Levett et al.'s is undoubtedly the most complete and sophisticated theory of lexical access in production. It offers a precise account of diverse data and situates that account in an original metatragic framework. We particularly like the emphasis on the natural rift between sound and meaning in the theory, and the way that the rift is associated with frequency effects, the tip-of-the- tongue phenomenon, and the grammatical features of words. Although we question Levett et al.'s claim that there is no interaction across the rift, we agree that the arbitrary mapping between words, semantic (and syntactic) properties and their phonological forms has profound effects on the production system.

We disagree with Levett et al.'s view of binding-by-timing, however. They are rightly impressed that people can name a picture accurately even when they hear a distractor word. (For example, in Cutting & Ferrera, in press, the distractor was erroneously spoken instead of the picture's name: a monkey re-activated instead of the picture name, a mere 14 ms after its presentation.) Binding-by-checking teachers us that the system does not select the distractor instead of the picture's name. In binding-by-timing theories (e.g., Dell 1986), what is selected is determined by activation. Retrieval mechanisms must make sure that the right things are activated at the right time. Thus a binding-by-timing theory would have to claim that the distractors are activated enough to slow things down, but not activated enough to generate more than a handful of errors. This seems implausible to Levett et al.

We think Levett et al. are confusing the need for attentional control with a need for binding-by-checking over binding-by-timing. The picture-word interference experiments are a kind of string task, where a response based on relevant information must be executed while the irrelevant information is ignored. The string literature is vast, and the mechanisms proposed for account for the effect of distracting information on the target are numerous, but at least some of these mechanisms involve control of the activation of distracting information through attention (e.g., Cohen et al. 1999). Thus, in these theories, binding is ultimately expressed in terms of activation, rather than binding-by-checking. In fact, a model of picture-word interference without binding-by-checking can be specified so that the semantic distractor's activation cannot slow things down without inducing errors (Cutting & Ferrera, in press). So we are not persuaded that the way word error rate in these tasks favors binding-by-checking over binding-by-timing.

The timing of retrieval as a determinant of all sorts of things in production. According to Levett et al., binding-by-checking is needed to account for why speakers don't say "kings erect pages" instead of "pages erect kings" when the timing of "kings" happens to be premature. However, much evidence suggests that timing is very influential in assignment of phrasal grammatical functions. Factors that cause a word to be retrieved more rapidly (e.g., semantic priming) also cause the corresponding phrase to be bound to an earlier sentence position (Bock 1982, 1986). In these cases, though, timing affects relative placement without causing errors because of syntactic flexibility (Petrita 1996). Rather than the activation of "kings" leading to an error like "kings erect pages," it can lead to an alternate structure such as "kings are erected by pages." In fact, this highlights how binding-by-timing, which is a principle of recombination, can account succinctly for both errorless order of mention effects and word exchange errors. These kinds of generalizations are not as neatly captured by binding-by-checking.

Levett et al. make a cogent observation when they identify speech errors with binding failures. However, their binding-by-checking mechanism leads them to an unusual (and we think, erroneous) account of speech errors. Their model does not naturally produce phonological exchanges, such as "sed rock" for "red rock." Our model, based on binding-by-timing, highlights how binding-by-timing, coupled with a principle of recombination, can account succinctly for both errorless order of mention effects and word exchange errors. Their model therefore appears to account for all types of errors that their model would fail to account for.

In summary, we share Levett et al.'s account of lexical access in speech production. We think that their model is a more powerful and more complete account of lexical access in speech production. We think that their model is a more powerful and more complete account of lexical access in speech production.
are superfluous for modulists but not for interactionists. Levelt's working hypothesis is that there is a strictly verbal, feed-forward, semantically holistic, localist model, with morphological decomposition and no inhibition is considerably simpler to describe, comprehend, and test than a distributed, interactive, semantically holistic, distributed model with inhibition and without morphological composition (see also Jacobs & Grainger, 1994, for a similar claim applied to visual word recognition). But since every model is at best an approximation (and a fortiori, so are simple models such as the one defended by Levelt et al.), it is hardly surprising that a more complex set of complementary models is generally needed to describe lexical access in speech production more accurately.

I will consider two features that are not needed in Levelt et al.'s model, although experimental evidence favor their inclusion in a more complete model of speech production. Consider first the 'holistic' feature. Levelt et al. did not include any inhibition, but their model is unable to explain the semantic inhibition effects in picture naming observed by Wheeler and Monsell (1994) who reported the astonishing result that the lexicalization process during picture naming is inhibited when a word likely to be a competitor has been primed by a recent production. Naming a pictured object (e.g., SHARK) was retarded when a competing word (e.g., WHALE) had been recently elicited by a definition, Borg and Schade (1992; Schade & Berg 1992; see also Dell & O'Seaghdha 1994) have summarized other empirical evidence in favor of a model that includes both excitation and inhibitory mechanisms. These data are hard to reconcile with the purely activation-based model of Levelt et al.

Consider now the 'interactivity' feature. A thorough examination of the cognitive neuropsychology literature reveals evidence for the coexistence of interactive processes between lexical selection and phonological encoding (Dell et al. 1997; Laine & Martin 1996; Martin et al. 1996). For example, Laine and Martin (1996) studied an amnesic patient who suffered from partial functional disconnection between lexical-semantic and lexical-phonological levels. Systematic manipulation of semantic and phonological relatedness between the to-be-named targets indicated that this patient's naming response patterns were sensitive to both types of lexical relatedness. Levelt et al.'s discrete theory of lexical access is unable to explain the observed effects of semantic and phonological relatedness. However, these results are consistent with an interactive activation model.

In sum, there is a great deal to disagree with in the target article. The moral of this story is that Oorschot's razor should not be wielded blindly.

Acknowledgment
I am grateful to Art Jacobs for his insightful comments.

Prosy and word production
Fernanda Ferreira
Department of Psychology, Michigan State University, East Lansing, MI 48824-1117, fernanda@eyebob.msu.edu

Abstract: Any complete theory of lexical access in production must address how words are produced in prosodic contexts. Levelt, Roelofs & Meyer make some progress on this point, for example, they discuss the 'interactivity' and 'holistic' contributions to the present work demonstrating that word articulation takes into account overall prosodic context. This review supports Levelt et al.'s hypothesis separated between metrical and segmental information.

Levelt, Roelofs & Meyer have done an enormous service to the field of psycholinguistics by articulating an explicit theory of lexical access in production. The model is motivated by almost the entire range of relevant empirical data: speech errors, patterns of reaction times in distracto r tasks, and so on. My goal in this commentary will be to amplify their main points by bringing in issues concerning word production in multworld utterances.

As Levelt et al. note, words are produced in phonological contexts. In Ferreira (1993), I examined how prosodic constituents created from syntactic structure might influence the characteristics of spoken words, especially their durations. In two separate experiments I asked speakers to perform the following task: They were to read a sentence on a computer monitor, commit the sentence to immediate memory, and produce the sentence upon receipt of a cue (the question "What happened?"). Two factors were varied in each experiment: the segmental context of a particular word within the sentence and the prosodic context in which that word occurred. The dependent measures were the duration of the critical word, the duration of the following pause, and the sum of those two durations (total duration).

In the first experiment, the critical word was either phrase-medial or phrase-final and had either a short or long intrinsic duration. Thus, the following four conditions were tested: the chauffeur thought he could stop (short duration) / drive (long duration) the car (phrase-medial position). Even though the chauffeur thought he could stop / drive, the passengers were worried (phrase-final condition).

The theory I proposed assumed that segmental context and position within a prosodic constituent would both influence duration, but independently. The results supported the prediction: word durations were longer for long words and for words in the phrase-final environments, and the two factors were additive. More importantly, total duration was affected only by position, indicating that pause time was used to compensate for intrinsic duration to allow the prosodic interval specified by the prosodic constituent structure to be filled out. I concluded that when words are produced in sentences, the production system assigns an abstract duration to each word based on its placement in a prosodic constituent structure. Words at the ends of prosodic constituents would be assigned longer durations than those in other locations (Selkirk 1984), word lengthening and pausing are used to fill out the interval. In addition, I argued that when the intervals are created the production system does not yet know about the segmental context of the words that will occupy them.

The second experiment used the same logic but varied prosodic constituency in a different way: the critical word was either spoken with constrastive prominence or without. The conditions are illustrated as follows:

The cat (short duration) / mouse (long duration) crossed the bony street.

The CAT / MOUSE crossed the bony street.

The discourse context preceding each item established the need for contrastive prominence. Again, it was expected that this prosodic variable and segmental context would both affect durations, but independently. The prediction was again confirmed, and furthermore the pattern of data was identical to that found in the experiment manipulating location within prosodic constituents: total duration were virtually identical in the two segmental contexts but longer when the word was spoken with contrastive prominence. This finding reinforces the point that abstract intervals are specified for words based on the properties of the prosodic constituent in which they occur. A word with contrastive prominence would be assigned a longer interval before the production system had any knowledge of the segmental context of the word that would end up occupying that interval. When the speaker actually produces the sentence, he has to assign a duration that maintains the size of the interval, thus, if the words segments give it a short intrinsic duration, a longer pause will be necessary.

These experiments make two relevant points. First, they sup-

BEHAVIORAL AND BRAIN SCIENCES (1999) 22:1 43