

## Integrating psycholinguistic and motor control approaches to speech production: where do they meet?

Ardi Roelofs\*

*Donders Institute for Brain, Cognition and Behaviour, Centre for Cognition, Radboud University Nijmegen, Spinoza Building B.02.34, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands*

*(Received 3 September 2013; accepted 2 October 2013)*

To integrate psycholinguistic and motor control approaches to speech production, Hickok proposes the hierarchical state feedback control (HSFC) model. Here, I contest the direct mapping between lemmas and syllable motor programmes proposed by Hickok, and argue that this mapping proceeds indirectly via morphemic and phonological word representations. The results of computer simulations suggest that phoneme-based mapping is a viable alternative to the lemma-based mapping of the HSFC model.

**Keywords:** articulatory programmes; lemmas; motor control; psycholinguistic; simulations; word production

Hickok (2012, 2013) proposes the hierarchical state feedback control (HSFC) model to integrate psycholinguistic and motor control approaches to speech production. I agree that integration is important but take issue with the specific proposal of Hickok instantiated in the HSFC model. In particular, I argue against the direct mapping between lemmas and syllable motor programmes proposed by Hickok and, instead, propose that the mapping of lemmas onto motor programmes is mediated by morphemic and phonological word representations, following Levelt, Roelofs, and Meyer (1999). I present the results of computer simulations using a modified version of Hickok's (2012) implemented model that instantiates phoneme-based mapping, which show that it is a viable alternative to the HSFC model.

Hickok proposes that lemmas are directly connected to 'motor syllable programmes', which are connected to 'motor phoneme programmes'. In contrast, Levelt et al. distinguish between 'phonological syllables', which are part of online constructed 'phonological word representations', and 'syllable motor programmes', which seem to correspond to Hickok's motor syllable programmes. According to the proposal of Levelt et al. (1999), computationally implemented in the WEAVER++ model, the conceptually driven planning of a spoken word involves conceptual preparation, lemma retrieval and word-form encoding, with the latter consisting of morphological, phonological and phonetic encoding. In line with Hickok (2012, 2013), the model assumes internal and external monitoring loops involving the speech comprehension system (Levelt et al., 1999; Özdemir, Roelofs, & Levelt, 2007; Roelofs, 2004; Roelofs, Özdemir, & Levelt, 2007).

Information about words is stored in a large associative network that is accessed by spreading activation. Lemmas specify the grammatical properties of words, crucial for their use in phrases and sentences. For example, the lemma of *get* specifies that the word is a verb. Lemmas also allow for the specification of morphosyntactic parameters, such as number (singular and plural) for nouns and number, person (first, second and third) and tense (past and present) for verbs so that the appropriate morphemes may be retrieved (e.g. first-person singular present-tense <get>). In phonological encoding, the associated phonemes are selected (i.e. /g/, /e/ and /t/) and syllable positions are assigned to the phonemes (e.g. /t/ will become syllable coda). The assignment of syllable positions allows for syllabification across morpheme and word boundaries (e.g. Booij, 1995; Kenstowicz, 1994), for example, creating *ge-ting* in 'it is getting better' or *ge-tit* in 'I don't get it'. Here, the /t/ of *get* occupies a syllable onset position (in *ting* and *tit*) rather than a coda position (in *get*). The outcome of phonological encoding is a phonological word representation, which specifies the phonemes grouped into syllables and the stress pattern across syllables in polysyllabic words. In phonetic encoding, the constructed phonological word representation is used to retrieve corresponding syllable motor programmes (i.e. [get]).

As argued extensively by Levelt et al. (1999) and Roelofs (1997), models that store words as (sequences of) syllable nodes, like Hickok's HSFC model, have difficulty dealing with the need for flexibility of syllable membership. In contrast, models that include online syllabification may take neighboring morphemes and words into account.

---

\*Email: [A.Roelofs@donders.ru.nl](mailto:A.Roelofs@donders.ru.nl)

In that case, syllable positions will be computed for phonological words rather than for lexical ones. Online syllabification has to be assumed for some languages such as Dutch and English, where syllable positions are context-dependent, but not for all languages. For example, in Mandarin Chinese, syllable positions of phonemes are fixed (i.e. there is no resyllabification in the language) so that phonological syllables may be stored in memory rather than computed online. Evidence suggests that during phonological encoding in Mandarin Chinese, lexical tones are assigned to atonal stored phonological syllables to construct phonological word representations, which are then used to address tonally specified syllable motor programmes (O’Seaghdha, Chen, & Chen, 2010). Similarly, evidence suggests that in a language like Japanese, tones are assigned to moras to construct phonological word representations, which are then used to address tonally specified syllable motor programmes (Kureta, Fushimi, & Tatsumi, 2006). The results of computer simulations using Mandarin Chinese and Japanese versions of WEAVER++ revealed that these assumptions about phonological and phonetic encoding account for key empirical findings on Dutch, English, Mandarin Chinese and Japanese (Roelofs, 2013).

To recapitulate, whereas Hickok proposes a direct mapping from lemmas onto syllable motor programmes, I instead propose that the mapping of lemmas onto motor programmes happens indirectly via morphemic and phonological word representations. To demonstrate the viability of phoneme-based mapping, I ran computer simulations using a modified version of Hickok’s (2012) implemented model. The purpose of the simulations was to show that the activation patterns observed by Hickok

are replicated with phoneme rather than lemma input. If phoneme-based and lemma-based mappings yield equivalent activation patterns, then phoneme input is to be preferred because it provides a better integration of motor control and psycholinguistic approaches, as outlined above. Of course, the observation that phoneme input works just as well as lemma input does not prove the necessity of an indirect mapping of lemmas onto syllable motor programmes. However, as indicated earlier, such indirect mapping is independently motivated.

In the simulations, phonemes rather than lemmas were connected to syllable motor programmes, as illustrated in Figure 1A for the word *get*. Thus, the modified model concerns phonetic encoding in the terminology of Levelt et al. (1999), that is, the mapping of a phonological word representation onto articulatory programmes. The results of the simulations showed that the activation patterns of the modified model were very similar to the patterns of Hickok’s original model (see Hickok, 2012, Figure 5). For example, as Figure 1B shows, an activated syllable motor programme (i.e. [get]) inhibits its predicted auditory consequences (i.e. /get/), which supports error detection and correction. However, unlike the original version of Hickok’s model, the modified model can deal with the flexibility of syllable membership because syllable motor programmes are activated by phonemes rather than lemmas.

To conclude, whereas the motor control side of Hickok’s HSFC model is specified in detail, its psycholinguistic side consists of a simplified version of existing psycholinguistic proposals (e.g. Levelt et al., 1999). To better integrate psycholinguistic and motor control approaches to speech production, the more detailed

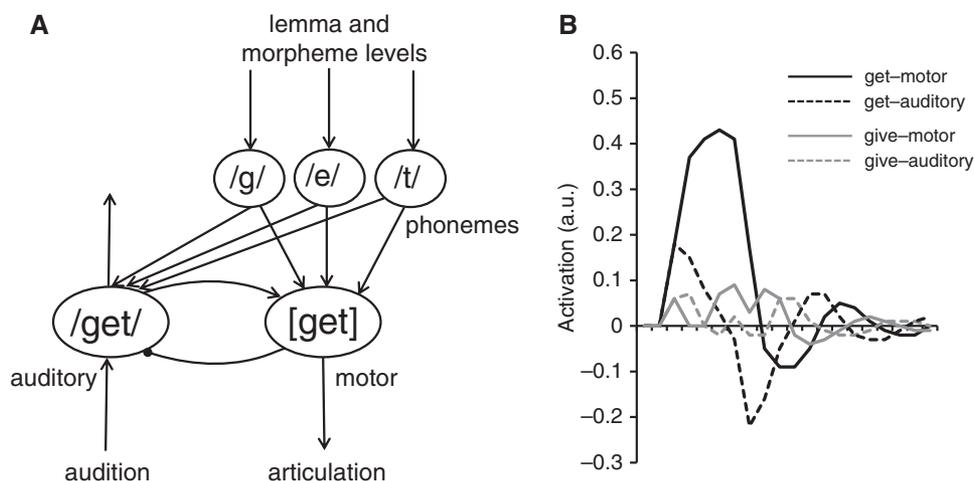


Figure 1. (A) Illustration of a modified version of Hickok’s (2012, 2013) model with a mapping of phonemes rather than lemmas onto syllable motor programmes. The motor-to-auditory connections (‘forward prediction’) are inhibitory and the auditory-to-motor connections (‘inverse correction’) are excitatory. (B) Activation patterns for motor and auditory representations when phonemes are mapped onto syllable motor programmes.

psycholinguistic proposals in the literature should be adopted.

## References

- Booij, G. (1995). *The phonology of Dutch*. Oxford, UK: Clarendon Press.
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*, 135–145. doi:10.1038/nrg3118
- Hickok, G. (2013). The architecture of speech production and the role of the phoneme in speech processing. *Language, Cognition and Neuroscience*.
- Kenstowicz, M. (1994). *Phonology in generative grammar*. Oxford, UK: Blackwell.
- Kureta, Y., Fushimi, T., & Tatsumi, I. F. (2006). The functional unit in phonological encoding: Evidence for moraic representation in native Japanese speakers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 1102–1119. doi:10.1037/0278-7393.32.5.1102
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38.
- O'Seaghdha, P. G., Chen, J.-Y., & Chen, T.-M. (2010). Proximate units in word production: Phonological encoding begins with syllables in Mandarin Chinese but with segments in English. *Cognition*, *115*, 282–302. doi:10.1016/j.cognition.2010.01.001
- Özdemir, R., Roelofs, A., & Levelt, W. J. M. (2007). Perceptual uniqueness point effects in monitoring internal speech. *Cognition*, *105*, 457–465. doi:10.1016/j.cognition.2006.10.006
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*, 249–284. doi:10.1016/S0010-0277(97)00027-9
- Roelofs, A. (2004). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: Comment on Rapp and Goldrick 2000. *Psychological Review*, *111*, 561–572. doi:10.1037/0033-295X.111.2.561
- Roelofs, A. (2013). Modeling of phonological encoding in word production: From Germanic languages to Mandarin Chinese and Japanese. *Japanese Psychological Research*.
- Roelofs, A., Özdemir, R., & Levelt, W. J. M. (2007). Influences of spoken word planning on speech recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 900–913. doi:10.1037/0278-7393.33.5.900