

COMMENT

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

Matthew Goldrick¹

Published online: 10 November 2015
© Psychonomic Society, Inc. 2015

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

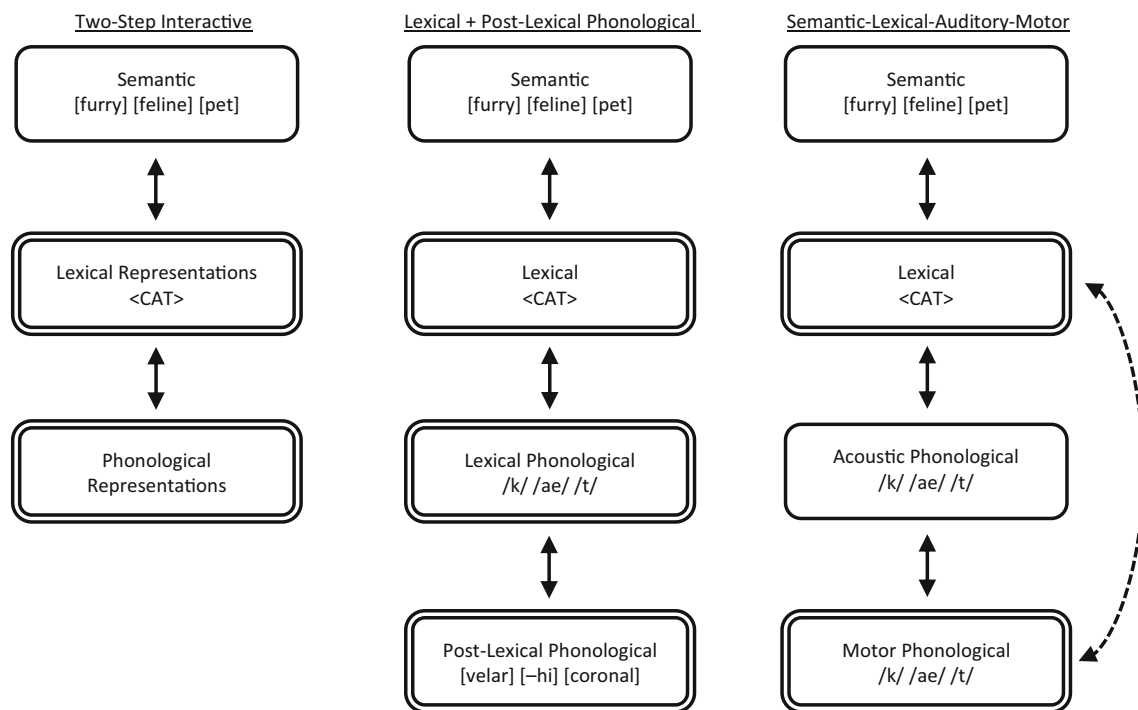


Fig. 1 Theoretical proposals contrasted in this commentary. Arrows indicate the directions of activation flow between representations connected by the lines. Dotted lines indicate weaker activation flow

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, & Rumel, 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.

Author note Thanks to Laurel Brehm for helpful comments on the manuscript.

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., & Rumel, 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., Benvegnù, 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.

title: "Integrating SLAM with existing evidence: Comment on Walker and Hickok (2015)" author: Matthew Goldrick journal: Psychonomic Bulletin Review year: 2016 volume: 23 pages: 648–652 DOI: 10.3758/s13423-015-0946-9

Goldrick (2016) SLAM と既存の証拠との統合: Walker and Hickok (2015) へのコメント

要旨

Walker と Hickok (Psychonomic Bulletin & Review, 2015) はシミュレーションを用いて新しい提案、意味-語彙-聴覚-運動モデル (SLAM) を既存の音声生成の説明モデル、2 段階相互活性化モデル (TSIA; Foygel & Dell, Journal of Memory and Language, 43:182--216, 2000) と比較した。本解説では SLAM の評価を批判的に検証する。SLAM が TSIA よりも優れているケースは、音韻処理の 2 つの段階 (語彙 + 後語彙説明) を組み込んだ音声生成の既存の理論を (不完全に) 近似する SLAM の能力を主に反映している。これまでの研究では、全体的な応答分布は理論的な代替案を確実に識別できないことが示されていたので SLAM と TSIA が失語症患者の全体的な応答分布に同等の適合度を示すことができるという事実は、驚くに値しない。最後に SLAM は表現レベル間の強いフィードバックという TSIA の仮定に関連した問題を継承している。このことは SLAM が既存の音声生成理論を超える進歩ではないことを示唆している。

keywords: Speech production; Phonology; Connectionist models

Walker & Hickok (2015; 以下 WH) は Hickok (2012) の階層的状態フィードバック制御理論の側面を実装した音声生成シミュレーション結果を発表した。彼らはこの提案を、Foygel & Dell (2000) の 2 段階相互活性化モデルによる説明 (TSIA; Dell, Lawler, Harris, & Gordon, 2004; Schwartz, Dell, Martin, Gahl, & Sobel, 2006 参照) と対比させた。これらは、図 1 の左右にそれぞれ描かれている。どちらの説も、音声生成には、意味、語彙、音韻表現の相互作用があるとしている。WH の提案では、聴覚情報に対応する第 2 の音韻表現が含まれており、これが語彙表現と (運動) 音韻表現の両方と相互作用する (SLAM: semantic-lexical-auditory-motor model) とされている。WH は、失語症患者の全体的な反応パターンを説明するために、両モデルのシミュレーションの相対的な能力を検証した。神経障害は、表現レベル間を流れる活性化の量を減らすことでモデル化されており、これによりランダムノイズの影響が相対的に大きくなり、エラーが発生する。

WH は 2 つの主要な発見を報告した。SLAM シミュレーションは TSIA シミュレーションと同様に全体的な回答分布にフィットしていたこと、TSIA シミュレーションに比べて SLAM シミュレーションは伝導失語という臨床的ラベルを付与された個人に対して比較的良いフィットを示していたことである。本解説では、TSIA の経験的問題や Foygel & Dell (2000) のアプローチの方法論的な問題を確立した過去の研究に照らして、これらの主張を再検討した。この研究から生まれた既存の理論的提案と SLAM を比較すると、この新しい提案には明らかな欠点があることがわかる。

1. TSIA の音構造処理の説明に対する経験的な挑戦

1.1 語彙 + 語彙後処理による説明

TSIA で説明するのが難しい全体的な成績パターンは、音韻に関連した誤り (cat → hat のような形に関連した誤りや, a cat → zat のような新造語など) のみの産出である。TSIA では、音韻に関連した誤り (特に新語や非単語の誤り) は、音韻処理中に生じる可能性が高い。しかし、カスケード活性化は音韻レベルで意味的に関連する単語を活性化するので、このレベルの処理に障害があると、音韻的な誤りだけでなく意味的な誤りも生じやすくなる (Rapp & Goldrick, 2000)。TSIA は音韻に関連した誤りだけのパターンを生成することはないと予測している。

多くの研究で、この予測に反する個体が記録されている (Galluzzi, Bureca, Guariglia, & Romani, 2015; Goldrick & Rapp, 2007; Romani & Galluzzi, 2005; Romani, Galluzzi, Bureca, & Olson, 2011; Romani, Olson, Semenza, & Granà, 2002)。さらに、このパターンを示す人の誤りは、音韻構造の音響的・聴覚的な複雑さに強く影響されるが (例えば、頻度の低い子音の並びで誤りを起こす)、語彙的特性 (例えば、単語頻度) には比較的影響されないという。これは、音韻エラーを起こしながらも、語彙的要素に敏感で (例えば、低頻度語では精度が低い)、音韻構造の複雑さには無頓着という補完的なパターンを示す他の患者とは対照的である。

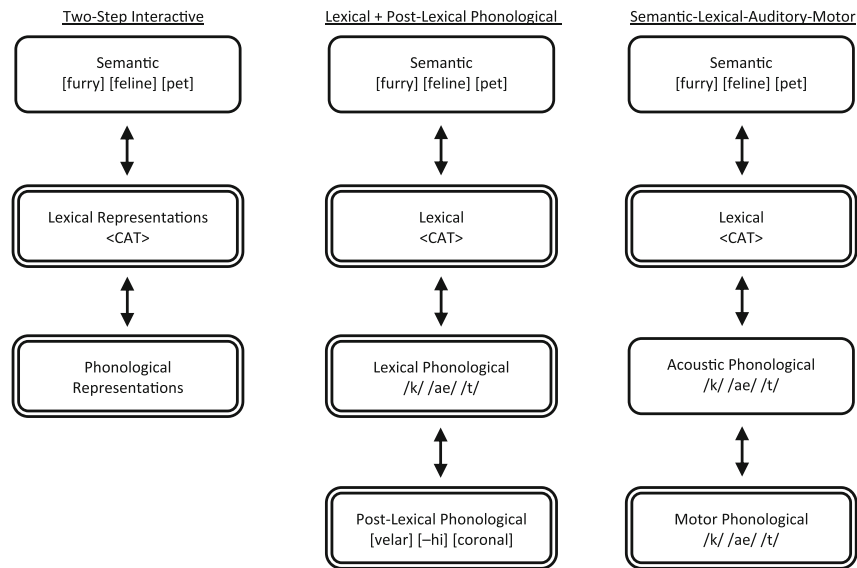


図 1. 本解説で対比される理論的提案。矢印は、線で結ばれた表象間の活性化の流れの方向を示す。点線は、実線に比べて表現間の活性化の流れが弱いことを示す。選択点は、表現のレベルを二重輪郭で示した。

これらの結果は、音声生成における音構造の処理を複数のレベルに分けて説明する理論によって説明することができる。図 1 中段に示したように、この理論は TSIA と類似しており、語彙選択後に、音韻構造の比較的抽象的な仕様を検索する処理段階がある(語彙的音韻処理)。その後、音韻処理の第 2 段階で、音構造のより詳細な側面(例えば、特徴的な表現)が検索され、選択される。このことから、語彙 + 語彙後処理 [LPL] 説明と呼ばれている。これは、抽象的な音韻表現を明示的に選択した後に行われるという点で、生成処理とは異なる段階であることに注意。一般的に、このような選択機構は、処理レベル間の相互作用を減らし、異なる下位処理が異なる障害パターンを示す度合いを高める役割を果たす (Rapp & Goldrick, 2000)。

語彙的な音韻処理が語彙表現の選択(および意味的に関連する単語の共役化)から始まるのに対し、音韻後処理は音韻表現の選択から始まり、その結果、複数の音韻構造が共役化される(例えば、ターゲット cat の場合、hat などの単語に対応する音節と zat などの単語ではない音節が併存する)。このように、音韻後処理が阻害されると、音韻に関連した単語と非単語が生成され、前述のような全体的な成績パターンが形成される。また、それぞれの処理段階で異なる表現タイプが存在することも、成績のより詳細な側面を説明する。語彙的音韻処理に障害がある人は、語彙的要因(語彙処理への入力に反映)に強く影響されるが、音韻的複雑さ(語彙的な音韻表現の抽象的な構造を反映)には影響されない。完全に指定された音韻構造間に関係に支配される語彙後段階に障害がある人は、語彙的要因の影響を受けないが、音韻的複雑さの影響を強く受けることになる。

最後に、語彙後処理は、長期記憶から抽象的な構造を取り出した後に行われるため、すべての音声生成課題に関与していると考えられる。この仮定と一致するように、絵画名前検査時に音韻的に関連した単語と非単語のエラーしか出さない人は、反復練習と音読成績に同じようなパターンを示す (Goldrick & Rapp, 2007; Romani et al. 2002)。

1.2 LPL に対する SLAM の評価

WH の主要な発見の 1 つは、SLAM シミュレーションが TSIA よりも伝導失語の人の成績によくフィットすることである。
-> 伝導失語の課題成績. 復唱と絵画命名課題における音韻エラーが特徴

WH の主要な発見の一つは、SLAM シミュレーションが TSIA よりも伝導失語患者の成績によく適合することである。伝導失語症とは、明瞭な発声や聴覚的な理解処理があるにもかかわらず、復唱や絵画命名時に音素的な誤りを犯すことが多い人のことで、前述の語彙後処理パターンと似ている。実際、個々の伝導失語症例を調べてみると、この適合性の向上は、主に音韻に関連した誤りを生じている人に対する SLAM の相対的な成功を反映していることがわかる。

これは WH のオンラインフィッティングアルゴリズム (<http://cogsci.uci.edu/~alns/webfit.html>) を用いて、SLAM および TSIA シミュレーション (2,321 個の地図点に基づく) を、Moss Aphasia Psycholinguistic Project Database (Mirman et al., 2010) のバージョン 2.0 に含まれる 50 人の伝導失語症患者の成績にフィットさせることで評価した(1)。表 1 に示すように、適合度が最も向上した 10 人は、上記で確認された語彙後パターンに類似した成績パターンを示している。これらの個人のエラーの大部分は、音韻的に(形式的に)関連する単語または非単語(音韻的に関連する形を含む可能性が高い回答カテゴリー)であ

る。実際、50 人の伝導失語症患者において、これら 2 つのカテゴリーに分類されるエラーの相対的な割合は、TSIA に対する SLAM の RMSD の改善量と有意な相関があった ($r(48)=0.61, p < 0.0001$)。このことは、SLAM が TSIA よりも優れているのは、主に形に関連した単語や非単語の誤りを生じさせる障害に、SLAM がよりよく適合しているからであることを示唆している。

Table 1. 意味的・語彙的・聴覚的・運動的理論の適合性が、2 段階相互的説明に比べて最も向上した伝導失語症 10 例のそれぞれの回答割合。

| 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% |

LPL は、語彙処理 (意味関連エラーを引き起こす) と語彙後処理 (音韻関連エラーの割合を増加させる) の両方に障害があると仮定することで、このような人々の全体的な反応分布を説明することができる。この一般的なパターンでは、どちらの説明も TSIA を明らかに上回っているが、既存のデータ全体をより包括的に説明できるのはどちらだろうか。この点を検討するために、音韻エラーのみが生じている典型的なケース (BON; Goldrick & Rapp, 2007) に対する SLAM の適合性を調べた。表 2 に示すように、SLAM はこのエラーパターンに適合することが非常に困難であり、意味エラーだけでなく、形態エラーも生成されるはずであると予測している。興味深いことに、SLAM は LPL の接続性を近似することでこれを適合させようとする。辞書・運動・音韻の接続は無視できる値 (0.0051) に設定され、その他の接続は高い値 (0.035) に設定されている。しかし、単にこの接続パターンを近似するだけでは不十分であり、LPL 説明を完全に実装するには、音韻処理の第一段階における明示的な選択処理も加える必要がある (Goldrick & Rapp, 2002 参照、これらの拡散活性化理論の中で選択を弱めたり排除したりした場合の結果について分析している)。

表 2. BON (Goldrick & Rapp, 2007) の観察された患者の応答と予測反応分布 (音韻関連エラーのみを生成する患者)

| | 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% |

これらのケースの全体的なエラー分布を一致させるための課題に加えて、SLAM は音韻の複雑さと語彙変数の異なる障害に対する差次的効果についての説明を提供しておらず、音韻処理の複数の段階がどのように算出に組み込まれるかについての一般的な説明を提供していない (より一般的な階層的な状態のフィードバック制御理論の文脈における問題の追加的な議論については、Rapp, Buchwald, & Goldrick, 2014; Roelofs, 2014 参照)。このように LPL は、SLAM が TSIA を上回るケースのタイプについて、明らかに優れた説明を提供している。

2. シミュレーション研究の方法論的課題

もう 1 つの主な結果は、SLAM が TSIA シミュレーションと同様に全体の回答分布にフィットする度合いを示したことである。これは、これまでの TSIA 研究では、シミュレーションが各被験者の全体的な応答分布 (正解率、意味エラー、音韻的エラーなど) にフィットする度合いが、各シミュレーションに対応する理論を区別する一般的な手段になると考えられてきた。これは、いくつかの理論的説明 (例えば、産出系の大域的混乱と局所的混乱 Foygel & Dell, 2000) には当てはまるかもしれないが、多くの場合は失敗する。

Goldrick (2011)は、TSIA シミュレーションのシミュレーションデータセットへの適合性を調べることで、このことを実証した。(a) Foygel and Dell (2000) の TSIA, (b) TSIA で想定されている 2 段階語彙アクセスの前に発話エラーが発生する理論, (c) Rapp and Goldrick (2000) の制限付き相互作用説明をシミュレーションして、人工的なケースシリーズを作成した。Dell ら (2004) のパラメータフィッティング手順を用いて、これらの人工的なケース・シリーズのそれぞれに TSIA をフィッティングしたところ、3つの人工的なケースシリーズすべてにおいてフィットの度合いは同等であった。このように、全体的な反応分布に関して、TSIA は TSIA シミュレーションによって生成されたデータを、異なる理論的説明のシミュレーションによって生成されたデータと同様に適合させることができた。このことは、全体の反応分布が、どのような理論に基づいてデータが生成されたかを判別できないことが多いことを示唆している。これらの結果から、SLAM と TSIA の回答分布の適合度が同等であることは当然のことであり、多くの場合、この尺度では理論の違いを判別することができない。理論的な対比に基づいて成績の特定の側面に焦点を当てることは、全体的な反応分布の測定よりも説明を区別するより効果的な手段である (Goldrick, 2011; Rapp & Goldrick, 2000)。

3. TSIA と SLAM の音構造処理以外の問題

Schwartz ら (2006) は、TSIA が説明するのが難しいもう一つの全体的な成績パターンとして、意味エラーを生じるだけの音声生成に対するモダリティ特有の障害を挙げている (Cuetos, Aguado, & Caramazza, 2000 も参照)。いくつかの研究で、このパターンの成績が記録されている (Basso, Taborelli, & Vignolo, 1978; Caramazza & Hillis, 1990; Cuetos et al., 2000; Miceli, Benvegnù, Capasso, & Caramazza, 1997; Nickels, 1992; Rapp & Goldrick, 2000 も参照)。Rapp & Goldrick (2000) は、このパターンは音韻表現から語彙表現への強いフィードバックを含んでいるため、TSIA では説明が難しいというシミュレーション結果を示した。このような強いフィードバックは、障害のない話者の時系列データや音声エラーのデータとも矛盾している (レビューはGoldrick, 2006 参照)。SLAM は、フィードバックに関して同様の仮定を採用しているため、これらの同じ問題に悩まされる可能性が高い。Rapp & Goldrick の制限付き相互作用の説明は、これらの課題をうまく解決する代替案を提供している。

4. 結論

WH は Hickok (2012) に続き、言語心理学的アプローチと音声運動制御的アプローチを統合しようとすることで、SLAM モデルの動機付けを行った。このような分野横断的な概念統合は賞賛すべき目標であるが、音声生成に対する心理言語学的アプローチの豊富なデータと理論を完全に統合する必要がある。SLAM はこれを達成できていない。SLAM は TSIA よりも優れているが、それは LPL の近似性が低いためであり、SLAM はこの既存の理論よりも行動データを完全に説明することができない。SLAM はまた、TSIA モデルを用いた既存の研究から得られた方法論的な問題を解決することができず、TSIA で問題となっている意味処理や語彙処理の問題も解決できていない。これらの問題は、言語心理学と音声運動制御理論の真の統合には、別のアプローチが必要であることを示唆している。