CrossMark

COMMENT

Integrating SLAM with existing evidence: Comment on Walker and Hickok (2015)

Matthew Goldrick[1]

Abstract Walker and Hickok (*Psychonomic Bulletin & Review* doi:10.3758/s13423-015-0903-7, 2015) used simulations to compare a novel proposal, the semantic–lexical–auditory–motor model (SLAM), to an existing account of speech production, the two-step interactive account (TSIA; Foygel & Dell, *Journal of Memory and Language*, 43:182–216, doi:10.1006/jmla.2000.2716, 2000). This commentary critically examines their assessment of SLAM. The cases in which SLAM outperforms TSIA largely reflect SLAM's ability to (poorly) approximate an existing theory of speech production incorporating two stages of phonological processing (the lexical+ postlexical account). The fact that SLAM and TSIA can exhibit equivalent fits to the overall response distribution of a set of aphasic patients is unsurprising, since previous work has shown that overall response distributions do not reliably discriminate theoretical alternatives. Finally, SLAM inherits issues associated with TSIA's assumption of strong feedback between levels of representation. This suggests that SLAM does not represent an advance over existing theories of speech production.

Keywords Speech production · Phonology · Connectionist models

Walker and Hickok (2015; henceforth, WH) have presented results from a simulation of speech production implementing aspects of Hickok's (2012) hierarchical state feedback control theory. They contrasted this proposal to Foygel and Dell's

(2000) two-step interactive account (TSIA; see also Dell, Lawler, Harris, & Gordon, 2004; Schwartz, Dell, Martin, Gahl, & Sobel, 2006). These are depicted on the right and left sides of Fig. 1 (respectively). Both accounts assume that speech production involves interaction between semantic, lexical, and phonological representations. WH's proposal also includes a second set of phonological representations, corresponding to auditory information, that interact with both lexical and (motor) phonological representations (leading to the moniker *SLAM*: the semantic–lexical–auditory–motor model). WH examined the relative abilities of simulations of both models to account for the overall response patterns of a set of individuals with aphasia. Neurological impairment was modeled by reducing the amount of activation flowing between levels of representation; this increased the relative influence of random noise, leading to errors.

WH reported two major findings: Their simulation of SLAM exhibited a degree of fit to overall response distributions similar to the fit of a simulation of the TSIA, and the simulation of SLAM exhibited a relatively better fit for individuals that were assigned the clinical label of conduction aphasia than did the TSIA simulations. This commentary reexamines these claims in light of previous work that has established empirical issues with TSIA and methodological issues with the approach of Foygel and Dell (2000). A comparison of SLAM with existing theoretical proposals arising out of this research reveals clear shortcomings of this new proposal.

Empirical challenges to TSIA's account of sound structure processing

The lexical + postlexical account An overall performance pattern that is difficult to account for under TSIA is the

✉ Matthew Goldrick
matt-goldrick@northwestern.edu

[1] Department of Linguistics, Northwestern University, 2016 Sheridan Road, Evanston, IL 60208, USA
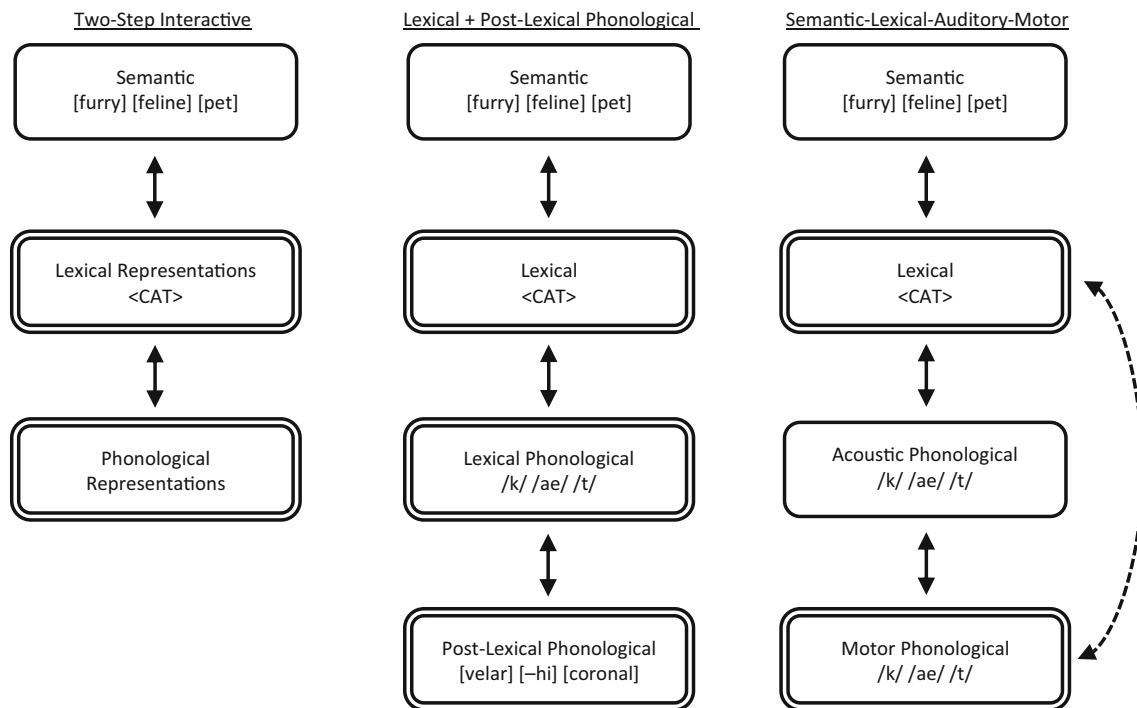
Fig. 1 Theoretical proposals contrasted in this commentary. Arrows indicate the directions of activation flow between representations connected by the lines. Dotted lines indicate weaker activation flow between representations than that marked by solid lines. Selection points are indicated by a double outline around a level of representation

production of only phonologically related errors (i.e., form-related errors such as *cat → hat*, as well as neologisms such as *cat → zat*; Caramazza, Papagno, & Ruml, 2000; Schwartz et al., 2006). In TSIA, phonologically related errors (in particular, neologisms or nonword errors) are most likely to arise during phonological processing. However, because cascading activation serves to activate semantically related words at the phonological level, impairments to this level of processing are likely to result in the production of semantic as well as phonological errors (Rapp & Goldrick, 2000). TSIA thus predicts that individuals should never produce a pattern of only phonologically related errors.

A number of studies have documented individuals that violate this prediction (Galluzzi, Bureca, Guariglia, & Romani, 2015; Goldrick & Rapp, 2007; Romani & Galluzzi, 2005; Romani, Galluzzi, Bureca, & Olson, 2011; Romani, Olson, Semenza, & Granà, 2002). Furthermore, the errors of individuals exhibiting this pattern are strongly influenced by the acoustic/articulatory complexity of phonological structures (e.g., exhibiting errors on less-frequent sequences of consonants), but relatively uninfluenced by lexical properties (e.g., word frequency). This contrasts with other individuals that produce phonological errors yet show a complementary pattern: sensitivity to lexical factors (e.g., lower accuracy on low-frequency words) and an insensitivity to the complexity of phonological structures.

These results can be accounted for by a theory that distinguishes multiple levels of sound structure processing in production. As is shown in the middle panel of Fig. 1, this account parallels the TSIA, in that lexical selection is followed by a stage of processing during which relatively abstract specifications of phonological structure are retrieved (lexical phonological processing). A second stage of (postlexical) phonological processing then retrieves and selects more detailed aspects of sound structure (e.g., featural representations; this leads to the moniker *lexical + postlexical* [LPL] account). Note that this is a distinct stage of production processing, in that it follows the explicit selection of an abstract phonological representation. In general, such selection mechanisms serve to reduce interactions across processing levels, increasing the degree to which distinct subprocesses can exhibit distinct patterns of impairment (Rapp & Goldrick, 2000).

Whereas lexical phonological processing begins with the selection of a lexical representation (and the coactivation of semantically related words), postlexical processing is initiated by the selection of a phonological representation—resulting in the coactivation of multiple phonological structures (e.g., for the target *cat*, syllables corresponding to words such as *hat*, as well as nonword syllables such as *zat*). Disruption to postlexical processing therefore results in the production of phonologically related words as well as nonwords, accounting for the overall performance pattern discussed above. The presence of distinct representational types at each discretely separated stage of processing also accounts for more detailed

aspects of their performance. Individuals with deficits arising in lexical phonological processing will be strongly influenced by lexical factors (reflecting the input to lexical processing), but not by phonological complexity (reflecting the abstract structure of lexical phonological representations). Individuals with deficits to a postlexical stage, governed by relationships among fully specified phonological structures, will not be influenced by lexical factors but will show strong effects of phonological complexity.

Finally, because postlexical processing occurs after the retrieval of abstract structures from long-term memory, it is assumed to be engaged by all spoken production tasks. Consistent with this assumption, individuals who produce only phonologically related word and nonword errors in picture naming produce similar patterns in performance of repetition and reading aloud (Goldrick & Rapp, 2007; Romani et al. 2002).

Assessing SLAM relative to LPL One of WH's major findings is that SLAM simulations show a better fit than TSIA to the performance of individuals with conduction aphasia. This clinical label is applied to individuals who typically (but not always) produce phonological errors in both repetition and picture naming in the context of intact articulatory and auditory comprehension processes—similar to the postlexical pattern reviewed above. In fact, inspection of individual conduction aphasia cases reveals that this improvement in fit largely reflects SLAM's relative success in accounting for individuals who produce primarily phonologically related errors.

This was assessed by using WH's online fitting algorithm (http://cogsci.uci.edu/~alns/webfit.html) to fit SLAM and TSIA simulations (based on 2,321 map points) to the performance of 50 individuals with conduction aphasia from version 2.0 of the Moss Aphasia Psycholinguistic Project Database (Mirman et al., 2010).[1] As is shown in Table 1, the ten individuals with the greatest improvement in fit show a performance pattern similar to the postlexical pattern identified above. The vast majority of these individuals' errors are phonologically (formally) related words or nonwords (a response category likely to include phonologically related forms). In fact, across the set of 50 individuals with conduction aphasia, the relative proportions of errors that fall into these two categories are significantly correlated with the amount of SLAM's improvement in *RMSD* relative to TSIA [$r(48) = .61$, $p < .0001$]. This suggests that SLAM is outperforming TSIA because it better matches deficits that result in the production of predominantly form-related word and nonword errors.

---

[1] Thanks to Dan Mirman and Stephen Faha for assistance in accessing these data.

Table 1 Response proportions for each of the ten conduction aphasia cases showing the greatest improvements in fit for the semantic–lexical–auditory–motor theory relative to the two-step interactive account

Participant ID	Correct	Semantic	Formal	Mixed	Unrelated	Nonword
LD0062	46%	4%	14%	3%	2%	30%
MR1230	49%	1%	18%	2%	1%	29%
MR0129	37%	1%	17%	1%	4%	40%
MR0333	42%	3%	16%	3%	5%	31%
MR0595	70%	1%	10%	1%	1%	16%
MR1185	68%	1%	10%	2%	1%	18%
MR0281	69%	1%	8%	1%	0%	20%
MR1336	50%	2%	17%	2%	1%	27%
MR1939	29%	2%	19%	2%	4%	43%
MR0315	49%	2%	14%	2%	3%	30%

The LPL account can also account for the overall response distribution of such individuals, by assuming deficits to both lexical processes (resulting in semantically related errors) and postlexical processes (which increases the rate specifically of phonologically related errors). Given that both of these accounts clearly outperform TSIA for this general pattern, which provides a more comprehensive account of the overall set of existing data? To examine this, the fit of SLAM to a prototypical case of *only* phonological errors in production (BON; Goldrick & Rapp, 2007) was examined. As is shown in Table 2, SLAM has great difficulty fitting this error pattern; it predicts that semantic as well as form-related errors should be produced. Interestingly, SLAM attempts to fit this by approximating the connectivity of LPL. The lexical–motor phonological connections are set to a negligible value (0.0051), whereas all other connections are set to a high value (0.035). However, merely approximating this connectivity pattern is insufficient; fully implementing the LPL account would require also adding in an explicit selection process during the first stage of phonological processing (see Goldrick & Rapp, 2002, for an analysis of the consequences of weakening or eliminating selection within these spreading-activation theories).

In addition to the challenges in matching the overall error distributions of these cases, SLAM offers no account of the differential effects of phonological complexity versus lexical variables on different deficits, and offers no general account of how multiple stages of phonological processing might be incorporated in production (for additional discussion of the issues in the context of the hierarchical state feedback control theory more generally, see Rapp, Buchwald, & Goldrick, 2014; Roelofs, 2014). Thus, the LPL account provides a clearly superior account of the types of cases on which SLAM outperforms TSIA.

Table 2 Observed versus predicted response distributions for BON (Goldrick & Rapp, 2007), an individual producing only phonologically related errors in production

	Correct	Semantic	Formal	Mixed	Unrelated	Nonword
Observed	77.8%	0.0%	6.1%	0.0%	0.0%	16.2%
Best SLAM fit	75.3%	2.0%	4.9%	1.6%	0.2%	16.1%

Methodological challenges to simulation studies

The other major result of WH is that SLAM exhibited a degree of fit to overall response distributions similar to that of a simulation of the TSIA. This follows previous studies of TSIA, which have assumed that the degree to which simulations fit the overall response distribution of each participant (e.g., proportions of correct responses, semantic errors, phonologically related errors, etc.) provides a general means of distinguishing between the theories corresponding to each simulation. Although this may be true of some theoretical accounts (e.g., global vs. local disruptions to the production system; Foygel & Dell, 2000), in many cases it fails.

Goldrick (2011) demonstrated this by examining the ability of TSIA simulations to fit simulated data sets. Artificial case series were generated using simulations of (a) Foygel and Dell's (2000) TSIA, (b) a theory in which speech errors arise prior to the two steps of lexical access assumed in TSIA, and (c) Rapp and Goldrick's (2000) restricted interaction account, which differs from TSIA in the strength and nature of feedback. When the parameter-fitting procedure of Dell et al. (2004) was then used to fit the TSIA to each of these artificial case series, and the degrees of fit were equivalent for all three artificial case series. Thus, with respect to overall response distributions, TSIA was able to fit data generated by a TSIA simulation just as well as data generated by simulations of distinct theoretical accounts. This suggests that overall response distributions frequently fail to discriminate what type of theory generated a given set of data. In light of these results, the fact that SLAM and TSIA exhibited equivalent fits to overall response distributions is unsurprising; in many cases, this measure will fail to discriminate alternative theories. Focusing on specific aspects of performance, motivated by theoretical contrasts, is a more effective means of distinguishing accounts than measures of overall response distributions (Goldrick, 2011; Rapp & Goldrick, 2000).

Issues outside of sound structure processing for TSIA and SLAM

Schwartz et al. (2006) have noted another overall performance pattern that is difficult for TSIA to account for: modality-specific impairments to speech production that result only in the production of semantic errors (see also Cuetos, Aguado, & Caramazza, 2000, for a discussion). Several studies have documented this pattern of performance (Basso, Taborelli, & Vignolo, 1978; Caramazza & Hillis, 1990; Cuetos et al., 2000; Miceli, Benvegnú, Capasso, & Caramazza, 1997; Nickels, 1992; see also Rapp & Goldrick, 2000). Rapp and Goldrick (2000) presented simulation results showing that this pattern is difficult for TSIA to account for because it incorporates strong feedback from phonological to lexical representations. Such strong feedback is also inconsistent with chronometric and speech error data from unimpaired speakers (see Goldrick, 2006, for a review). Because SLAM adopts similar assumptions regarding feedback, it is likely that it suffers from these same issues. Rapp and Goldrick's restricted interaction account provides an alternative that successfully addresses these challenges.

Conclusions

WH, following Hickok (2012), motivated the SLAM model by attempting to integrate psycholinguistic and speech motor control approaches to speech production. Although such cross-disciplinary conceptual integration is a laudable goal, it requires a full integration with the rich set of data and theory from psycholinguistic approaches to speech production. SLAM fails to achieve this. To the extent that SLAM outperforms the TSIA, it does so by poorly approximating the LPL account; SLAM is less successful than this existing theory in accounting for the full range of behavioral data. SLAM also fails to address methodological issues from existing work with the TSIA model, and fails to address issues in semantic and lexical processing that are problematic for TSIA. These issues suggest that a true integration of psycholinguistic and speech motor control theories will require a different approach.

References

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.

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

Caramazza, A., Papagno, C., & Ruml, W. (2000). The selective impairment of phonological processing in speech production. *Brain and Language, 75,* 428–450. doi:10.1006/ brln.2000.2379

Cuetos, F., Aguado, G., & Caramazza, A. (2000). Dissociation of semantic and phonological errors in naming. *Brain and Language, 75,* 451–460.

Dell, G. S., Lawler, E. N., Harris, H. D., & Gordon, J. K. (2004). Models of errors of omission in aphasic naming. *Cognitive Neuropsychology, 21,* 125–145.

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language, 43,* 182–216. doi:10.1006/jmla.2000.2716

Galluzzi, C., Bureca, I., Guariglia, C., & Romani, C. (2015). Phonological simplifications, apraxia of speech and the interaction between phonological and phonetic processing. *Neuropsychologia, 71,* 64–83. doi:10.1016/j.neuropsychologia.2015.03.007

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

Goldrick, M. (2011). Theory selection and evaluation in case series research. *Cognitive Neuropsychology, 28,* 451–465.

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. (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition, 102,* 219–260.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience, 13,* 135–145. doi:10.1038/nrn3158

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.

Mirman, D., Strauss, T. J., Brecher, A., Walker, G. M., Sobel, P., Dell, G. S., & Schwartz, M. F. (2010). A large, searchable, web-based database of aphasic performance on picture naming and other tests of cognitive function. *Cognitive Neuropsychology, 27,* 495–504. doi:10.1080/02643294.2011.574112

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

Rapp, B., Buchwald, A., & Goldrick, M. (2014). Integrating accounts of speech production: The devil is in the representational details. *Language, Cognition and Neuroscience, 29,* 24–27.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107,* 460–499. doi:10.1037/0033-295X.107.3.460

Roelofs, A. (2014). Integrating psycholinguistic and motor control approaches to speech production: Where do they meet? *Language, Cognition and Neuroscience, 29,* 35–37.

Romani, C., & Galluzzi, C. (2005). Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory and phonological difficulties. *Cognitive Neuropsychology, 22,* 817–850.

Romani, C., Galluzzi, C., Bureca, I., & Olson, A. (2011). Effects of syllable structure in aphasic errors: Implications for a new model of speech production. *Cognitive Psychology, 62,* 151–192.

Romani, C., Olson, A., Semenza, C., & Granà, A. (2002). Patterns of phonological errors as a function of a phonological versus an articulatory locus of impairment. *Cortex, 38,* 541–567.

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language, 54,* 228–264.

Walker, G. M., & Hickok, G. (2015). Bridging computational approaches to speech production: The semantic–lexical–auditory–motor model (SLAM). *Psychonomic Bulletin & Review.* doi:10.3758/s13423-015-0903-7. Advance online publication.