbors (e.g., /n/ /ʌ/ /n/). During processing of the target picture, cascading activation from L-level representations of semantic neighbors (e.g., <NUN>) serves to drive these representation over threshold, leading to substitution errors.

<u>Evidence for limits on cascading activation to phonology</u>. Although a significant degree of cascading activation is clearly required, other evidence supports upper bounds on its strength. Simulation results show that systems with very strong cascade activation predict patterns inconsistent with the empirical data.

Figure 2 reports new simulation results documenting how systems with very strong cascading activation predict that phonological processing will be indirectly disrupted following damage localized to the L-level (see Rapp & Goldrick, 2000, for implementation details[6]). Following simulated L-level damage (activation noise 0.7), the rate of nonword errors was used to index the disruption of phonological processing. Such errors can only be produced during phonological selection (at the L-level, the representational units are words). The extent of cascade was manipulated by altering selection strength—here, the predetermined high level to which a selected representation's activation is boosted over that of competitors ('jolt' strength). Note that weaker L-level selection entails a greater amount of cascading activation. As shown in

Figure 2, the rate of nonword errors in the simulation is related to the strength of L-level selection; significant numbers of nonword errors are found when selection strength is weakened. This shows that in systems with strong cascading activation phonological processing can be indirectly disrupted when L-level selection strength is too weak. Similar results have been observed in a variety of systems with weak selection points (e.g., when feedback is incorporated or in attractor-based distributed connectionist systems; Rapp & Goldrick, 2000; Plaut & Shallice, 1993).

(Figure 2 about here)

This predicted consequence of systems with strong cascading activation is inconsistent with the empirical data; individuals with deficits localized to the L-level do not produce nonword errors. A number of studies have documented speakers with modality-specific impairments to speech production that produce only semantic errors (Basso, Taborelli, & Vignolo, 1978; Caramazza & Hillis, 1990; Miceli, Benvegnú, Capasso, & Caramazza, 1997; Nickels, 1992; Rapp, Benzing, & Caramazza, 1997; Rapp & Goldrick, 2000). Rapp and Goldrick (2000; see also Caramazza & Hillis, 1990) argued that this pattern of performance can be best explained by postulating a deficit to selecting the target L-level representation. The possibility that their errors result from a deficit to amodal lexical semantic representations is precluded by the modality-specific nature of their deficits. Furthermore, detailed testing of semantic processing in these individuals has revealed no significant impairment (see, e.g., Goldrick & Rapp, 2002). Second, the production of only semantic errors suggests that the deficit cannot arise at the phonological level. Given that the units of phonological processing are sub-lexical, it is highly likely that a deficit at this level will result in the production of phonologically related word and nonword errors. For example, in all of the architectures simulated by Rapp & Goldrick (2000), damage to phonological processing invariably led to the production of phonologically related word and nonword errors. The most likely locus of their deficit is to L-level processing.

These case studies therefore provide an upper bound on the strength of cascading activation. As shown above, if cascading activation was very strong, individuals with L-level deficits are predicted to produce nonword errors. These case studies show that they do not; their errors are limited to semantic word substitutions.

Limited cascading activation: Chronometric and accuracy evidence

Facilitation of words phonologically related to semantically activated items. A number of studies have shown phonological facilitation (reflected by reduced naming latency and/or higher accuracy) for words that overlap in form with lexical items that have been semantically activated (e.g., semantic neighbors of the target). Each of these studies provides evidence for cascading activation from L-level to phonological representations. In each experiment, a non-target L-level representation is activated via semantic processes. Activation from this L-level unit cascades to phonological representations, facilitating production of words that overlap with its phonological representation. Without cascading activation, no pre-activation would be present, and no facilitation would be observed.

To consider a specific example, in Peterson and Savoy's (1998) paradigm, near-synonyms of a target picture name become activated at the lexical semantic level (e.g., for target "couch", {SOFA}). These conceptual representations activate their corresponding L-level representations (e.g., <SOFA>), and, via cascade, their phonological representations (e.g., /s/ /oʊ/ /f/ /ə/). In their cued-response paradigm, this facilitates production of visually presented words related to the near-synonym (e.g., "soda") relative to phonologically unrelated controls (e.g., "horse").

These studies are summarized in Table 1. The first column lists the citation. The second column briefly describes the paradigm. The third column lists the (non-target) lexical item that is claimed to be activated via semantic processes. Finally, the fourth column describes the word phonologically related to the semantically activated lexical item that shows evidence of facilitation.

(Table 1 about here)

<u>Interference for words phonologically unrelated to semantically activated items</u>. Several studies have shown the complementary effect to those above: increased reaction times for targets that are phonological <u>dissimilar</u> to semantically activated words. Cascading activation provides a plausible mechanism for this effect. Activation from semantically related words enhances the activation of non-target phonological representations; this may make retrieval of target phonological representations more difficult (e.g., by increasing the competition for selection at the phonological level).

To consider a specific example, Jescheniak and Schriefers' (1998) paradigm relied on the phonological activation of near-synonyms of a target picture (e.g., for target "couch," the activation of "sofa" at the phonological level; see the discussion of Peterson and Savoy, 1998, above). Cascade from the near-synonym increases the activation of the phonological representation of distractor words which are related to it but dissimilar to the target picture name (e.g., the phonological representation of "soda"—similar to "sofa" but not "couch"—will be more active than that of "horse"). The increased activation of these distractors during the picture/word interference task yields more competition for selection at the phonological level, slowing reaction times in picture naming.

These studies are summarized in Table 2. The first three columns follow the format of Table 1. The fourth column identifies the critical distractor word. In these studies, retrieval of the phonological representation of the picture name was inhibited. This inhibition was attributed to a distractor word (identified in column 4) whose activation was enhanced by its phonological relationship to the semantically activated lexical item.

(Table 2 about here)

Evidence for upper limits on cascading activation to phonology. Although the above chronometric studies provide clear evidence for cascading activation, this does not imply that it is unlimited in its power and influence. Across these studies, it appears that cascading activation effects in chronometric studies are only observed in three conditions.

1. Very strong semantic similarity. *Near-synonyms* (Jescheniak & Schriefers, 1998; Peterson & Savoy, 1998) have nearly complete overlap with target semantic representations. A similar situation holds for *translation equivalents* (Costa et al., 2003; Hermans, et al., 1998) as well as *alternative target names at different levels of categorization* (Jescheniak, Hantsch, & Schriefers, 2005).
2. Very high phonological overlap. Words cannot be any more phonologically similar than *homophones* (Cutting & Ferreira, 1999; Taylor & Burke, 2002).
3. Simultaneous overlap with the semantic and phonological structure of the target. *Mixed* neighbors simultaneously overlap the target at both levels of structure (Damian & Martin, 1999; Starreveld & La Heij, 1995, 1999; Taylor & Burke, 2002). *Cognates* exhibit very strong, simultaneous overlap at both levels (Costa et al., 2000; Gollan & Acenas, 2002).

With respect to strong semantic or phonological similarity, parallel studies with stimuli that exhibit neither simultaneous overlap at semantic and phonological nor high degrees of overlap at either single level fail to find interactive effects. For semantic associates that are not synonyms (e.g., for target “sheep,” “goat”), phonological facilitation effects on non-homophones (e.g., “goal”) are either absent (Levelt, Schriefers, Vorberg, Meyer, Pechmann, & Havinga, 1991a; Peterson & Savoy, 1998) or extremely weak (O’Seaghdha & Marin, 2000). With respect to simultaneous overlap, Melinger and Abdel Rahman (2004) utilized a version of the picture-word interference task with two distinct distractor words (e.g., naming a picture of a pig with PILL and SOCK simultaneously superimposed on the picture). When semantic and phonological overlap equivalent to that of a single mixed neighbor was carried by two different distractors, interactive effects are not observed. For example, the naming latency of “pig” accompanied by both a semantic and a phonological distractor (e.g., BEAR, PILL) reflected a simple additive interaction of semantic interference (from BEAR) and phonological facilitation (from PILL).

These contrasting findings suggest that although cascading activation from L-level to phonological representations is present, it must be limited. Specifically, the divergent results across studies reflects the restriction of chronometric effects due to cascading activation to situations with strong, convergent activation. For example, in most situations, the source of the cascading activation is fairly weak; limited cascading activation from this weak source cannot significantly influence processing at the phonological level. In certain exceptional cases of strong semantic similarity (e.g., near-synonyms, translation equivalents), activation of the non-selected L-level unit is very strong, allowing it to produce interactive effects in spite of weak cascade. With respect to phonological priming, cascading activation usually provides at best partial support for the response. However, in the case of homophones and cognates, virtually the entire response is supported by cascading activation, allowing these stimuli to exhibit an effect.

Why then are effects observed with mixed neighbors (which have neither strong semantic activation nor complete phonological overlap)? Unlike stimuli in the other studies, mixed neighbors do not provide the sole source of activation for the response on which interactive effects are observed. Activation from the mixed neighbor converges with that of the (already activated) target (Dell & O'Seaghdha, 1991). In other studies, the phonological representation which is used to indicate the presence of interaction (e.g., "goal" in Levelt et al., 1991a) receives no activation from the target; weak cascading activation is its only source of support. In contrast, when mixed neighbors are activated, they modulate the activity of representations that are already above threshold. Therefore, although the phonological and L-level representation of mixed neighbors themselves are not highly activated, they can exert an influence on representations that receive other sources of support. This provides a ready account of the lack of interactive effects in Melinger and Abdel Rahman (2004). In this study, activation from the L-level representation of the semantic distractor fails to converge with the phonological representation of the target (e.g., BEAR fails to converge with "pig"). Since the phonological and semantic distractors fail to converge on the target's phonological representation, they fail to enhance its activation any more than predicted by each individual distractor. Finally, convergent activation also accounts for the positive results of Morsella and Miozzo (2004) and Navarette and Costa (2005); as with mixed distractors, activation from the pictorial distractor converges with that of the already activated target word.

Role of the monitor

Some researchers (e.g., Levelt, 1983; Levelt et al., 1999; Roelofs, 2004a) have questioned some of the evidence for cascading activation. For example, they have attributed mixed error biases solely to the influence of monitoring systems (see Hartsuiker, 2006, for further discussion). As described by Roelofs (2004a), prior to articulation the phonological form of the speaker's utterance is fed back through the comprehension system. The speaker monitors for mismatches between the concept activated by this phonological form and the intended concept. Since mixed neighbors partially activate the intended concept, it is more difficult to detect them compared to purely semantic errors (e.g., mixed error "calf" will partially activate the target "cat" within the perceptual system, while pure semantic neighbor "dog" will not).

There are two major problems with this account. Monitoring theories have remained quite underspecified, making their precise predictions under particular processing conditions rather unclear (see Goldrick & Rapp, 2002; Rapp & Goldrick, 2004, for further discussion). Furthermore, even perception-based monitoring accounts (e.g., Levelt et al., 1999; Roelofs, 2004a) appear to require cascading activation in addition to monitoring to account for phonological effects of pictorial distractors (Morsella & Miozzo, 2004; Navarette & Costa, 2005). In order for distractors or mixed errors to induce monitoring effects, their phonological representations must be activated (if they are not, they cannot engage perceptual processes). However, pictorial distractors can only activate their phonology via semantic and L-level representations (Hillis, 2001). In other words, the monitor requires cascade to be present to account for these data—without it, the phonology of distractors would simply be inactive and the monitor would not be engaged. Given that all accounts require this mechanism, it is more parsimonious to attribute the various effects discussed here to cascading activation.

Implications for the strength of cascading activation

There is abundant evidence from both traditions of speech production research to support the presence of both upper and lower bounds on cascading activation. Note the discussion above is limited to those effects that can be accounted for solely by the presence of cascading activation; namely, phonological effects attributable to semantically-activated L-level representations. As discussed below, cascading can also interact with feedback to allow other L-level representations (i.e., formally related words) to influence phonological processing.

Although the data support both upper and lower bounds on cascade, it is unclear how to more precisely formulate these limitations. One concrete proposal that is not consistent with the data is that of Levelt et al. (1999). They claimed the upper bound of cascade is defined by words that are semantically similar and contextually appropriate targets (e.g., synonyms). Empirical data suggest this is too strong a limitation on cascade. With respect to errors, phonological level errors are facilitated for members of the same semantic category (e.g., "cat" and "rat"), which are clearly neither synonymic nor contextually appropriate. With respect to chronometric data, effects are also observed for contextually inappropriate items (e.g., cognates or translation equivalents in a monolingual context; alternative target names at inappropriate levels of categorization; see Jescheniak et al., 2005, for further discussion). It is unclear what alternative formulation would provide a better account of restrictions on cascade. This is in part due to lingering disagreements between chronometric and error data. For example, while direct chronometric effects of cascading activation are only observed under very high degrees of conceptually-driven activation (e.g., cognates, near-synonyms), speech error effects are observed in a wider range circumstances (e.g., categorically-related semantic associates). This issue is returned to in the general discussion.

Evidence Supporting Limited Feedback: Phonological to L-level Representations

As discussed below, adding feedback from phonological to L-level representations (Figure 1C) accounts for a number of interactive effects that cannot be produced by cascading activation. Following the limitation of the discussion to data from speech production, arguments for feedback based on shared representations for production and perception (see Martin & Saffran, 2002, for a review) will not be considered here.

Evidence supporting limited feedback effects on L-level selection

Above, cases of deficits to L-level processing were reviewed. These individuals had modality-specific impairments to speech production that resulted in the production of only semantic errors. Deficits to lexical semantic processing were ruled out, as these individuals did not exhibit comprehension deficits. Phonological impairments were ruled out because they failed to produce purely phonologically related word and nonword errors.

Rapp and Goldrick (2000) examined the performance of two of these individuals (PW and RGB) and found that both exhibited a significant mixed error effect. Specifically, their semantic errors showed greater phonological overlap than predicted by chance (i.e., a higher degree of phonological overlap than randomly paired targets and errors; see Rapp & Goldrick, 2000, for a detailed description of the analysis). This supports the presence of feedback. Feedback from phonological representations serves to facilitate the L-level representation of mixed neighbors relative to those of purely semantic neighbors. This facilitation translates into a greater likelihood of errors, producing the mixed error effect. Note that unlike the mixed error data reviewed above these effects cannot be accounted for by cascading activation. Cascading activation only allows for a mixed error benefit at the phonological level, and in these cases there is independent evidence that their deficit is to L-level processing.

Upper limits on feedback to L-level selection. The many cases of L-level deficits reviewed above also provide an upper bound on the strength of feedback. As described above, despite relatively low accuracy levels (e.g., 28% error in untimed picture naming), these individuals produce only semantic errors. This suggests a limit to the strength of feedback; very strong feedback predicts that these individuals should also produce formally related errors (Rapp & Goldrick, 2000). Figure 3 reports new simulation results that further document the relationship between feedback strength and the production of formal errors following L-level damage. The simulation reported here implements a system with cascading activation and feedback from the phoneme to the L-level (see Rapp & Goldrick, 2000, for implementation details[7]). Damage to the L-level was simulated by setting L-level noise to a high level (0.7). Feedback strength was set to two very low levels (.01 and .025) and then varied from 0.05 to 0.50 in 0.05 increments. As shown in Figure 3, when feedback strength is limited, formal errors were produced at negligible levels. However, when feedback strength was increased beyond this point, a highly significant number of formal errors occurred at L-level selection. If feedback is not sufficiently limited, theories of speech production cannot account for the production of only semantic errors following L-level damage[8].

(Figure 3 about here)

Evidence supporting indirect effects of feedback on phonological processing

The results reviewed above suggest that formally related neighbors are not highly active at the point of L-level selection. However, following L-level selection, feedback can continue to exert an influence on L-level representations, allowing phonological related words (e.g., <HAT> for target <CAT>) to grow in activation. If cascading activation is present, these formally related L-level representations will pass on activation to their phonological representations. The positive feedback loops created by this interaction of feedback and cascade allow L-level properties to bias the outcome of phonological processing (Dell, 1986)[9].

Phonologically related error outcomes are biased by lexical properties. To concretely illustrate the influence of positive feedback loops, consider their most well-studied consequence: the lexical bias effect. Suppose that subsequent to L-level selection, phonological processing is randomly disrupted, increasing the activation of a non-target phoneme that creates a word outcome (e.g,. while naming "cat", the activation of /h/ in onset is randomly increased, creating the potential word error "hat"). Via feedback, this non-target phoneme can pass its increased activation back to the L-level unit corresponding to the word outcome (e.g., <HAT>). This drives up the activation of the L-level unit. Via cascade, this L-level unit then reactivates its constituent phonemes (e.g., /h/). The cycle then repeats, increasing activation levels of both phonological and L-level representations corresponding to this word outcome—hence, a positive feedback loop. In contrast, if random disruptions boost the activity of phoneme that creates a nonword outcome (e.g., /z/), it will be unable to enter into such a strong feedback relationship. Since there is no L-level unit corresponding to this outcome (e.g., <ZAT> is not present), this phoneme will be unable to benefit from the influence of strong feedback loops[10]. Phonemes that create word outcomes will therefore be more active than those that create nonword outcomes. This will bias errors at the phonological level to produce more word outcomes than matched nonword outcomes—creating the lexical bias effect. Simulation results consistent with this analysis have been reported in multiple architectures (e.g., Dell, 1986, Rapp & Goldrick, 2000).

Empirically, the lexical bias effect has been reported in studies of spontaneous speech errors (Dell & Reich, 1981; Harley, 1984; Nooteboom, 2005b; Stemberger, 1985; but see del Viso et al., 1991 and Garrett, 1976, for null results), experimentally-induced errors (Baars, Motley, & MacKay, 1975; Dell, 1986, 1990; Humphreys, 2002; Hartsuiker, Corley, & Martensen, 2005; Nooteboom, 2005a), as well as in errors produced by aphasic speakers (Best, 1996; Gagnon, Schwartz, Martin, Dell, & Saffran, 1997; but see Nickels & Howard, 1995, for a null result).

It should be noted that the lexical bias effect may be influenced not only by automatic mechanisms such as feedback, but may also be modulated by mechanisms monitoring the outcome of speech production processes. Current evidence from experimentally induced errors suggest that both mechanisms are required to provide a full account of performance (see Harstuiker et al., 2005, for discussion).

If cascading activation from the L-level plays a crucial role in the lexical bias effect, the strength of the bias should vary with lexical frequency. Most current theories of speech production (e.g., Dell 1990; Jescheniak & Levelt, 1994; Miozzo & Caramazza, 2003; Roelofs, 1997), assume that lexical frequency[11] effects are localized to L-level representations (i.e., in their selection thresholds; time required for L-level representations to accumulate activation; verification times for lexical representations). Since the L-level representations of low frequency words are in general more weakly activated than those of high frequency words, they should transmit less cascading activation to the phonological level. This activation difference should cause errors to exhibit a frequency bias (i.e., errors at the phonological level should be more likely to result in high frequency than lower frequency words). The frequency bias has been observed in errors in neurologically intact speakers (see Harley & MacAndrew, 2001, for a review of positive results in spontaneous speech errors; but see Dell, 1990, for a null result in experimentally induced errors), as well as in a number of studies of individuals with acquired neurological impairments (Blanken, 1990, 1998; Gagnon et al., 1997; Goldrick & Rapp, in press; Martin, Dell, Saffran, & Schwartz, 1994; but see Best, 1996, for a null result).

Note that these effects cannot be produced by a system with only cascading activation. In such a system, the only L-level units that are active are semantic neighbors of the target (see Figure 1B). As discussed above, this serves to the boost the activation of mixed neighbors at the phonological level. It does not boost the activation of purely phonologically neighbors. Thus, in a system with only cascading activation, non-target phonemes that correspond to word outcomes not semantically related to the target (e.g., “hat” for “cat”) are no more active than those that correspond to nonword outcomes (e.g., “zat”). Similarly, with respect to a frequency bias, all formally related outcomes will be equally activated, regardless of their lexical frequency. Feedback is required to activate purely phonologically related words at the L-level; without it, these lexical effects cannot be produced at the phonological level.

Neighborhood density effects. The previous section documented how feedback enhances the representation of non-target words. It can enhance the phonological representation of the target as well. For example, as shown in Figure 1C, feedback activates <HAT> and <CAB>, and they enhance the activation of phonological structure they share with the target “cat” (e.g., /ae/). This facilitatory effect of formal neighbors predicts that words with many neighbors should more quickly and accurately retrieve their phonological representations compared to words with few neighbors (see Dell & Gordon, 2003, for simulation results supporting this prediction).

Several empirical findings are consistent with this claim. With respect to chronometric experiments, picture naming latency for high density words is shorter than that of matched low density words (Vitevitch, 2002; Vitevitch, Armbrüster, & Chu, 2004). These effects are arguably driven by the lexical properties of the stimuli. Density effects on latency are still present when responses are initiated via button press, suggesting that the effects are not attributable to articulatory differences between stimuli (Vitevitch, 2002). Furthermore, the effect is still found when sub-lexical properties of the stimuli are explicitly controlled (e.g., phonotactic probability; Vitevitch, 2002). With respect to the accuracy of lexical retrieval, words in high density neighborhoods (i.e., words that are phonologically similar to many other words) are less susceptible to TOT states than low density words (Harley & Bown, 1998; Vitevitch & Sommers, 2003). Experimentally-induced (Stemberger, 2004; Vitevitch, 2002) and spontaneous (Vitevitch, 1997) speech errors are more likely to occur on forms in low density neighborhoods. Finally, individuals with acquired speech production deficits produce more errors on words in low density as compared to those in high density neighborhoods (Goldrick & Rapp, in press; Gordon, 2002).

Implications for the strength of feedback and cascading activation. Although these lexical effects suggest that both feedback and cascading activation are present, they are also consistent with restrictions on both forms on interaction. Since these effects are due to positive feedback loops that allow activation to build up over time, they can be realized by fairly weak levels of interaction.

Assuming that feedback is weak implies that formally related words will require time to become significantly active at the L-level. This is consistent with the claim that prior to L-level selection formally related words are not strongly active. As processing continues, the positive feedback loops will allow the formally related words to become more active and influence phonological processing. Similarly, assuming weak cascade implies that only strongly activated L-level representations will be able to influence phonological processing. This is consistent with the restrictions on cascade from semantically-related items noted in the first section of the paper.

With respect to formally related words, weak cascade interacts with weak feedback to gradually boost the activation of their L-level representations, allowing them to influence phonological processing following L-level selection. In contrast, since purely semantically related neighbors do not share the phonology of the target, they are not re-activated by feedback. Their influence therefore decays over time, while formally related neighbors grow in strength.

As an existence proof that limited feedback and cascade can produce the proper mix of interactivity and discreteness, Rapp and Goldrick (2000) presented simulation results from systems where phonological to L-level feedback was set to 1/2 the strength of feedforward connections. Cascade was also limited by incorporating a strong L-level selection process. These simulations demonstrated attested effects of interaction (e.g., lexical bias, mixed error effects at the L-level) without resulting in unattested effects (e.g., following L-level damage, there were very few formal errors or nonword errors produced).

Evidence Supporting Absent/Functionally Insignificant Feedback from L-level to Lexical Semantic Representations

Increasing feedback by allowing L-level representations to feed back and activate lexical semantic representations creates the potential for information at lower levels in the production system (e.g., L-level information such as lexical frequency, or phonological information such as phonological overlap) to influence semantic processing. For example, as shown in Figure 1D, feedback from <RAT> transfers its activation (enhanced by phonological overlap with the target) back to the lexical semantic level[12]. Note that following the discussion above, feedback motivated by shared representations across production and perception (cf. Roelofs, 2004a, footnote 2) will not be considered.

Lexical frequency does not influence semantic processing. A number of studies have found that when pictures have to be recognized, but not named, lexical frequency effects are not found (see Jescheniak & Levelt, 1994, for a review; for a recent replication of this null result [for items matched for name agreement], see Schatzman & Schiller, 2004)[13]. For example, Jescheniak and Levelt (1994: Experiment 1) found a robust effect of lexical frequency on picture naming latency (pictures with high frequency names were named on average 62 milliseconds faster than those with low frequency names). In contrast, in a picture-word matching task (with basic-level word targets and distractors), no significant effect was found (pictures were matched with high frequency names only 6 milliseconds faster than with low frequency names). Jescheniak and Levelt attributed the difference in performance to differences in the processing stages required to perform each task. Picture naming requires access to L-level and phonological representations (to support oral production). In contrast, basic level word-picture matching is accomplished by comparing the semantic representations evoked by the pictorial and written word stimuli (DeCaro & Reeves, 2002). Thus, they argue, lexical frequency effects are observed only when post-semantic processing is required—consistent with the absence of feedback between these L-level and lexical semantic representations.

Mixed semantic-phonological error effects are absent at the lexical semantic level. Rapp and Goldrick (2000) examined the performance of KE, an individual with a deficit to lexical semantic processes. Like the individuals with L-level deficits described above, KE produced only semantic errors in speech production. As argued above, this suggests that his impairment was not at the phonological level. Unlike individuals with L-level deficits, KE exhibited remarkably similar performance across a range of perception and production tasks in a variety of modalities. For example, he produced similar rates of semantic errors in both writing and speaking, and made similar rates of semantic confusions in auditory, reading, and tactile comprehension tasks. This is consistent with damage to an amodal lexical semantic system shared across perception and production (see Hillis, Rapp, Romani, & Caramazza, 1990, for further analyses arguing against multiple coincident deficits).

Using the same techniques applied to the analysis of individuals with L-level deficits (see above), Rapp and Goldrick (2000) found that KE's semantic errors exhibited no more phonological overlap than predicted by chance. This suggests that at the lexical semantic level, mixed semantic-phonological neighbors are no more active than those that are purely semantically related. This is consistent with the absence of feedback to lexical semantic representations. If such feedback were present, the activation advantage that mixed semantic-phonological neighbors have at the L-level (as evidence by mixed semantic-phonological error effects at L-level selection) should be transferred to the lexical semantic level.

<u>Selective deficits to speech production</u>. As discussed above, if cascading activation from L-level to phonological representations is too strong, L-level deficits can indirectly disrupt phonological processing. Similar effects can arise due to feedback from L-level to semantic representations. If feedback is too strong, disruptions to L-level processing can feed back and indirectly disrupt lexical semantic processing (Rapp & Goldrick, 2000). For example, suppose L-level noise randomly boosts the activation of a semantic competitor at the L-level prior to lexical semantic selection (e.g., <DOG> is boosted prior to the selection of {CAT}). If feedback is strong, this increased activity can be fed back to its lexical semantic representation (e.g., {DOG}). Positive feedback loops between these two representations can drive the activity of the lexical semantic representation over that of the target, resulting in errors during lexical semantic selection. This would lead to the prediction that individuals with deficits to the L-level should

show semantic deficits; as discussed above, they do not. They produce semantic errors in speech production, but do not have deficits to comprehension processes.

Figure 4 presents new simulation results further documenting the indirect disruption of lexical semantic processing under conditions of strong feedback from the L-level. The simulation implements a system with cascading activation and feedback throughout the production system (see Rapp & Goldrick, 2000, for implementation details[14]). Damage to the L-level was simulated by setting L-level noise to a high level (0.7). Feedback strength from the L-level to lexical semantic representation was set to two very low levels (.01 and .025) and then varied from 0.05 to 0.50 in 0.05 increments. As shown in Figure 4, when feedback strength was very limited, lexical semantic processing was not disrupted. However, when feedback strength was increased beyond this point, many errors occurred at lexical semantic selection. Feedback from the L-level must be limited; if it is not, the occurrence of selective deficits to speech production following L-level damage cannot be explained.

(Figure 4 about here)

<u>Implications for the strength of feedback to lexical semantic representations</u>. Existing evidence places an upper bound on the amount of feedback from L-level to lexical semantic representations. However, it is possible for theories with weak feedback between these two levels of processing to account for these data. First, the simulation studies of Rapp and Goldrick (2000) as well as those reported here showed that feedback strength can be attenuated to such a degree that it is functionally insignificant. For example, as shown in Figure 4, at very low levels of feedback there is no indirect disruption of lexical semantic processing by L-level damage. Feedback may therefore be technically present, but it may have no consequence on processing. (For further discussion of functionally irrelevant feedback, see Rapp & Goldrick, 2000; Roelofs,

2003.) A second possibility is that feedback does exert an influence, but that this can only be found after lexical concepts have been selected. This would be analogous to the situation observed at the L-level for phonologically related neighbors. One area for future research is therefore the question of what lower bounds can be placed on the strength of feedback from L-level to semantic representations.

A related point is that theories assuming multiple levels of lexical representation (e.g., Levelt et al., 1999) can account for these findings while still allowing for strong feedback from some level of lexical representation to lexical semantics. This is possible so long as feedback is blocked from lexical representations that encode lexical frequency (e.g., feedback from frequency-insensitive lemmas is acceptable, so long as feedback from frequency-sensitive lexemes is blocked).

General Discussion

Consideration of data from both major traditions of speech production research reveals support for limited interactivity in the speech production system. These studies provide strong constraints on both the upper and lower bounds of interaction in production theories. First, both chronometric and speech error studies set a lower limit on the degree of cascading activation from L-level to phonological representations. Unselected, semantically activated L-level representations must be able to enhance the activation of their phonological representations. However, there is an upper bound on this influence, as revealed by restrictions on the situations in which chronometric data show cascading activation effects and the lack of indirect disruptions to phonological processing following L-level damage. The second section reviewed evidence concerning feedback from phonological to L-level representations. Such feedback must be strong enough to facilitate the activation of mixed neighbors prior to L-level selection. It must also be able to combine with weak cascading activation to produce lexical effects within phonological processing. As with cascading activation, evidence was reviewed providing upper bounds on this type of interaction; prior to L-level selection, formal neighbors must not be significantly activated. Finally, existing evidence suggests that prior to selection lexical semantic representations are not influenced by feedback from lower levels in the speech production system. Such feedback is either absent or extremely weak.

Implications for theories of speech production

The results are most consistent with the Restricted Interaction Account (RIA) proposed by Rapp and Goldrick (2000, 2004; Goldrick & Rapp, 2002). This theory claims that: limited cascading activation is present throughout the production system; limited feedback exists between the phonological and L-levels; and feedback is absent between L-level and semantic representations. These previous papers have motivated this theory primarily through discussion of speech error data; as shown here, it is consistent with the existing chronometric data as well.

Although RIA provides the most explicit claims about restrictions on interaction, these claims resonate with properties of many other proposals in the literature. Theories with cascade and feedback often assume that the influence of these interactive mechanisms is restricted. Harley (1993, 1995) presented results from a simulation with cascading activation throughout the production system and feedback from phonological representations to the L-level. His results show that the influence of cascade and feedback can be minimized (or even eliminated) by the presence of inhibition and activation decay. Dell and O'Seaghdha (1991, 1992) pointed out that many spreading activation accounts are only locally interactive; these theories restrict interaction to adjacent representational levels and enforce rather strict selection points to increase the discreteness of spoken production processing[15]. Coming from the chronometric tradition, discrete theories often assume that cascade and feedback effects can appear in limited situations (see above for discussion of issues associated with these particular proposals). As noted above, Levelt et al. (1999) assume that cascading activation from L-level to phonological representations can occur in cases of near synonymy. Roelofs (2004a, b) argues that feedback from phonological to L-levels can (optionally) occur, mediated by perceptual monitoring.

Directions for future research

In short, empirical data support the presence of limited interaction. Not surprisingly, then, all current theoretical proposals incorporate various forms of limited interaction. This suggests that the presence versus complete absence of interaction in the production system is no longer an issue under debate. If this is indeed the case, the empirical focus of research on interaction should focus on finer-grained questions on the manner and degree of interaction between speech production processes. Four such avenues for further investigation are sketched below.

Differences in the strength of interactive effects in chronometric and speech error data. One question that comes naturally out of the review above is why errors are more likely to shown an influence of interaction than reaction times. As noted above, chronometric and speech error data disagree on the extent to which activation cascades from the L-level; specifically, chronometric data fail to reveal the activation of non-synonymic semantically related words. Two possible accounts of this difference have been proposed.

The first of these focuses on methodological differences between the two types of studies. As noted above, chronometric studies reporting null effects often attempt to observe mediated effects, while speech error studies often focus on cases where activation largely converges with that of the target (Dell & O'Seaghdha, 1991). The presence versus absence of convergent activation may be sufficient to account for these differences.

A second possibility is that when errors occur, the dynamics and organization of processing are shifted into a "deviant" processing state that allows for greater interaction (for discussion of errors in reading aloud as the consequence of shifts in processing dynamics, see Farrar & VanOrden, 2001). A version of this hypothesis is Levelt et al.'s (1999: 35; see also Roelofs, 2004a) claim that cascading activation effects with near-synonyms are due to the erroneous selection of two L-level representations. In this proposal, interactive effects are due to extraordinary processing circumstances and do not reflect the normal state of processing. Note that this hypothesis is not identical to the thoroughly-debunked claim that errors bear no relation to normal language processing (see, e.g., Fromkin, 1971, for a critique of this proposal). Rather, this hypothesis claims that certain aspects of errors are not indicative of normal processing.

The former hypothesis is more parsimonious in the sense that errors are assumed to be a natural consequence of spoken production processing; special mechanisms are not invoked purely to account for errors. However, it has not gone unchallenged. Levelt et al. (1991b) argued that the empirically observed strength of semantic priming is so great that interactive theories (specifically, Dell & O'Seaghdha, 1991) predict that some amount of priming should be observed in mediated situations. Resolution of this issue requires the development of accounts that simultaneously make fine-grained predictions regarding reaction times and errors. Analysis of such systems will allow a better test of these two proposals.

Interaction: Dynamic or static feature of production processes? The question of whether the processing dynamics shift during error productions relates to the broader question of whether interactivity is itself dynamic. Is the degree of interaction a hard-wired architectural feature of the production system, or does the extent of interaction across representations and processes shift across different processing situations?

For example, Kello, Plaut, & MacWhinney (2000) claimed that the strength of cascading activation to articulation could be modulated by task demands. A similar task-demand account might explain some of the results reviewed above. For example, the absence of lexical frequency effects in "semantic" tasks (e.g., picture-word matching) might be due to the amount of resources they require. These tasks may simply be easier to perform than picture naming. The lower demand on processing resources, not a lack of feedback connections, may account for the absence of interactive effects. Note that it is unlikely that such an account can be offered for all the results reviewed here. For example, patients with lexical semantic vs. L-level deficits were all asked to perform the same task—untimed picture naming—yet their performance yielded contrasting patterns. Future research should more carefully explore how interactive effects can or cannot be modulated by the demands of different tasks.

Architectural accounts of the absence of feedback from L-level to semantic representations. Instead of contrasting task demands, is there instead some functional or architectural feature of the production system that could motivate this difference in interaction? One such account is connectivity distance. As pointed out above, feedback mechanisms may be present, but they may be too weak to exert an influence on lexical semantic selection processes. This is highly plausible for the case of phonological information. It must traverse many links in the production system before it can influence semantic processing; it may come too late or be too weak to influence selection of a lexical concept. It is less clear if this account could explain the lack of lexical frequency effects at the semantic level; in the three-level system discussed here these effects arise at a representational level adjacent to semantics. However, if there are instead two levels of lexical representation, and frequency effects are associated with the level more distant from semantics (e.g., a lexeme; Levelt et al., 1999), connectivity distance may provide a ready account of the lack of frequency effects.

A second proposal is that the lack of interaction stems from a representational distinction between semantic representations, on one hand, and L-level and phonological representations, on the other. If we assume (following Caramazza, 1997) that L-level representations are modality specific, the latter two levels of representation can be easily distinguished from amodal semantic representations. The lack of feedback can then be attributed to a more independent design principle—stronger interaction is restricted to similar representational types (Rapp & Goldrick, 2000).

Again, resolving these two proposals will most probably rely on the development of more detailed theories and simulations of speech production processing. Following the first proposal, is it even possible to set up such an interactive system? That is, can a system with a constant strength of cascade and feedback rely solely on connectivity distance to explain the absence of effects at the lexical semantic level? Future simulation studies of interactive theories are clearly required.

Mechanisms mediating feedback. This review explicitly avoided the issue of whether feedback was mediated by mechanisms and processes internal to speech production, or if speech perception processes may also play a role in mediating such effects. Roelofs (2004a, b) claimed that feedback mediated by speech perception and monitoring systems provided the best account of the data supporting feedback. Rapp and Goldrick (2004) discussed some of the difficulties associated with this particular proposal. Since these papers, new data from Hartsuiker et al. (2005) provides further support for the claim that theories attributing all feedback effects to monitoring processes are clearly insufficient. Future theoretical and empirical work should continue to address the relative contributions of production-internal, perceptual and monitoring mechanisms in supporting feedback (see also Hartsuiker, 2006, for further discussion).

Conclusions

Speech production theories have tended to focus on a single type of evidence—based around reaction times or speech errors—to inform theory development. It has therefore been difficult to determine if disagreements between theories developed within each empirical tradition were merely due to differences in the type of evidence driving each theory. Consideration of both types of data shows that this is partially true—interactive effects appear to be more easily observed in speech errors than chronometric data. Recent work, however, reveals that interactive effects are not peculiar to speech error data; furthermore, discrete effects are not peculiar to chronometric data. Both types of evidence argue for the presence of interactivity; both types of evidence point to significant constraints on interactivity. These findings reveal the limitations of the "highly discrete/highly interactive" dichotomy, and challenge us to explore the continuum of possibilities offered by limited interaction.

Footnotes

[1] More recent studies (see Indefrey & Levelt, 2004, for a recent review) have utilized neuroimaging techniques to examine word production. This review, however, is limited to speech error and chronometric data (for discussion of neuroimaging evidence specifically concerning the interaction between speech production processes, see deZubicaray, Wilson, McMahon, & Muthiah, 2001).

[2] Assuming a single level of lexical representation leads one to attribute the effects of homophonic primes to overlapping phonological, not lexical, representations. It is generally assumed that homophones have distinct lexical representations at some level in the processing system (Bonin & Fayol, 2002; Caramazza, Bi, Costa, & Miozzo, 2004; Caramazza, Costa, Miozzo, & Bi, 2001; Dell, 1990; Jescheniak & Levelt, 1994; Jescheniak, Meyer, & Levelt, 2003; Levelt et al., 1999). Assuming a single lexical level means that homophones share no lexical representation; their only overlap is at the phonological level.

[3] It is possible that errors such as blends reflect the simultaneous selection of multiple representations (e.g., Levelt et al., 1999).

[4] As noted by Vigliocco and Hartsuiker (2002: footnote 8), del Viso et al's claim that their data do not show a mixed error bias is debatable. del Viso et al. did find a significant mixed error effect in their overall corpus. However, the effect was carried by a small subset of highly similar items; when these were excluded no significant mixed error effect was found. Although this subset analysis is justifiable (the mixed error effect should not be carried by only a few items), it may in fact be too conservative (obscuring a true mixed error effect).

[5] To guard against strategic effects, Ferreira & Griffin discouraged participants from covertly completing each sentence by including a condition where pictures were never an appropriate sentence completion (Experiment 3). Effects of similar magnitude were observed here, arguing against a strategic basis for the findings.

[6] The "Cascading Feedforward Account" simulation of Rapp & Goldrick (2000) was used. With the exception of the L-level jolt (which was varied as described above), parameters were set at the default levels described in that paper.

[7] The "Restricted Interaction Account" simulation of Rapp & Goldrick (2000) was used. Parameter settings that deviated from default settings were: feedforward connections were set to 0.05; L-level jolt was set to 5.0; and feedback connections were varied as specified above.

[8] Limited feedback predicts that some formal errors should be produced with sufficiently high levels of L-level damage. The pattern of performance exhibited by EA (Shelton & Weinrich, 1997) is consistent with this prediction.

[9] Related mechanisms can produce similar effects within attractor-based systems (Plaut & Shallice,1993).

[10] Of course, the phoneme will enter into feedback loops with some L-level units (e.g., <ZIP>); however, these units are more distant from the target and thus will receive less feedback support from the target (e.g., <HAT> will receive support from the /ae/ and /t/ of target "cat").

[11] Note that lexical frequency is highly correlated with other lexical properties such as age of acquisition (for recent reviews, see Bonin, Barry, Méot, & Chalard, 2004; Zevin & Seidenberg, 2002).

[12] This feedback would have similar consequences in systems with unitary lexical semantic representations and those with distributed feature representations. In the former, the activation

of a single unit would be enhanced, making it a stronger competitor for selection at the lexical semantic level. In the latter, a set of semantic features would have their activation enhanced, allowing their corresponding lexical concept to compete more strongly for selection.

[13] Note that Bonin and Fayol (2002) found a reverse frequency effect in a picture categorization task (e.g., pictures with low frequency names were more quickly categorized than those with high frequency names).

[14] The "High Interaction Account" simulation of Rapp & Goldrick (2000) was used. Parameter settings that deviated from default settings were: feedforward and feedback connections were set to 0.05 (except when varied as specified above); and the L-level jolt was set to 5.0.

[15] Note that these restrictions on interaction are not a priori features of production theories. Plaut and Shallice (1993) proposed a theory of (semantically-based) reading without strict selection points. In a similar context, Van Orden and colleagues (Farrar and Van Orden, 2001; VanOrden, op de Haar, & Bosman, 1997) have proposed that interactions can occur between non-adjacent representational levels.

References

Baars, B. J., & Motley, J. T. & MacKay, D. (1975). Output editing for lexical status from artificially elicited slips of the tongue. Journal of Verbal Learning and Verbal Behavior, 14, 382- 391.

Basso, A., Taborelli, A., & Vignolo, L. A. (1978). Dissociated disorders of speaking and writing in aphasia. Journal of Neurology, Neurosurgery and Psychiatry. 41, 556-563.

Berg, T. , & Schade, U. (1992). The role of inhibition in a spreading-activation model of language production I: The psycholinguistic perspective. Journal of Psycholinguistic Research, 21, 405-434.

Best, W. (1996). When racquets are baskets but baskets are biscuits, where do the words come from? A single case study of formal paraphasic errors in aphasia. Cognitive Neuropsychology, 13, 443-480.

Blanken, G. (1990). Formal paraphasias: A single case study. Brain and Language, 38, 534–554.

Blanken, G. (1998). Lexicalisation in speech production: Evidence from form-related word substitutions in aphasia. Cognitive Neuropsychology, 15, 321-360.

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. Journal of Memory and Language, 48, 468-488.

Bloem, I., van den Boogaard, S., & La Heij, W. (2004). Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access. Journal of Memory and Language, 51, 307-323.

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? Journal of Memory and Language, 50, 456-476.

Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. European Journal of Cognitive Psychology, 14, 289-313.

Brédart, S. & Valentine, T. (1992). From Monroe to Moreau: An analysis of face naming errors. Cognition, 45, 187-223.

Burke, D. M., Locantore, J. K., Austin, A. A., & Chae, B. (2004). Cherry pit primes Brad Pitt: Homophone priming effects on young and older adults' production of proper names. Psychological Science, 15, 164-170.

Caramazza, A. (1997). How many levels of processing are there in lexical access? Cognitive Neuropsychology, 14, 177-208.

Caramazza, A., Bi, Y. Costa, A., Miozzo, M. (2004). What determines the speed of lexical access: Homophone or specific-word frequency? A reply to Jescheniak et al. (2003). Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, 278-282.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. Journal of Experimental Psychology: Learning, Memory, and Cognition, 27, 1430-1450.

Caramazza, A. & Hillis, A. E. (1990). Where do semantic errors come from? Cortex, 26, 95-122.

Costa, A., Caramazza, A., Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 1283-1296

Costa, A., Colomé, A., Gómez, O., & Sebastián-Gallés, N. (2003). Another look at cross-language competition in bilingual speech production: Lexical and phonological factors. Bilingualism: Language and Cognition, 6, 167-179.

Costa, A., Mahon, B., Savova, V., & Caramazza, A. (2003). Levels of categorization effect: A novel effect in the picture-word interference paradigm. Language and Cognitive Processes, 18, 205-233.

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25, 318-344.

Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. Psychonomic Bulletin & Review, 10, 111-117.

Damian, M. F., & Martin, R. M. (1999). Semantic and phonological codes interact in single word production. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25, 345–361.

DeCaro, S., & Reeves, A. (2002). The use of word–picture verification to study entry-level object recognition: Further support for view-invariant mechanisms. Memory & Cognition, 30, 811-821.

del Viso, S., Igoa, J. M. & García-Albea, J. E. (1991). On the autonomy of phonological encoding: Evidence from slips of the tongue in Spanish. Journal of Psycholinguistic Research, 20, 161-185.

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

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. Language and Cognitive Processes, 4, 313-349.

Dell, G. S., & Gordon, J. K. (2003). Neighbors in the lexicon: Friends or foes? In N. O. Schiller & A. S. Meyer (Eds.) Phonetics and phonology in language comprehension and production: Differences and similarities. New York: Mouton de Gruyter.

Dell, G. S. & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). Psychological Review, 98, 604–14.

Dell, G. S. & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. Cognition, 42, 287–314.

Dell, G. S., & O'Seaghdha, P. G. (1994). Inhibition in interactive activation models of linguistic selection and sequencing. In D. Dagenbach & T. H. Carr (Eds.) Inhibitory processes in attention, memory, and language (pp. 409-453). San Diego: Academic Press.

Dell, G. S. & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. Journal of Verbal Learning and Verbal Behavior, 20, 611-629.

Dell, G. S., Schwartz, M., Martin, M., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. Psychological Review, 104, 801–838.

de Zubicaray, G. I., Wilson, S. J., McMahon, K. L., & Muthiah, S. (2001). The semantic interference effect in the picture-word paradigm: An event-related fMRI study employing overt responses. Human Brain Mapping, 14, 218-227.

Farrar IV, W. T., & Van Order, G. C. (2001). Errors as multistable response options. Nonlinear Dynamics, Psychology, and Life Sciences, 5, 223-265.

Ferreira, V. S., & Griffin, Z. M. (2003). Phonological influences on lexical (mis-)selection. Psychological Science, 14, 86–90.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. Language, 47, 27-52.

Gagnon, D. A., Schwartz, M. F., Martin, N., Dell, G. S. & Saffran, E. M. (1997). The origins of formal paraphasias in aphasics' picture naming. Brain and Language, 59, 450-472.

Garrett, M. F. (1976). Syntactic processes in sentence production. In R.J. Wales & E. Walker (eds.) New approaches to language mechanisms (pp. 231-255). Amsterdam: North Holland.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.) Language production (vol. I): Speech and talk (pp. 177-220). New York: Academic Press.

Goldrick, M., & Rapp, B. (2002). A restricted interaction account (RIA) of spoken word production: The best of both worlds. Aphasiology, 16, 20-55.

Goldrick, M., & Rapp, B. (in press). Lexical and post-lexical phonological representations in spoken production. Cognition.

Gollan, T. H., & Acenas, L.-A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, 246-269.

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

Harley, T. A. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. Cognitive Science, 8, 191-219.

Harley, T. A. (1993). Phonological activation of semantic competitors during lexical access in speech production. Language and Cognitive Processes, 8, 291-309.

Harley, T. A., (1995). Connectionist models of anomia: A comment on Nickels. Language and Cognitive Processes, 10, 47-58.

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

Harley, T. A., & MacAndrew, S. B. G. (1995). Interactive models of lexicalization: Some constraints from speech error, picture naming, and neuropsychological data. In J. P. Levy, D. Bairaktaris, & J. A. Bullinaria (Eds.) Connectionist models of memory and language (p. 311-331). London: UCL Press.

Harley, T. A., & MacAndrew, S. B. G. (2001). Constraints upon word substitution speech errors. Journal of Psycholinguistic Research, 30, 395-417.

Harstuiker, R. J. (2006). Are speech error patterns affected by a monitoring bias? Language and Cognitive Processes, 21, xxx-xxx.

Hartsuiker, R. J., Corley, M., & Martensen, H. (2005), The lexical bias effect is modulated by context, but the standard monitoring account doesn't fly: Related beply to Baars et al. (1975). Journal of Memory and Language, 52, Pages 58-70.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? Bilingualism: Language and Cognition, 1, 213-229.

Hillis, A. E. (2001). The organization of the lexical system. In B. Rapp (Ed.) The handbook of cognitive neuropsychology: What deficits reveal about the human mind (pp. 185-210). Philadelphia: Psychology Press.

Hillis, A. E., Rapp, B., Romani, C., & Caramazza, A. (1990). Selective impairment of semantics in lexical processing. Cognitive Neuropsychology, 7, 191–243.

Humphreys, G. W., Lloyd-Jones, T. J., & Fias, W. (1995). Semantic interference effects on naming using a postcue procedure: Tapping the links between semantics and phonology with pictures and words. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 961-980.

Humphreys, K. R. (2002). Lexical bias in speech errors. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Igoa, J. M. (1996). The relationship between conceptualization and formulation processes in sentence production: Some evidence from Spanish. In M. Carreiras, J. García-Albea & N. Sebastián-Galles (eds.) Language processing in Spanish (pp. 305-351). Hillsdale, NJ: Lawrence Erlbaum Associates.

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. Cognition, 92, 101-144.

Jescheniak, J. D., Hantsch, A., & Schriefers, H. (2005). Context effects on lexical choice and lexical activation. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31, 905-920.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20, 824–843.

Jescheniak, J. D., Meyer, A. S., & Levelt, W. J. M. (2003). Specific-word frequency is not all that counts in speech production: Comments on Caramazza, Costa, et al. (2001) and new experimental data. Journal of Experimental Psychology: Learning, Memory, and Cognition, 29, 432-438.

Jescheniak, J. D., & Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: Further evidence from the coactivation of near-synonyms. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24, 1256-1274.

Jescheniak, J. D., & Schriefers, H. (2001). Priming effects from phonological related distractors in picture-word interference. The Quarterly Journal of Experimental Psychology, 54A, 371-382.

Kello, C. T., Plaut, D. C., & MacWhinney, B. (2000). The task dependence of staged versus cascaded processing: An empirical and computational study of Stroop interference in speech production. Journal of Experimental Psychology: General, 129, 340–360.

Kulke, F., & Blanken, G. (2001). Phonological and syntactic influences on semantic misnamings in aphasia. Aphasiology, 15, 3-15.

Laine, M., Tikkala, A., & Juhola, M. (1998). Modelling anomia by the discrete two-stage word production architecture. Journal of Neurolinguistics, 11, 275-294.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. Cognition, 14, 41-104.

Levelt, W. J. M. (1989). Speaking: From intention to articulation. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1999). Models of word production. Trends in Cognitive Sciences, 223-232.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. Behavioral and Brain Sciences, 22, 1-75.

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

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991b). Normal and deviant lexical processing: Reply to Dell and O'Seaghdha (1991). Psychological Review, 98, 615-618.

Martin, N., Gagnon, D. A., Schwartz, M. F., Dell, G. S. & Saffran, E. M. (1996). Phonological facilitation of semantic errors in normal and aphasic speakers. Language and Cognitive Processes, 11, 257-282.

Martin, N., Dell, G. S., Saffran, E. M., & Schwartz, M. F. (1994). Origins of paraphasias in deep dysphasia: Testing the consequences of a decay impairment to an interactive spreading activation model of lexical retrieval. Brain and Language, 47, 609-660.

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

Martin, N., Weisberg, R. W. & Saffran, E. M. (1989). Variables influencing the occurence of naming errors: Implications for models of lexical retrieval. Journal of Memory and Language, 28, 462-485.

Melinger, A., & Abdel Rahman, R. (2004). Investigating the interplay between semantic and phonological distractor effects in picture naming. Brain and Language, 90, 213-220.

Meyer, D. E., & Gordon, P. C. (1985). Speech production: Motor programming of phonetic features. Journal of Memory and Language, 24, 3-26.

Miceli, G., Benvegnú, B., Capasso, R., & Caramazza, A. (1997). The independence of phonological and orthographic lexical forms: Evidence from aphasia. Cognitive Neuropsychology, 14, 35-70.

Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in the picture-word interference paradigm. Journal of Experimental Psychology: General, 132, 228-252.

Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. Journal of Experimental Psychology: Learning, Memory, and Cognition, 28, 555-563.

Navarette, E., & Costa, A. (2005). Phonological activation of ignored pictures: Further evidence for a cascade model of lexical access. Journal of Memory and Language, 53, 359-377.

Nickels, L. (1992). The autocue? Self-generated phonemic cues in the treatment of a disorder of reading and naming. Cognitive Neuropsychology, 9, 307-317.

Nickels, L. (1995). Getting it right? Using aphasic naming errors to evaluate theoretical models of spoken word production. Language and Cognitive Processes, 10, 13-45.

Nickels, L. & Howard, D. (1995). Phonological errors in aphasic naming: Comprehension, monitoring, and lexicality. Cortex, 31, 209-237.

Nooteboom, S. G. (2005a). Lexical bias revisited: Detecting, rejecting, and repairing speech errors in inner speech. Speech Communication, 47, 43-58.

Nooteboom, S. G. (2005b). Listening to oneself: Monitoring speech production. In R. Hartsuiker, Y. Bastiaanse, A. Postma, & F. Wijnen (Eds.), Phonological encoding and monitoring in normal and pathological speech (pp. 167-186). Hove, UK: Psychology Press.

Page, M. (2000). Connectionst modelling in psychology: A localist manifesto. Behavioral and Brain Sciences, 23, 443-467.

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

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. Cognitive Neuropsychology, 10, 377-500.

Rapp, B., Benzing, L., & Caramazza, A. (1997). The autonomy of lexical orthography. Cognitive Neuropsychology, 14, 71-104.

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

Rapp, B., & Goldrick, M. (2004). Feedback by any other name is still interactivity: A reply to Roelofs' comment on Rapp & Goldrick (2000). Psychological Review, 111, 573-578.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. Cognition, 64, 249-284.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. Psychological Review, 110, 88-125.

Roelofs, A. (2004a). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: Comment on Rapp and Goldrick (2000). Psychological Review, 111, 561-572.

Roelofs, A. (2004b). Comprehension-based versus production-internal feedback in planning spoken words: A rejoinder to Rapp and Goldrick (2004). Psychological Review, 111, 579-580.

Schade, U. & Berg, T. (1992). The role of inhibition in a spreading-activation model of language production II: The simulation perspective. Journal of Psycholinguistic Research, 21, 435-462.

Schatzman, K. B., & Schiller, N. O. (2004). The word frequency effect in picture naming: Contrasting two hypotheses using homonym pictures. Brain and Language, 90, 160-169.

Shelton, J. R., & Weinrich, M. (1997). Further evidence of a dissociation between output phonological and orthographic lexicons: A case study. Cognitive Neuropsychology, 14, 105-129.

Starreveld, P. (2000). On the interpretation of phonological context effects in word production. Journal of Memory and Language, 42, 497-525.

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 686–698.

Starreveld, P. A., & La Heij, W. (1999). Word substitution errors in the picture-word task. American Journal of Psychology, 112, 521-553.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (ed.) Progress in the psychology of language (Vol. 1, pp. 143-186). Hillsdale, NJ: Erlbaum.

Stemberger, J. P. (2004). Neighbourhood effects on error rates in speech production. Brain and Language, 90, 413-422.

Taylor, J.K., & Burke, D.M. (2002). Asymmetric aging effects on semantic and phonological processes: Naming in the picture-word interference task. Psychology and Aging, 17, 662-676.

Van Orden, G. C., op de Haar, M. A. J., & Bosman, A. M. T. (1997). Complex dynamic systems also predict dissociations, but they do not reduce to autonomous components. Cognitive Neuropsychology, 14, 131-165.

Vigliocco, G., & Harsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. Psychological Bulletin, 128, 442-472.

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

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

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

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. Journal of Memory and Language, 47, 1-29.

Experiment(s)	**Task**	**Semantically activated lexical item:**	**Phonological facilitation observed in:**
Peterson & Savoy (1998)	Cued response: After picture presentation, either name the picture or a visually presented word.	Near-synonym of picture name (e.g., “sofa” for “couch”).	Naming latency of visually presented word phonologically related to near synonym of picture name (e.g., naming of target “soda,” similar to “sofa”).
Cutting & Ferreira (1999); Taylor & Burke (2002)	Picture naming during auditory picture-word interference	Semantic associate of distractor (e.g., “ball,” a formal dance, for distractor “dance”).	Picture naming latency of homophone of semantic associate of distractor (e.g., target “ball,” a toy, is a homophone of “ball,” a formal dance).
Damian & Martin (1999); Starreveld & La Heij (1995, 1999); Taylor & Burke (2002)	Picture naming during auditory or orthographic picture word interference	Mixed (semantic-phonological) neighbor of picture target (e.g., “calf” for target “camel”).	Picture naming latency of target[T1] (Damian & Martin; Starreveld & La Heij, 1995; Taylor & Burke, 2002); target accuracy (Starreveld & La Heij, 1999).
Costa, Caramazza, & Sebastián-Gallés (2000); Gollan & Acenas (2002)	Picture naming	Cognate (phonologically similar translation) of target (e.g., for target Spanish “gato” ‘cat,’ Catalan cognate “gat”).	Picture naming latency of target (Costa et al.); target accuracy (Gollan & Acenas)
Morsella & Miozzo (2002); Navarrete & Costa (2005)	Picture naming during picture-picture interference	Phonologically related non-target picture (e.g., “bell” for target “bed”).	Picture naming latency of target[T2]

Table 1. Studies showing phonological facilitation by semantically-activated non-target words.

Experiment(s)	**Task**	**Semantically activated lexical item:**	**Target phonologically inhibited by:**
Jescheniak & Schriefers (1998)	Picture naming during auditory picture-word interference	Phonologically dissimilar near-synonym of picture name (e.g., "sofa" for "couch").	Distractor, dissimilar to target, but phonologically related to near-synonym of picture name (e.g., "soda", similar to "sofa" but not target "couch").
Hermans, Bongaerts, De Bot, & Schreuder (1998); Costa, Colomé, Gómez & Sebastián-Gallés, (2003)	Picture naming during auditory picture-word interference	Non-cognate translation equivalent (e.g., Dutch "berm" for English target "mountain").	Distractor, dissimilar to target, but phonologically related to non-cognate translation equivalent (e.g., "bench", similar to "berm" but not target "mountain").[T3]
Jescheniak, Hantsch, & Schriefers (2005)	Picture naming during auditory picture-word interference	Phonologically dissimilar basic or subordinate level label for picture (in subordinate or basic level naming context; e.g., "fish" in context where "shark" is appropriate).	Distractor, dissimilar to target, but phonologically related to basic or subordinate level label (e.g., "finger," similar to "fish" but not target "shark").

Table 2. Studies showing phonological inhibition by semantically-activated non-target words.

Footnotes

[T1] Typically, these findings have been interpreted as reflecting modulation of the activation of L-level representations, not phonological representations (e.g., Damian & Martin, 1999; Roelofs, 2004a; but see Vigliocco & Hartsuiker, 2002, for a cascading activation account of these results). This interpretation is based on the observation that pure phonological distractors in these studies do not exhibit facilitation at the SOAs (stimulus-onset asynchronies) at which interactive effects are observed (Roelofs, 2004a). However, phonological facilitation effects have been observed at a wide range of SOAs (Jescheniak & Schriefers, 2001; Starreveld, 2000). Second, the facilitation of formally related neighbors may simply be too weak to be observed.

[T2] Other studies have claimed that non-target pictures are not activated post-semantically. Interference effects produced by semantically related distractor words in picture naming (Damian & Bowers, 2003; Humphreys, Lloyd-Jones, & Fias, 1995) or word translation (Bloem & La Heij, 2003; Bloem, van den Boogaard, & La Heij, 2004) are absent when pictorial distractors are used. If semantic interference effects arise during lexical selection processes, the absence of interference here suggests that pictorial input does not automatically activate post-semantic representations. However, recent work has challenged the post-semantic account of semantic interference (see Costa, Mahon, Savova, & Caramazza, 2003, for discussion), undermining these claims.

[T3] Note that Hermans et al. (1998) interpreted this inhibition as arising at the L-level.

Figure Captions

Figure 1. Schematic depictions of four theoretical positions regarding interaction within the speech production system. The first set of units denotes lexical semantic representations (amodal representations of the meaning of lexicalized concepts). The second denote L-level units (word representations mediating the mapping between the other two representational levels). The third set denotes phonological representations (sub-lexical representations of word form). Note that each representational level is only partially depicted (e.g., onset /k/ and coda /t/ from “cat” are not shown). To show the consequences of different types of interaction on the activation of each unit, as well as the flow of activation between units, semantic activation driven by the target is depicted with long dashes; phonological activation driven by the target is depicted with square dots.

A. Theory with no cascading activation or feedback. Multiple semantic competitors are activated at the L-level, but only the target activates its phonological representation.

B. Theory with cascading activation from the L-level to phonological representations. Semantic competitors activate their phonological representations.

C. Theory with feedback from phonological representations to the L-level. Phonological neighbors of the target are activated at the L-level; through cascading activation, these phonological neighbors activate their phonological representations.

D. Theory with feedback from the L-level to semantics. Activation differences based on L-level and/or phonological distinctions are transmitted back to semantic representations.

Figure 2. Indirect disruption to phonological processing due to decreases in L-level selection strength (jolt size). Results show percentage of nonword responses over 10,000 simulated naming attempts in a system with cascading activation. L-level noise was fixed at 0.7.

Figure 3. Production of formal errors due to increases in feedback strength. Results show percentage of formal related errors selected at the L-level over 10,000 simulated naming attempts in a system with cascading activation and feedback from the phoneme to L-level. L-level noise was fixed at 0.7.

Figure 4. Indirect disruption to lexical semantic processing due to increases in L-level to lexical semantic feedback strength. Results show percentage of errors at lexical semantic selection over 10,000 simulated naming attempts in a system with both cascading activation and feedback between all levels in the production system. L-level noise was fixed at 0.7.

Figure 1A

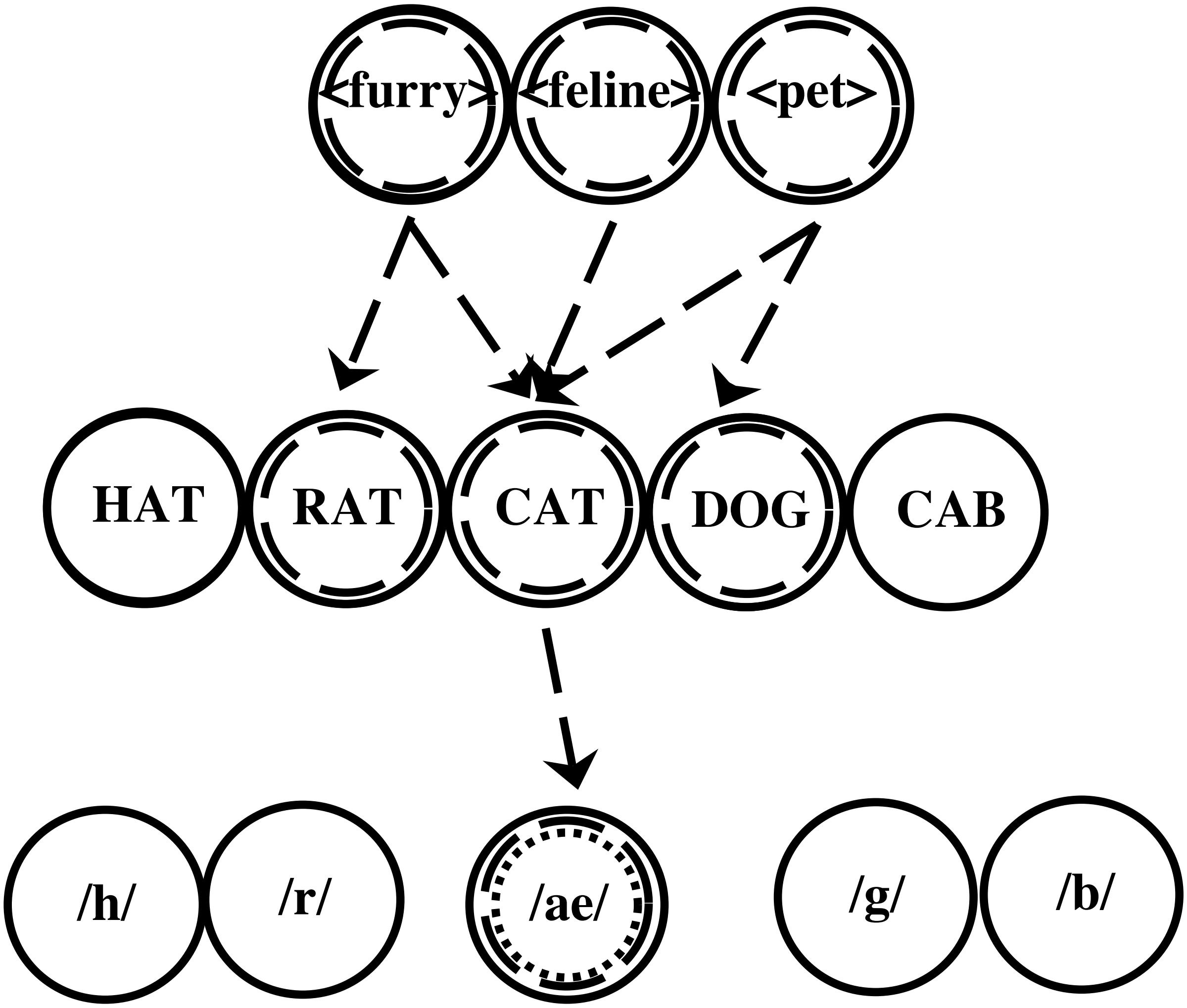


Figure 1B

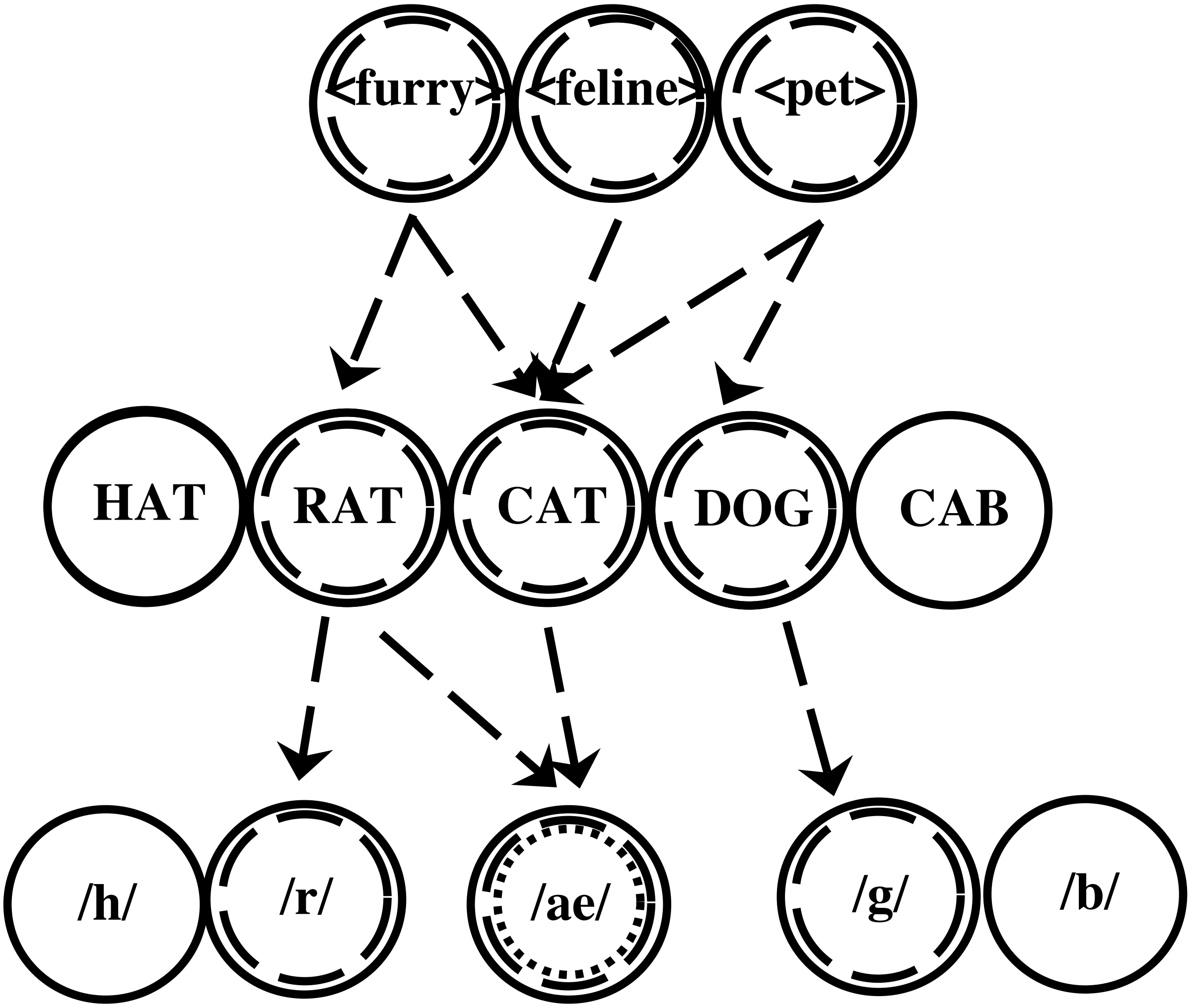


Figure 1C

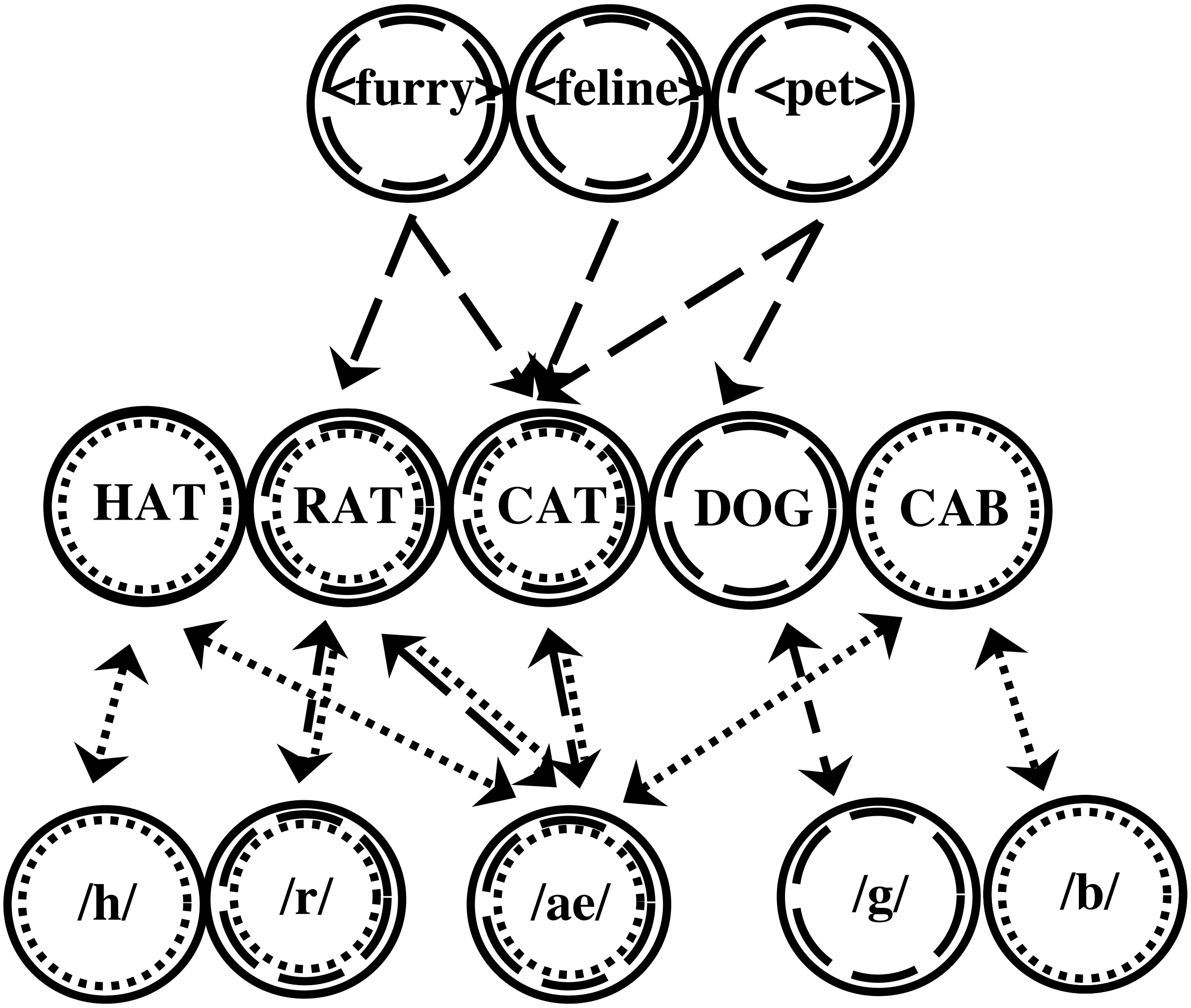


Figure 1D

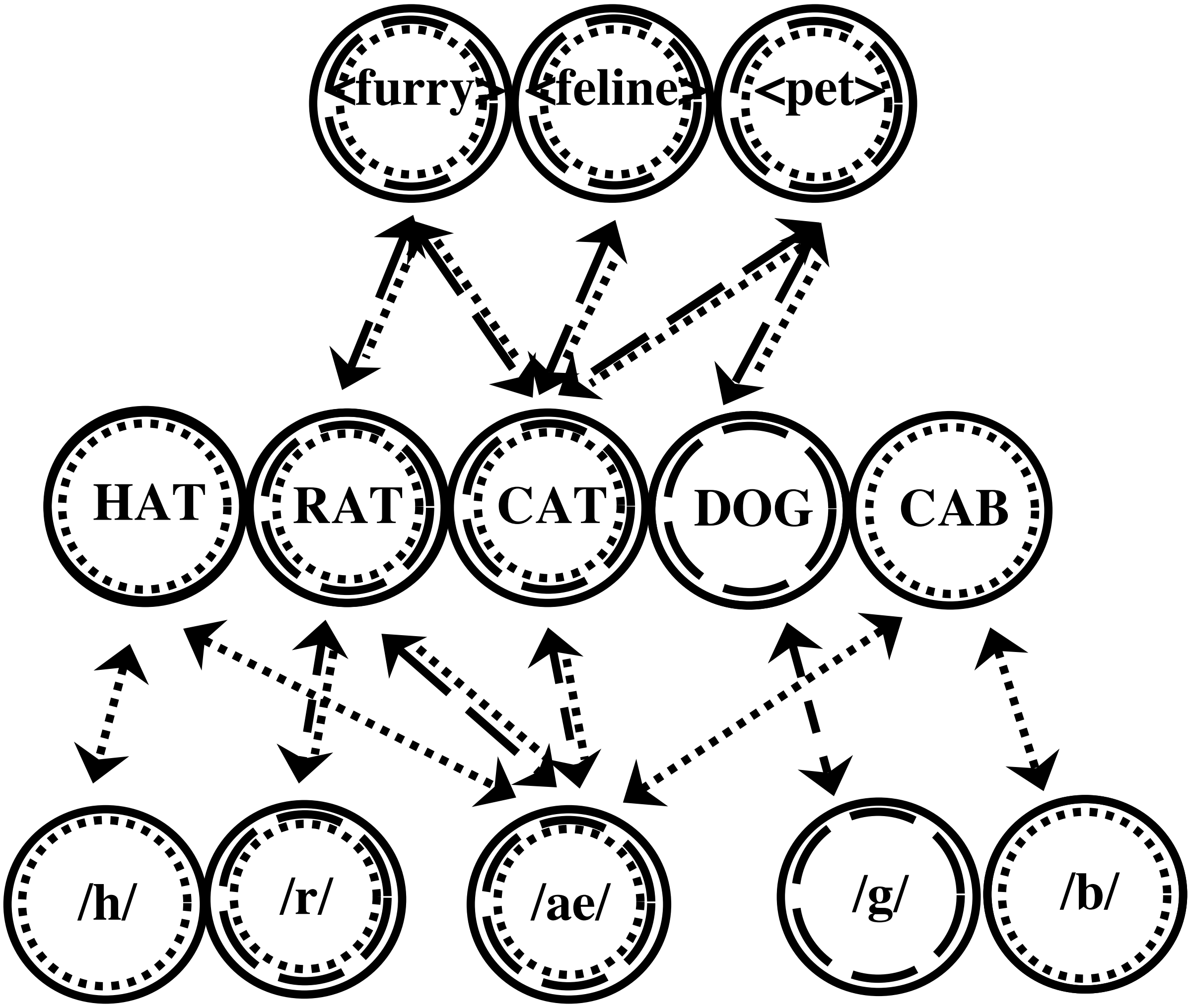


Figure 2

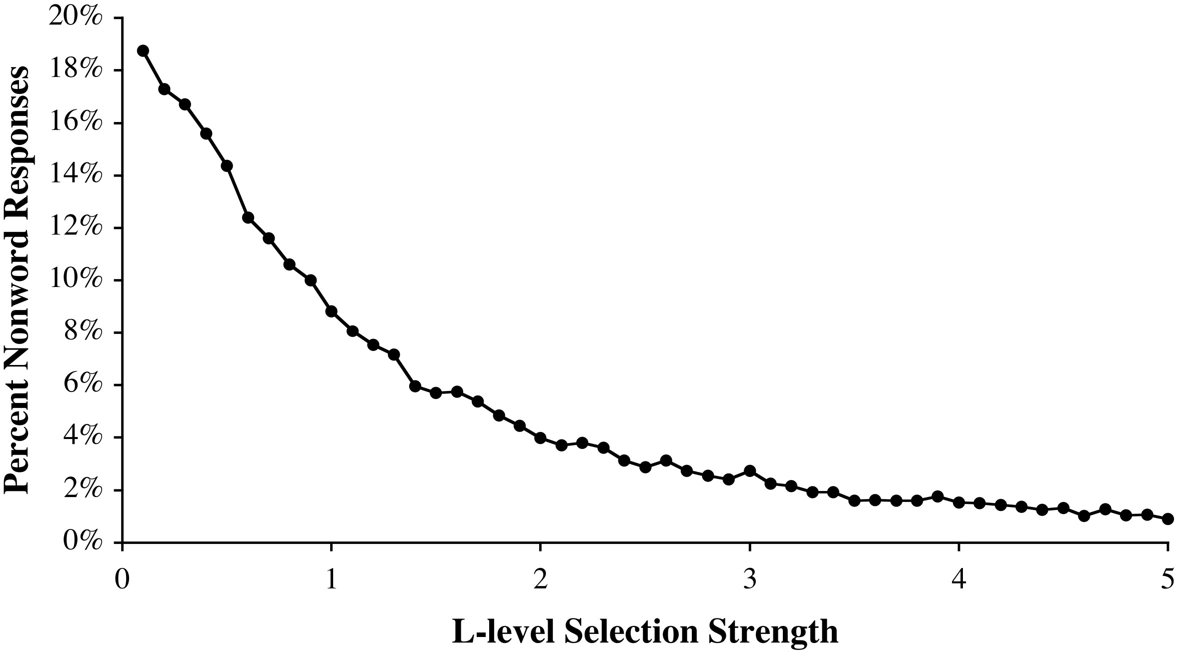

Percent Nonword Responses
20%
18%
16%
14%
12%
10%
8%
6%
4%
2%
0%
0
1
2
3
4
5
L-level Selection Strength

Figure 3

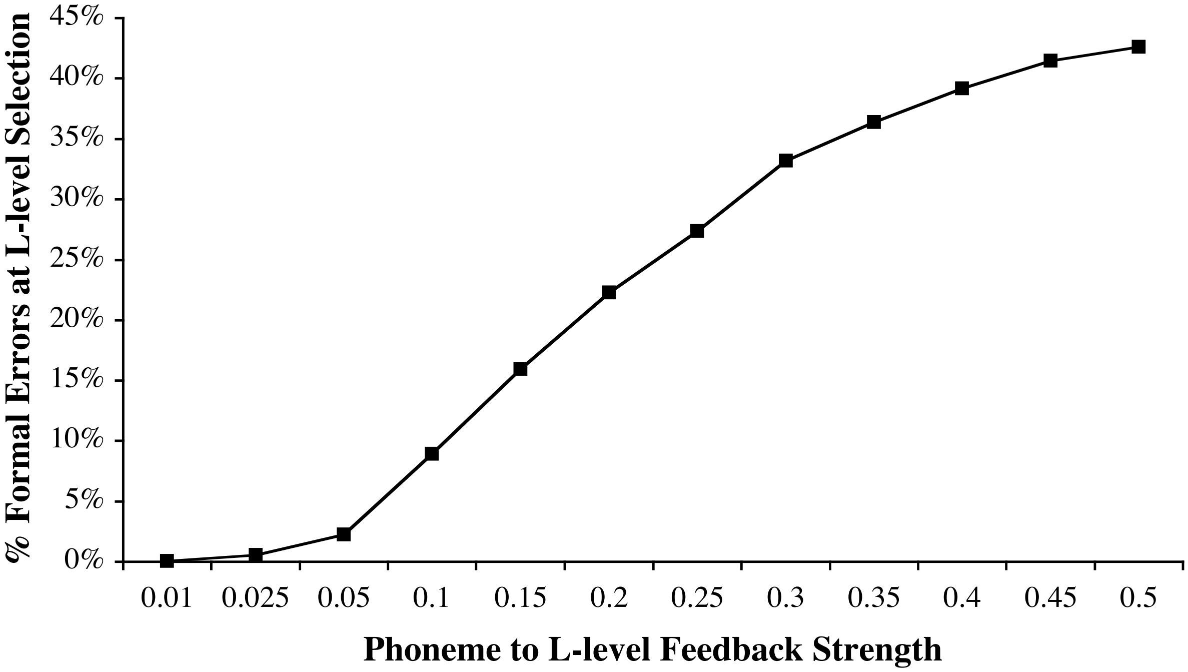
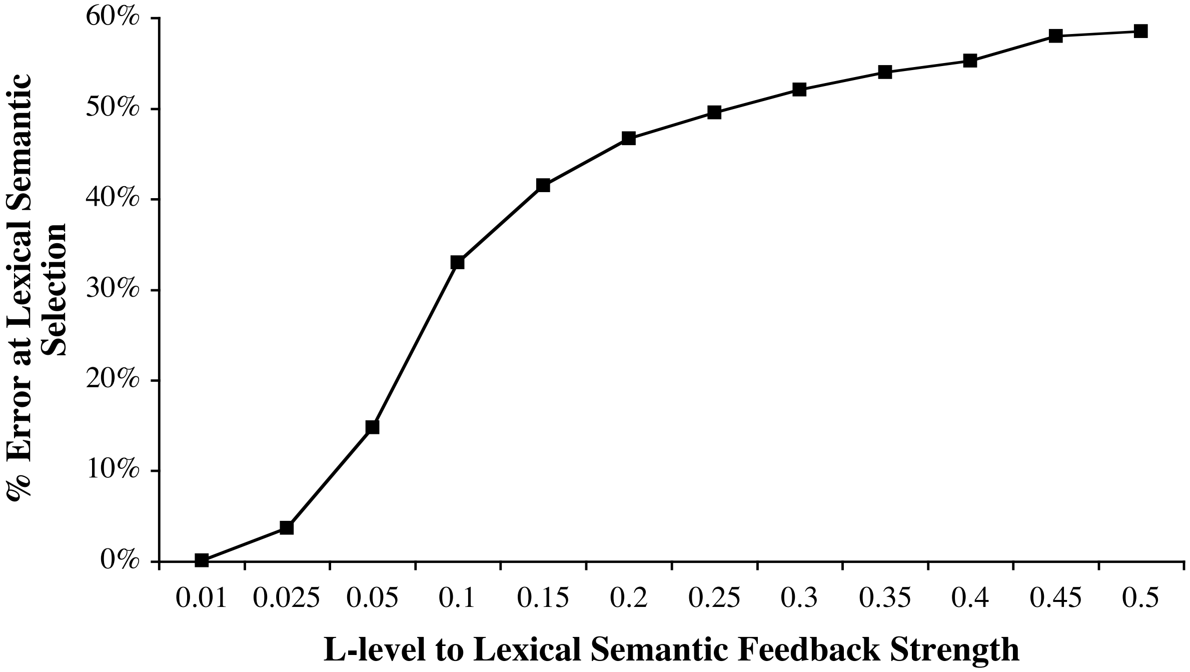

% Formal Errors at L-level Selection
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
0.01
0.025
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
Phoneme to L-level Feedback Strength

Figure 4


60%
50%
40%
30%
20%
10%
0%
% Error at Lexical Semantic Selection
0.01
0.025
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
L-level to Lexical Semantic Feedback Strength