How do speakers choose a word for production? One general idea is that they accumulate evidence until one word emerges as an acceptable option. According to this noncompetitive approach, the speed of lexical selection should depend on how strongly the strongest word is activated, independent of any alternatives. In this case, the selection process may be modeled as activation toward a simple threshold, whereby the first candidate to reach that threshold will be selected via a winner-take-all mechanism (e.g., Oppenheim, Dell, & Schwartz, 2010, Simulation 6). A competitive extension takes this idea further, suggesting that speakers accumulate evidence until one word emerges as clearly better than any alternative, for instance by surpassing a relative threshold (Roelofs, 1992, 2018). Although this competitive extension attracted unquestioning support for several decades, even serving as the basis of one of the most prominent theories of language production (Levelt, Roelofs, & Meyer, 1999), its necessity has more recently become the subject of robust debate on both empirical and computational grounds. As a step toward resolving this debate, Nozari and Hepner (2018) suggest that their hypothesized conflict monitoring mechanism (Nozari, Dell, & Schwartz, 2011) could provide a basis for assessing and possibly resolving task-incompatible conflict in lexical selection, essentially by scaling a relative threshold according to some function of baseline conflict and task demands.

Perhaps the clearest evidence that strongly activated alternatives can delay lexical selection comes from picture-word interference (Glaser & Glaser, 1989; Schriefers, Meyer, & Levelt, 1990; Starreveld & Heij, 1995), a paradigm in which participants are directed to name pictures using pre-specified names (e.g., “dog”) while suppressing responses to other stimuli (e.g., the visually superimposed name of another item in the response set, “cat”). When the picture is semantically related to the distractor, correct productions of its intended name are typically slower than when it is not. Delays are typically assumed to reflect competition during lexical selection, where activation from the distractor somehow combines with activation from the normal retrieval process, making it harder for the target’s activation to surpass the distractor’s. However, because the interference is less consistent than one might expect (e.g., Miozzo & Caramazza, 2003), and because the experimental paradigm is rather complex and contrived, alternative explanations have proliferated (e.g., Del-l’Acqua, Job, Peressotti, & Pascali, 2007; Dhooge & Hartsuiker, 2010 et passim; Mahon, Costa, Peterson, Vargas, & Caramazza, 2007 et passim; Navarrete & Mahon, 2013 et passim) including the idea that such behavioral results may simply reflect ad hoc monitoring processes.

With such mechanisms in dispute, converging evidence from simpler paradigms, with less obvious manipulations, becomes more important. Cumulative semantic interference is a behavioral effect where naming a picture of a dog as “dog” makes speakers persistently slower and more error-prone when subsequently attempting to name a picture of a cat as “cat”. This interference occurs even in simple picture naming, so it is tempting to conclude that it provides important converging evidence for the competitive extension (e.g., Howard, Nickels, Coltheart, & Cole-Virtue, 2006). However, after demonstrating that a simple model of lexical retrieval and incremental
learning could account for major behavioral manifestations of this interference, we (Oppenheim et al., 2010, Simulations 5–6) used it to assess consequences of three possible selection rules: (1) a fully noncompetitive rule (selection time depends on the strongest word’s activation), (2) a somewhat competitive rule (selection time depends on the strongest word’s activation, relative to the mean of all alternatives), and (3) a strongly competitive rule (selection time is a function of strongest word’s activation, relative to the second strongest; Nozari and colleagues describe this quantity as *conflict*). Remarkably, none of the simulation results required the competitive extension. Although researchers have since claimed novel empirical findings as evidence for competitive selection (e.g., Belke, 2013), in a recent integrated model of picture-word interference and cumulative semantic interference that hinged on competitive selection, Roelofs’s (2018) major empirical support for the feature still came from picture-word interference.

If we assume that picture-word interference patterns reveal core production processes, then there is an important challenge in reconciling data that seem to support competitive lexical selection with those that seem to challenge it. One approach is to assume that apparently discordant findings reveal a kind of flexibility within a single coherent system, characterizing observed variation as a predictable consequence of certain moderating factors. Speed/accuracy tradeoffs are well-documented throughout experimental psychology, and although debates over lexical selection mechanisms typically hinge on naming latency effects—with the implicit assumption that words come out as soon as they can—speakers can usually speak sooner when required (e.g., Dell, 1986; Vitkovitch & Humphreys, 1991). Thus, recognizing goal-driven flexibility in selection criteria (Nozari & Hepner, 2018) offers a path toward reconciling models that explain error patterns by assuming arbitrary selection times (Dell, 1986) with those that explain response times by assuming fixed thresholds (Levelt et al., 1999; Oppenheim et al., 2010, also addressing errors).

Defining such dimensions of flexibility is a natural next step in incrementally understanding how language production normally works. Much like statistical model building first defines obvious main effects, then interactions, and so on, while production research of the 1950s–90s was primarily concerned with characterizing mature systems (and separately studying acquisition), more recent work considers language as a continually learning system (e.g., Chang, Dell, & Bock, 2006; Dell, Reed, Adams, & Meyer, 2000; Oppenheim, 2018; Oppenheim et al., 2010), predictably changing with experience. Incorporating principled adjustability in lexical selection criteria similarly may allow a single model to account for a wider range of seemingly inconsistent data, including those from tasks with quite novel demands. In fact there is even some basis for such flexibility in the competitive selection rules specified for our 2010 model (e.g., Eq. 12: \( t_{selection} = \log_2(t/(a_i - a_{strongest\ competitor})) \)): the selection threshold, \( t \), decreases, the time required for lexical activation and selection reduces to the time required for lexical activation alone. Thus, the Dark Side model is readily amenable to incorporating adjustable selection criteria (Anders, Riés, Van Maanen, & Alario, 2017), including criteria based on competition, if principled means of estimating them can be delineated.

One challenge in defining flexibility, however, remains in distinguishing cases where a single mechanism operates in multiple ways (perhaps with only a tweak of a single parameter) from those where multiple mechanisms contribute broadly similar functions as needed. For instance, although both begin with a picture stimulus and conclude with a verbal response, it is not clear that the same selection processes operating in picture-word interference tasks need similarly contribute to simple picture naming. It is possible that selecting the only externally-defined-as-correct response from two very accessible alternatives, and/or suppressing a preferred name, could represent the same process as selecting within a range of similarly appropriate alternatives, but we wonder if these might be better characterized as distinct processes. If we assume that lexical activation, integrated over time, implements a non-competitive retrieval process, while a secondary process gates (or even monitors) further processing according to more flexible task demands (cf. Mahon et al., 2007), then this scenario begins to resemble Nozari and Hepner’s (2018) distinction between lexical activation and an adjustably competitive criterion for subsequent selection. Particularly in a cascading activation framework, it may be less useful to distinguish between early and late processes than between obligatory and optional processes.
Thus, there remains a tension between studying language production as it is and modifying it to fit particular laboratory constraints that we think might highlight particular aspects of the process. A gold standard in speech error research, therefore, is to demonstrate that patterns that emerge from controlled manipulations also emerge in error corpora (e.g., Dell & Reich, 1981). Similarly, with naming latency research, the same factors that emerge from experiments with obvious manipulations, like picture-word interference, should also at least hold in simpler paradigms, like normal picture naming. However, experiments with obvious manipulations, like latency research, the same factors that emerge from experiments with obvious manipulations, like picture-word interference, should also at least hold in simpler paradigms, like normal picture naming (e.g., Balatsou, Fischer-Baum and Oppenheim, in prep; Oppenheim, in prep). When effects are limited to particular paradigms, then although it is possible that those paradigms are uniquely suited to reveal special features of the system, we suggest it is also worth considering whether they actually tell us about core processes that are typically involved in language production.

Notes
1. But note that observed latencies may not actually be independent (e.g., Oppenheim, 2017).
2. Given its provenance as a model of error detection, it is worth noting that a conflict-assessment function could produce pseudocompetitive latency effects, by triggering a time-consuming monitoring process (e.g., Dhooge & Hartsuiker, 2010) that is distinct from typical notions of an early lexical selection process.

Disclosure statement
No potential conflict of interest was reported by the authors.

ORCID
Gary M. Oppenheim https://orcid.org/0000-0001-8121-0074

References
Nozari, N., & Hepner, C. (2018). To select or to wait? The importance of criterion setting in debates of competitive lexical


Oppenheim, G. M. (in prep.). Competition in normal word production?


