

# Inner speech as a forward model?

GARY M. OPPENHEIM\*

Center for Research in Language  
University of California, San Diego  
goppenheim@crl.ucsd.edu

## Abstract

*Pickering and Garrod consider the possibility that inner speech might be a product of forward production models. Here I consider the idea of inner speech as a forward model in light of empirical work from the past few decades, concluding that, while forward models could contribute to it, inner speech nonetheless requires activity from the implementers.*

Pickering and Garrod argue that coarse predictions from forward models can help detect errors of overt speech production before they occur. This error detecting function is often assigned to inner speech (e.g. Levelt, 1983; Levelt, Roelofs, & Meyer, 1999; Nooteboom, 1969), the little voice in one's head, better known for its role in conscious thought. It is therefore tempting to identify inner speech as a product of these forward models, with  $\hat{p} \rightarrow \hat{c}$  providing what we know as the internal loop. In fact, conceiving of inner speech as a forward model could elegantly address three key questions. First, why do we have inner speech at all? Inner speech is a by-product of the speaker's need to control their overt verbal behavior. Second, why does inner speech develop so long after overt speech (e.g. Vygotsky, 1962)? Inner speech develops as the speaker learns to simulate their verbal behavior, which may lag behind the ability to produce that behavior. And third, how are people able to produce inner speech without actually speaking aloud? If inner speech is simply the offline use of forward models ( $\hat{p} \rightarrow \hat{c}$ ), then speakers never need to engage the production and comprehension implementers ( $p \rightarrow c$ ) that are the traditional generators and perceivers of inner speech.

Pickering and Garrod's framework would specifically address two more recently demonstrated qualities of inner speech. First, inner speech involves attenuated access to subphonemic representations. When people say tongue-twisters in their heads, their reported errors are less influenced by subphonemic similarities than their reported errors when saying them aloud (Oppenheim & Dell, 2008, 2010; also Corley, Brocklehurst, & Moat, 2011, as noted by Oppenheim, 2012) For instance, /g/ shares more features with /k/ than with /v/, so someone trying to say GOAT aloud would more likely slip to COAT than VOTE, but this tendency is less pronounced for inner slips. As Pickering and Garrod note, this finding is predicted if the forward models underlying inner speech produce phonologically impoverished predictions (and thus might not reflect the production implementer). Second, inner speech is flexible enough to incorporate additional detail. Although inner slips show less pronounced similarity effects than overt speech, adding silent articulation is sufficient to boost their similarity effect, apparently coercing inner speech to include more subphonemic detail (Oppenheim & Dell, 2010). Such flexibility could be problematic for models that assign in-

---

\*Gary M. Oppenheim, Center for Research in Language, University of California, San Diego, 9500 Gilman Dr., Dept 0526, La Jolla, CA 92093-0526

ner speech to a specific level of the production process (e.g. Levelt *et al.*, 1999), but Pickering and Garrod's account specifically suggests that forward models simulate multiple levels of representation, so it might accommodate the subphonemic flexibility of inner speech by adding motoric predictions ( $\hat{p}$ [*sem,syn,phon,art*]; forward models' more traditional jurisdiction) that are tied to motor planning.

But forward model simulations cannot provide a complete account of inner speech. You still need to use what Pickering and Garrod would call 'the production implementer'. First, inner rehearsal facilitates overt speech production (MacKay, 1981; Rauschecker, Pringle, & Watkins, 2008; but *cf.* Dell & Repka, 1992), suggesting that some aspects of the production implementer are also employed in inner speech. Second, there is abundant evidence that people easily detect their inner speech errors (Corley *et al.*, 2011; Dell, 1978; Dell & Repka, 1992; Hockett, 1973; Meringer & Meyer, 1895, cited in MacKay, 1992; Oppenheim & Dell, 2008, 2010; Postma & Noordanus, 1996). But since monitoring is described as the resolution of predicted and actual percepts (from forward models and implementers, respectively), it is unclear how one could detect and identify inner slips without having engaged the production implementer. (Conflict monitoring, e.g. Nozari, Dell, & Schwartz, 2011, within forward models might at least allow error detection, but its use there seems to lack independent motivation, and still leaves the problem of how a speaker could identify the content of an inner slip). Third, analogues of overt speech effects are often reported for experiments substituting inner speech-based tasks. For instance, inner slips tend to create words, just like their overt counterparts (Corley *et al.*, 2011; Oppenheim & Dell, 2008, 2010), and their distributions resemble overt slips in other ways (Dell, 1978; Postma & Noordanus, 1996). And though inner and overt speech can diverge, they tend to elicit similar behavioral and neurophysiological effects in other domains (e.g. Kan & Thompson-Schill, 2004), and their impairments are highly correlated (e.g. Geva, Bennett, War-

burton, & Patterson, 2011). Though more ink is spilled cautioning differences between inner and overt speech, similarities between the two are the rule rather than the exception (at least for pre-articulatory aspects).

Given the impoverished character of Pickering and Garrod's forward models, it seems difficult to account for such parallels without assuming a role for production implementers in the creation of inner speech. Therefore, we could posit that inner speech works much like overt speech production, recalling Pickering and Garrod's acknowledgment that offline simulations could engage the implementers, actively truncating the process before articulation; forward models would supply a necessary monitoring component. This more explicit account of inner speech allows us to question Pickering and Garrod's suggestion that the subphonemic attenuation of inner speech might reflect impoverishment of the forward model instead of the generation of an abstract phonological code by the production implementer. Having clarified the role of forward models as error detection, their suggestion now boils down to the idea that inner slips might be hard to 'hear'. Empirical work suggests that is not the case. Experiments using noise-masked overt speech (Corley *et al.*, 2011) and silently mouthed speech (Oppenheim & Dell, 2010) showed that each acts much like normal overt speech in terms of similarity effects (see also Oppenheim, 2012). And, by explicitly modeling biased error detections, Oppenheim and Dell (2010) formally ruled out the suggestion that their evidence for abstraction merely reflected such biases. Thus, better specifying the role of forward models in inner speech allows the conclusion that the subphonemic attenuation of inner speech does have its basis in the production implementer. More generally, conceiving of forward models as *components of* inner speech can wed strengths of the forward model account with the fidelity of implementer-based simulations.

## REFERENCES

- [1] Corley, M., Brocklehurst, P. H., & Moat, H. S. (2011). Error biases in inner and overt speech: Evidence from tongue twisters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(1), 162–75. doi:10.1037/a0021321
- [2] Dell, G. S. (1978). Slips of the mind. *The fourth Lacus forum* (pp. 69–75). Columbia, S.C.: Hornbeam Press.
- [3] Dell, G. S., & Repka, R. J. (1992). Errors in inner speech. *Experimental slips and human error: Exploring the architecture of volition* (pp. 237–262). New York: Plenum.
- [4] Geva, S., Bennett, S., Warburton, E. a., & Patterson, K. (2011). Discrepancy between inner and overt speech: Implications for post-stroke aphasia and normal language processing. *Aphasiology*, 25(3), 323–343. doi:10.1080/02687038.2010.511236
- [5] Hockett, C. F. (1973). Where the tongue slips, there slip I. In V.A. Fromkin (Ed.), *Speech errors as linguistic evidence*. The Hague: Mouton.
- [6] Kan, I. P., & Thompson-Schill, S. L. (2004). Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cognitive, Affective, & Behavioral Neuroscience*, 4(1), 43–57.
- [7] Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14(1), 41–104.
- [8] Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, 22(01), 1–38.
- [9] MacKay, D. G. (1981). The problem of rehearsal or mental practice. *Journal of motor behavior*, 13(4), 274–285.
- [10] MacKay, D. G. (1992). Constraints on theories of inner speech. In D. Reisberg (Ed.), *Auditory imagery* (pp. 121–149). Lawrence Erlbaum Associates.
- [11] Nooteboom, S. G. (1969). The tongue slips into patterns. In A. G. Sciarone, A. J. van Essen, & A. A. van Raad (Eds.), *Leyden studies in linguistics and phonetics* (pp. 114–132). The Hague: Mouton.
- [12] Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive psychology*, 63(1), 1–33. doi:10.1016/j.cogpsych.2011.05.001
- [13] Oppenheim, G. M. (2012). The case for subphonemic attenuation in inner speech: Comment on Corley, Brocklehurst, and Moat (2011). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 502–512. doi:10.1037/a0025257
- [14] Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition*, 106(1), 528–537. doi:10.1016/j.cognition.2007.02.006
- [15] Oppenheim, G. M., & Dell, G. S. (2010). Motor movement matters: the flexible abstractness of inner speech. *Memory & cognition*, 38(8), 1147–1160. doi:10.1016/j.cognition.2007.02.006
- [16] Postma, A., & Noordanus, C. (1996). Production and detection of speech errors in silent, mouthed, noise-masked, and normal auditory feedback speech. *Language and Speech*, 39(4), 375–392.
- [17] Rauschecker, A. M., Pringle, A., & Watkins, K. E. (2008). Changes in neural activity associated with learning to articulate novel auditory pseudowords by covert repetition. *Human brain mapping*, 29(11), 1231–42. doi:10.1002/hbm.20460
- [18] Vygotsky, L. S. (1962). *Thought and language* (Hanfmann E and Vakar G, Trans.). Cambridge, Mass.: M.I.T. Press.