How Feedback in Interactive Activation Improves Perception

James S. Magnuson (james.magnuson@uconn.edu)
Psychological Sciences, U. Connecticut, Storrs, CT 06269-1020, USA
& BCBL. Basque Center on Cognition Brain and Language, Donostia-San Sebastián, Spain

Samantha Grubb (samantha.grubb@uconn.edu)
U. Connecticut, Storrs, CT 06269-1020, USA

Anne Marie Crinnion (anne.crinnion@uconn.edu)
U. Connecticut, Storrs, CT 06269-1020, USA

Sahil Luthra (sahilluthra@cmu.edu)
Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213-3815, USA

Phoebe Gaston (phoebe.gaston@uconn.edu)
Psychological Sciences, U. Connecticut, Storrs, CT 06269-1020, USA

Abstract

We follow up on recent work demonstrating clear advantages of lexical-to-sublexical feedback in the TRACE model of spoken word recognition. The prior work compared accuracy and recognition times in TRACE with feedback on or off as progressively more noise was added to inputs. Recognition times were faster with feedback at every level of noise, and there was an accuracy advantage for feedback with noise added to inputs. However, a recent article claims that those results must be an artifact of converting activations to response probabilities, because feedback could only reinforce the "status quo." That is, the claim is that given noisy inputs, feedback must reinforce all inputs equally, whether driven by signal or noise. We demonstrate that the feedback advantage replicates with raw activations. We also demonstrate that lexical feedback selectively reinforces lexically-coherent input patterns – that is, signal over noise – and explain how that behavior emerges naturally in interactive activation.

Keywords: computational modeling; interactive activation; spoken word recognition; speech perception

Introduction

Feedback from lexical to sublexical levels in interactive activation models (e.g., McClelland & Rumelhart, 1981) provides an intuitive explanation of lexical influences on sublexical processing – so-called *top-down* effects. A classic example is shown in Figure 1, where an identical visual pattern is interpreted as "H" in the context of "T_E" but as "A" in the context of "C_T". Lexical contexts implied by neighboring letters seem to influence the *perception* of the ambiguous form. Top-down feedback (from lexical to letter representations) would appear to provide an explanation of this and many other top-down effects in visual and spoken word recognition and perception (e.g., Elman & McClelland, 1988; Ganong, 1980; Luthra et al., 2021; Reicher, 1969; Rubin, Turvey, & Van Gelder, 1976; Samuel, 1997; Wheeler, 1970) and perception more generally (e.g., Fenske, Aminoff, Gronau, & Bar, 2006).

Figure 1: A variant of a famous example (Selfridge, 1955) of apparent top-down modulation of perception. Identical forms are interpreted as "H" between "T" and "E" but as "A" between "C" and "T".

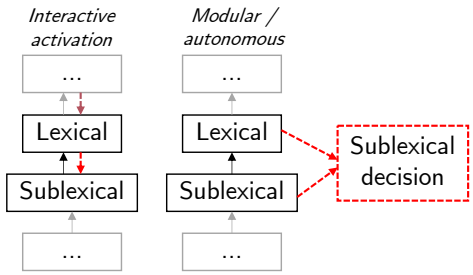

Figure 2: Interactive (left) vs. modular/autonomous architectures (right), exemplified by TRACE (McClelland & Elman, 1986) and Merge (Norris et al., 2000). Black outlines and connections are key in both TRACE and Merge; grey boxes and connections are either not implemented in *either* model (supralexical) or Merge (input to sublexical). Red elements highlight differences between approaches. The dashed red arrow from the grey supralexical box back to the lexical level highlights a key difference that would be included in an extension of interactive activation beyond the lexical level.

Norris et al. (2000) argue that intuition misleads us in such cases, and that top-down effects are not evidence for top-down feedback during processing. They claim that *any* such effect can be explained by modular or autonomous architectures where feedback is absent from perceptual pathways and top-down knowledge and other constraints are integrated post-perceptually. The differences between interactive and modular/autonomous approaches are schematized in Figure 2. In the interactive architecture, direct top-down feedback from the lexical level modulates sublexical activations. This leads naturally

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

to lexical representations influencing sublexical processing. For example, suppose that the ambiguous H/A form in Figure 1 activates "H" and "A" equally. Because "TAE" rarely occurs (e.g., *AORTAE*) but "THE" is a high-frequency pattern, "H" will receive strong feedback in that context. Because "CHT" would receive little support (only from low-frequency words like *YACHT* or *WATCHTOWER*) but *CAT* is a high-frequency word (and many words contain the pattern "CAT"), "A" would receive strong feedback in that context.

In the modular/autonomous approach, direct feedback is posited to be unnecessary, and detrimental. (The claim is that once top-down and bottom-up inputs are mixed, the system risks hallucinating, since it can no longer distinguish activations driven by bottom-up inputs vs. top-down influences.) The reason feedback is argued not to be necessary is that decisions could be made outside the primary processing pathway, as depicted on the right side of Figure 2. Here, as in the Merge model (Norris et al., 2000), the Sublexical layer is duplicated as a special-purpose set of sublexical decision nodes that receive input from both bottom-up and top-down sources. Crucially, the activations in the Sublexical and Lexical layers only send and receive activation in the feedforward direction, protecting lower levels from top-down contamination. The ultimate claim is that such an architecture can simulate anything a feedback (interactive) architecture can, with special-purpose decision paths generating context-specific metalinguistic decisions. Norris et al. (2000) argued that a system without feedback is simpler than a system with feedback (a claim we will revisit in the Conclusions), and therefore one should prefer the modular/autonomous architecture if the two systems are equally capable of simulating human performance.

Furthermore, Norris et al. (2000) have argued that feedback cannot possibly improve a system's performance in any way. They argue that the best that a spoken word recognition system can possibly do is activate the sublexical forms (phonemes) that best correspond to the input and then the lexical forms that best correspond to the activated phonemes. This idea is related to the *data processing inequality* theorem, which holds that the information in a signal cannot be *increased* through subsequent manipulation. However, while it is certainly the case that the information in the signal cannot be increased, it is possible for systems to *use* information differently (e.g., via different implicit or explicit decision policies) or for systems to perform noise reduction, which would have clear benefits. So: is it possible that feedback allows a system to make qualitatively different use of information *or* effectively to reduce noise (possibly by enhancing coherent patterns in signals)?

Magnuson, Mirman, Luthra, Strauss, and Harris (2018) found support for such possibilities in the form of comprehensive demonstration proofs that feedback improves word recognition. They measured word recognition accuracy and recognition time for words presented to TRACE with or without feedback enabled, with clear inputs and then progressively noisier inputs. With clear or noisy inputs, recognition times were faster on average with feedback than without. Accuracy was also substantially higher with feedback than without as noise was added to inputs.

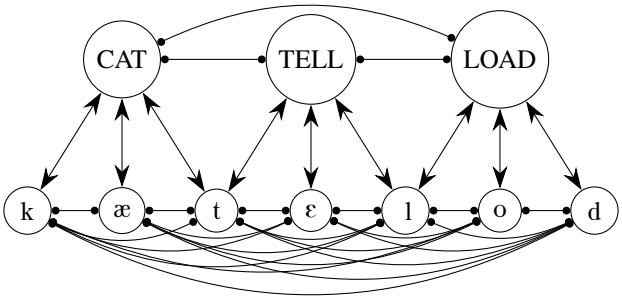

Figure 3: Interactive activation example. Arrows denote excitatory connections (7 input phonemes feed forward to 3 words, which send feedback to constituent phonemes). Edges with bulb connectors indicate lateral inhibition links within layers.

While recognition times were faster with feedback than without for most words at each level of noise (including zero noise), some items were recognized more quickly without feedback. (Note that such comparisons can only include words meeting the recognition definition both with and without feedback.) Whether a word was recognized more quickly or more slowly with feedback depended on a variety of factors, such as the makeup of a word's similarity neighborhood. But what drives the propensity for faster, more accurate processing with feedback than without?

Consider Figure 3, which provides a schematic of a very simple interactive activation network for modeling spoken word recognition with just seven phonemes and three words (*CAT, TELL, LOAD*). In this simple network, with no ability to encode temporal order, the input /kæt/ would strongly activate *CAT* and also weakly activate *TELL* (since /t/ connects to both words). Now consider what happens when noise is added to the initial input. Feedback will not simply reinforce the original feedforward pattern (of signal plus noise). For example, given the input /kæt/ + *noise*, noise will likely provide more bottom-up input to *TELL* than the signal alone did. However, noise will not be reinforced by feedback to the same extent as coherent input patterns that map onto words. Even slightly greater activation of *CAT* relative to *TELL* from the coherent input pattern to which noise has been added will allow *CAT* to inhibit *TELL* significantly (depending on exact parameters; sufficient noise will overwhelm the signal). Over iterations, this will lead *CAT* to send increasingly more feedback to its constituent phonemes relative to the feedback *TELL* sends to its constituent phonemes; *coherent patterns in the input will be reinforced to the detriment of nodes activated primarily by noise*, rather than simple "reinforcement of the status quo." In other words, the joint probability of {k, æ, t} embodied in the lexicon drives selective reinforcement of coherent lexical (and sublexical, phonotactically probable) patterns in the input. This will also drive faster and stronger activation of words consistent with the input, as Magnuson et al. (2018) reported.

However, Norris and Cutler (2021) claim that the apparent advantages for feedback reported by Magnuson et al. (2018) are due to an artifact: "The effect of noise was simulated by adding a constant amount of noise to a decision process – the Luce choice rule – operating on the output of the network ... In other words, because of a workaround in the model, simulations using TRACE can give the impression that feedback can improve performance" (Norris & Cutler, 2021, p. 3). We note, however, Magnuson et al. (2018) specify (p. 3) that they added Gaussian noise independently to each input element (following the procedure of McClelland, 1991). They subsequently converted activations to response probabilities in accordance with the procedures used by Frauenfelder and Peters (1998) to simulate lexical decisions (pp. 118-119), which were based on procedures used by McClelland and Rumelhart (1981) and McClelland and Elman (1986) for letter and phoneme recognition. While it seems unlikely that converting to response probabilities could change the rank ordering of recognition times, we confirm this by replicating the results of Magnuson et al. (2018) with raw activations. We go beyond their analyses and examine whether phoneme activations indicate that feedback selectively reinforces lexically-coherent input patterns.

Simulations

We use the same approach as Magnuson et al. (2018) to compare word recognition accuracy and recognition time in TRACE under increasing levels of noise with and without feedback – with the important difference that we use raw activations rather than response probabilities.

Procedure

We conducted simulations using jsTRACE, a recent reimplementation of TRACE in JavaScript (Magnuson, Curtice, Grubb, Crinnion, & Sossounov, in preparation). We used the default *slex* TRACE lexicon, consisting of 212 words (as well as the "silence" word used to represent a state of no input; the silence word was not included in analyses). We used three levels of feedback (0.00, 0.015, and 0.03, the last being the default level with small lexicons in TRACE). Inputs were default TRACE inputs (distributed representations of pseudo-spectral transformations of acoustic-phonetic features over time). We combined each level of feedback with seven levels of Gaussian noise (with mean of zero and standard deviation ranging from 0.0 to 1.5 in steps of 0.25). A value sampled from the distribution was added independently to each cell of the input matrix prior to the simulation. Any negative input values were replaced with zero. To ensure that results under noise were robust, we conducted 10 simulations of every word in the lexicon at all levels of noise greater than zero. We allowed simulations to run for 100 time steps (cycles) in TRACE.

Decision policy Note that the proposal that the best a word recognition system can do is choose the word with highest activation (Norris & Cutler, 2021) does not specify *when* a decision should be made. We cannot simply take the time of peak activation, as a target's activation may continue increasing for some time after it has become the dominantly activated item (potentially resulting in longer recognition times for words that are more strongly activated, relative to a word that might have an earlier but lower peak). A further complication is that if we were to present the model with nonword (or novel word) inputs, simply taking the word with maximum activation as the winner would lead to erroneous "recognition." We followed the example of Magnuson et al. (2018) and used a simple threshold-based policy, where a correct identification is defined as the target reaching or exceeding that threshold and no other item reaching it. Recognition time is the cycle where the target's activation first reaches or exceeds the threshold. We first identified the activation threshold that would maximize accuracy for zero feedback without noise; this was 0.4. We then applied that threshold to every simulation (that is, at every level of feedback and noise). Note that any potential bias in this policy favors simulations without feedback, since the threshold optimizes accuracy with zero noise and zero feedback. *Crucially, all analyses were applied to raw activations. We did not transform activations to response probabilities.*

Results

In Figure 4, we see a clear replication of the results of Magnuson et al. (2018) based on raw activations. Recognition time (left panel, only including recognized words) is faster with feedback than without at every level of noise (up to $sd = 1.25$ or 1.5, where there were too few correct trials with feedback off to make meaningful comparisons). Feedback also yields higher accuracy (right panel) once noise is added.

In Figure 5, we plot recognition times at each noise level for words that were correctly recognized with feedback at the default value of 0.03 and without feedback. Again, we repeated the simulation of each word 10 times at each level of noise greater than zero. A clear, consistent advantage is observed for the majority of words at each level of noise.

Contra assertions by Norris and Cutler (2021), feedback promotes more robust word recognition performance. Thus, the results of Magnuson et al. (2018) were not an artifact of using response probabilities rather than raw activations. TRACE recognized words more quickly with feedback than without at every level of noise (including no noise), and feedback promoted higher accuracy as more noise was added. How would this be possible if feedback simply reflected signal and noise equally, as Norris and Cutler (2021) claim? We can explore this by examining how phoneme and word activations change as noise is added with and without feedback.

First, consider phoneme activations. In Figure 6, we plot the mean activation of each input phoneme (solid red lines) averaged over all items (all words in *slex*) along with the mean activation of the next-most-activated phoneme at that position (dashed blue lines). Although the maximum word length is 9, the number of words contributing to means for later phonemes decreases, since most words are shorter. There is a clear pattern: with feedback (right panels), there is greater separation between Target phonemes and Next phonemes than

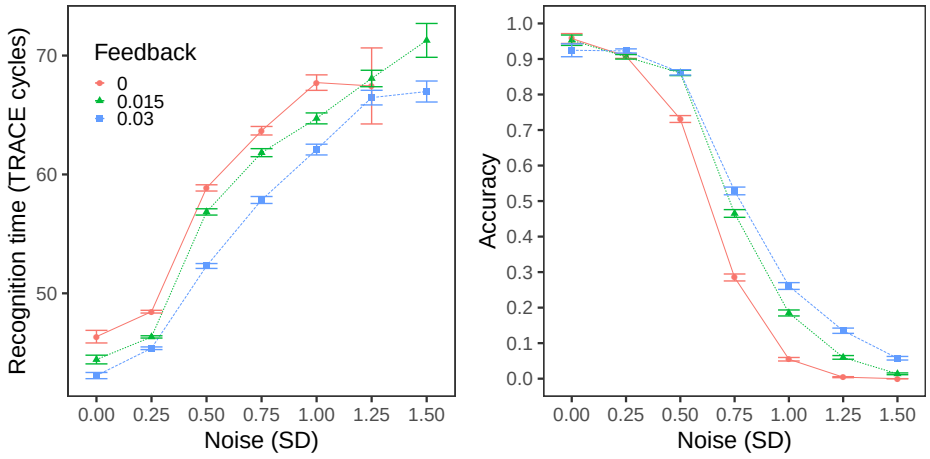

Figure 4: Replication of Magnuson et al. (2018) using activations instead of response probabilities. Each point represents the outcome of simulating every word in the 212-word *slex* lexicon, with 10 simulations conducted with each word at each noise level greater than zero. The threshold was set to 0.4, which maximized accuracy without feedback and with noise set to zero.

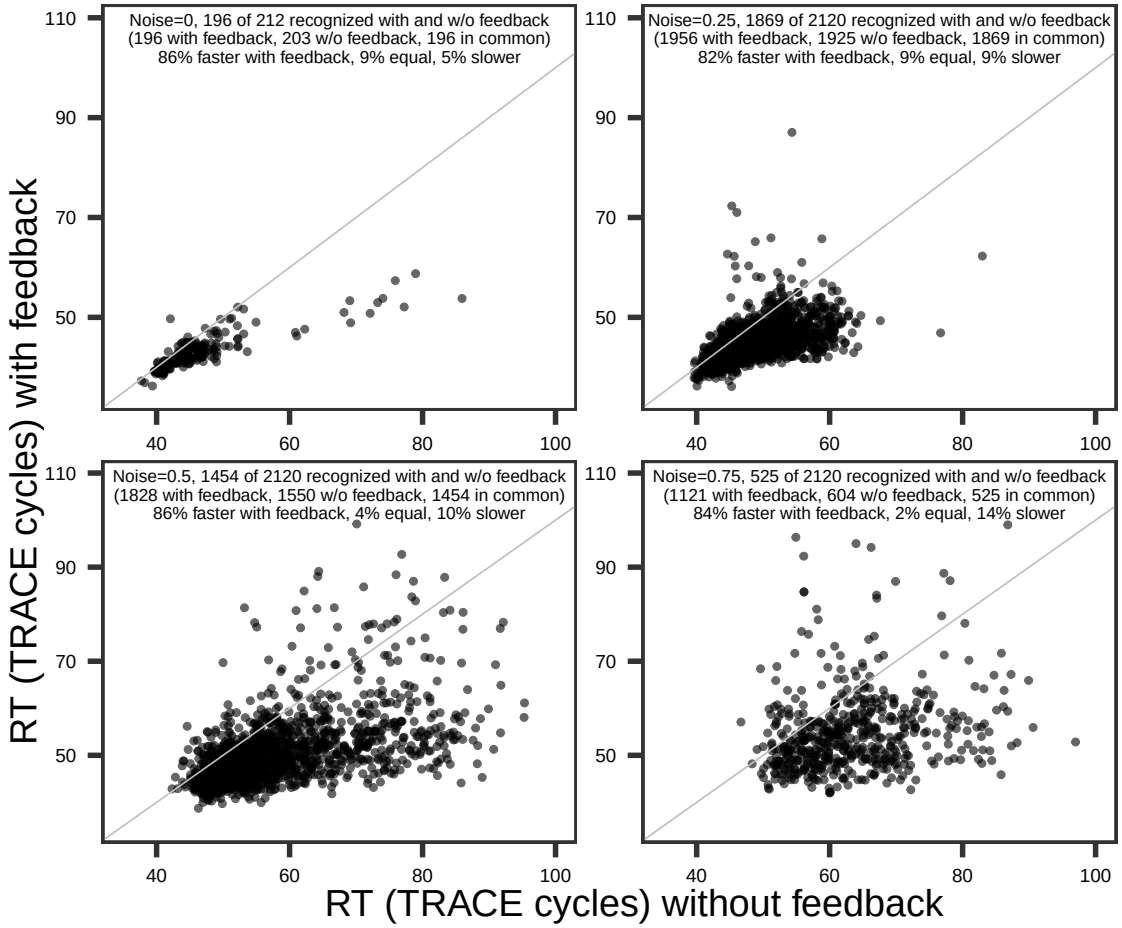

Figure 5: Comparing activation-based recognition time with feedback (set to 0.03) and without feedback (set to 0.0). Only words that were correctly recognized both with and without feedback are included. Without noise (top left), we conducted one simulation per word with and without feedback. At noise levels greater than 0, there were 2120 simulations (10 repetitions of each word with Gaussian noise added to input). Items classified as "faster" were recognized more quickly (reached the threshold more quickly) with feedback than without and are below the identity line; "equal" items are on the identity line; "slower" items reached the threshold later with feedback than without and are above the identity line.

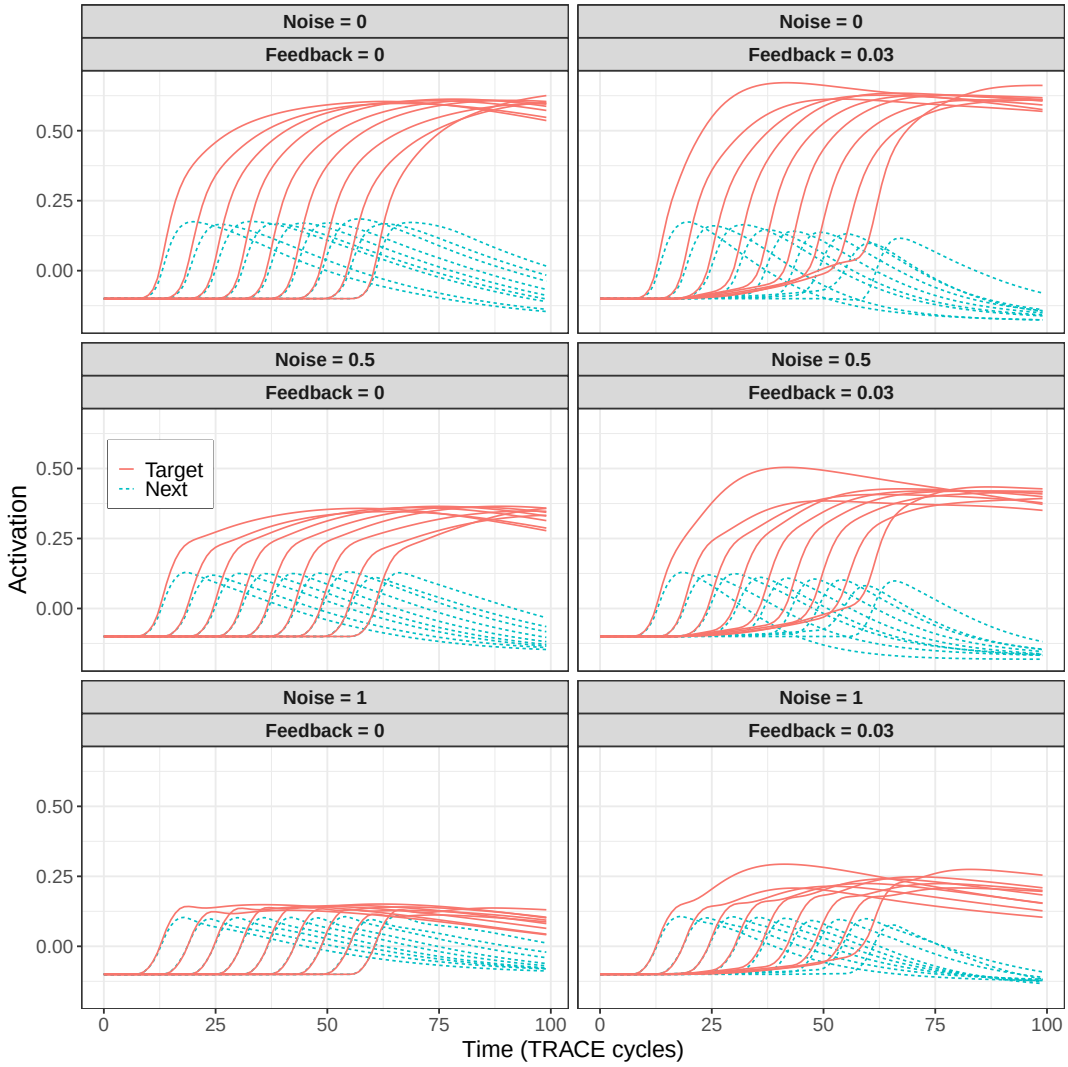

Figure 6: Phoneme activations over time for words at two levels of feedback and three levels of noise averaged over all 212 words in the *slex* lexicon. In each plot, the activations of the phoneme corresponding to the input (solid red lines) and the next-most activated phoneme (dashed blue lines) are plotted. Later phonemes are based on fewer words (those up to nine phonemes in length). Target phoneme activations are generally higher and Next activations generally lower with feedback on.

without (left panels). This is a result both of greater Target phoneme activation with feedback vs. without and modestly lower activation of Next phonemes with feedback vs. without (mainly observable as sharper, lower peaks for Next phonemes with feedback, indicative of greater lateral inhibition from Target phonemes). If signal and noise were equally reinforced by feedback, we would expect to see similar amplification of both Target and Next phonemes with feedback. Consistent with our discussion of Figure 3, feedback selectively boosts those phonemes that constitute lexically-coherent patterns – that is, series of phonemes that constitute words.

Discussion

The simulations we presented demonstrate clear benefits of feedback in the interactive framework. Feedback promotes faster recognition time and makes the model robust against noise, as demonstrated by preservation of accuracy with feedback as compared to without. Contra assertions made by Norris and Cutler (2021), this is not an artifact of using response probabilities rather than raw activations. Magnuson et al. (2018) used the standard approach of using response probabilities, but here we used raw activations. Our results were extremely similar (see Figures 4 and 5), with faster recognition with feedback than without at every noise level and substantially higher accuracy with feedback than without with moderate to high levels of noise (SD ≥ 0.5).

We extended the earlier work by tracking phoneme activations. In Figure 6, we plotted the mean activations of position-specific phonemes from target words and the next-most-activated phonemes at those positions. Relative to the no-feedback simulations, feedback promoted greater differences between target phonemes and next-most-activated phonemes.

Feedback drives selective reinforcement of lexically-probable patterns, which leads to faster recognition times as well as higher accuracy with feedback than without at multiple levels of noise added to inputs. Examining the impact of feedback on the phoneme level (Figure 6) reveals that feedback drives a greater separation between input phonemes and next-most-activated phonemes when noise is added to inputs. Feedback provides an effective and efficient mechanism to distinguish signal from noise.

Conclusions

To return to the larger theoretical issues we touched on in the Introduction, let us consider again the claims that feedback (a) cannot improve perceptual processing, (b) entails hallucination, (c) is more complex than a system without feedback, and (d) is not necessary. We have shown here that feedback *does* improve perceptual processing,[1] replicating Magnuson et al. (2018) and extending their results with our new examination of selective reinforcement of lexically-probable signals over noise. Previous papers have addressed hallucination, beginning with the original TRACE paper (McClelland & Elman, 1986), where feedback was set to a level that promotes bottom-up priority.

The final two claims are closely linked: We agree that many apparent top-down effects can be simulated by special-purpose, non-perceptual pathways (as in the right side of Figure 2). However, if the two architectures can both account for relevant data (though note that this has not been fully established), then it only matters that feedback is not necessary if there is a reason to prefer a system without feedback. Thus, a crucial issue is whether an interactive framework (as in Figure 2) is more complex than an analogous modular/autonomous system.

Ironically, one of the criticisms Norris et al. (2000) leveled against feedback in TRACE is that it appears only to be there to simulate top-down effects, while serving no functional purpose. This assertion was based on the finding that motivated the Magnuson et al. (2018) simulations: a report from Frauenfelder and Peters (1998) that in simulations comparing TRACE with feedback on vs. off, using 21 carefully selected words, about half the words were recognized more quickly with feedback, and about half were recognized more quickly *without* feedback. While Frauenfelder and Peters (1998) had good reasons for selecting their items, their results do not generalize beyond those items (as our simulations and those of Magnuson et al., 2018, demonstrate, while including their words and many more). The irony is that the sublexical decision *cul-de-sac* in the modular/autonomous architecture *serves no functional purpose*. While Norris et al. (2000) argue that it serves a necessary role as a readout or decision layer, this is simply a stipulation. An unspecified process still must "read" the activations and apply a decision policy. We can make the same stipulation about any layer in any model (i.e., we can stipulate that a layer is accessible to decision processes).

The interactive architecture is intuitively simpler; it requires one layer fewer than the modular/autonomous architecture. As Magnuson et al. (2018) discuss, this also implies that the modular/autonomous system would require more nodes and connections than a corresponding interactive system. Thus, there is no basis for claiming that a system with feedback is more complex than one without feedback.

Of course, these issues have been argued extensively without apparent progress by proponents of interactive frameworks (e.g., Magnuson, McMurray, Tanenhaus, & Aslin, 2003a, 2003b; McClelland, 1991, 2013; Samuel, 1997; Samuel & Pitt, 2003) and proponents of modular/autonomous frameworks (e.g., McQueen, 2003; Norris & Cutler, 2021; Norris et al., 2000; Norris, McQueen, & Cutler, 2016). Our goal here has been to address the key issue of *how* feedback improves perception, in direct response to claims made by Norris and Cutler (2021). In light of the demonstration proof that feedback promotes accuracy and speed under noise and the computational case we have made for *how* feedback achieves this, we have advanced the theoretical debate. A challenge is how to resolve the debate using experimental results.

Notably, empirical support for feedback is growing. This includes neural evidence consistent with feedback (e.g., Getz & Toscano, 2019; Gow & Olson, 2015; Gow, Segawa, Ahlfors, & Lin, 2008; Myers & Blumstein, 2008; Noe & Fischer-Baum, 2020) as well as new findings using the classic *lexically-mediated compensation for coarticulation* (LCfC) paradigm of Elman and McClelland (1988). The two camps have accepted LCfC as a crucial test for feedback (e.g., McQueen, Jesse, & Norris, 2009; Pitt & McQueen, 1998). While McQueen et al. (2009) failed to replicate the findings of Magnuson et al. (2003b) using their original materials, Luthra et al. (2021) noted that a precondition for observing LCfC is having materials that can drive both robust phoneme restoration and robust compensation for coarticulation – but few previous LCfC studies had actually established that their materials elicited these effects separately before combining them for the LCfC paradigm. When Luthra et al. (2021) first pretested items to ensure that they yielded both effects separately, they found robust LCfC results (with an initial sample of 40 participants and a direct replication with a second sample of 40 participants).

Theoretical, computational, neural, and behavioral findings are all converging towards the conclusion that feedback provides an efficient mechanism for using prior probabilities (embodied in lexical knowledge, in the case of word recognition) to promote fast, noise-resistant processing.

Acknowledgments

This research was supported by NSF BCS-PAC 1754284, NSF BCS-PAC 2043903, and NSF NRT 1747486 (PI: JSM). SL was supported by an NSF Graduate Research Fellowship. AMC and PG were supported by NIH T32 DC017703 (E. Myers and I-M. Eigsti, PIs). We thank Arty Samuel, Thomas Hannagan, and Rachel Theodore for helpful discussions.

[1] A reviewer pointed out that our RT analyses use a decision threshold. However, top-down impact on raw activations is clear in Figure 6.

References

Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, *27(2)*, 143–165.

Fenske, M., Aminoff, E., Gronau, N., & Bar, M. (2006). Top-down facilitation of visual object recognition: Object-based and context-based contributions. *Progress in Brain Research*, *155*, 3-21. doi: 10.1016/S0079-6123(06)55001-0

Frauenfelder, U. H., & Peters, G. (1998). Simulating the time course of spoken word recognition: An analysis of lexical competition in TRACE. In *Localist connectionist approaches to human cognition.* (pp. 101–146). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6(1)*, 110–125. doi: https://doi. org/10. 1037/0096-1523. 6. 1. 110

Getz, L. M., & Toscano, J. C. (2019). Electrophysiological evidence for top-down lexical influences on early speech perception. *Psychological Science*, *30*(6), 830–841.

Gow, D., & Olson, B. (2015). Lexical mediation of phonotactic frequency effects on spoken word recognition: A granger causality analysis of mri-constrained meg/eeg data. *Journal of Memory and Language*, *82*, 41–55.

Gow, D., Segawa, J., Ahlfors, S., & Lin, F.-H. (2008). Lexical influences on speech perception: A granger causality analysis of meg and eeg source estimates. *NeuroImage*, *43*, 614–623.

Luthra, S., Peraza-Santiago, G., Beeson, K., Saltzman, D., Crinnion, A. M., & Magnuson, J. S. (2021). Robust lexically-mediated compensation for coarticulation: Christmash time is here again. *Cognitive Science*, *45*(4), e12962.

Magnuson, J. S., Curtice, A., Grubb, S., Crinnion, A. M., & Sossounov, N. (in preparation). jsTRACE: A javascript reimplementation of the TRACE model of speech perception and spoken word recognition.

Magnuson, J. S., McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2003a). Lexical effects on compensation for coarticulation: A tale of two systems? *Cognitive Science*, *27*, 795–799.

Magnuson, J. S., McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2003b). Lexical effects on compensation for coarticulation: The ghost of christmash past. *Cognitive Science*, *27*, 285–298.

Magnuson, J. S., Mirman, D., Luthra, S., Strauss, T., & Harris, H. (2018). Interaction in spoken word recognition models: Feedback helps. *Frontiers in Psychology*, *9*(369). doi: 10. 3389/fpsyg. 2018. 00369

McClelland, J. L. (1991). Stochastic interactive processes and the effect of context on perception. *Cognitive Psychology*, *23*, 1-44. doi: 10. 1016/0010-0285(91)90002-6

McClelland, J. L. (2013). Integrating probabilistic models of perception and interactive neural networks: a historical and tutorial review. *Frontiers in Psychology*, *4*, 503. doi: 10. 3389/fpsyg. 2013. 00503

McClelland, J. L., & Elman, L., J. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1-86. doi: 10. 1016/0010-0285(86)90015-0

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375–407. (Place: US Publisher: American Psychological Association) doi: 10. 1037/0033-295X. 88. 5. 375

McQueen, J. M. (2003). The ghost of christmas future: didn't scrooge learn to be good? commentary on magnuson, mcmurray, tanenhaus, and aslin. *Sci*, *27*, 795–799. doi: 10. 1207/s15516709cog2705$_6$

McQueen, J. M., Jesse, A., & Norris, D. (2009). No lexical-prelexical feedback during speech perception or: Is it time to stop playing those christmas tapes? *Journal of Memory and Language*, *61*(1), 1–18.

Myers, E., & Blumstein, S. (2008). The neural bases of the lexical effect: An fmri investigation. *Cerebral Cortex*, *18*(2), 278–288.

Noe, C., & Fischer-Baum, S. (2020). Early lexical influences on sublexical processing in speech perception: Evidence from electrophysiology. *Cognition*, *197*, 1–14.

Norris, D., & Cutler, A. (2021). More why, less how: What we need from models of cognition. , *213*, 104688. Retrieved from https://www.sciencedirect.com/science/article/pii/S001 doi: https://doi.org/10.1016/j.cognition.2021.104688

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: feedback is never necessary. *Behavioral and Brain Sciences*, *23(3)*, 299-370. doi: 10. 1017/s0140525x00003241

Norris, D., McQueen, J. M., & Cutler, A. (2016). Prediction, Bayesian inference and feedback in speech recognition. *Language, Cognition and Neuroscience*, *31*(1), 4–18.

Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation meditated by the lexicon? *Journal of Memory and Language*, *39*, 347–370.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus materials. *Journal of Experimental Psychology*, *81*, 275-280.

Rubin, P., Turvey, M. T., & Van Gelder, P. (1976). Initial phonemes are detected faster in spoken words than in spoken nonwords. *Perception & Psychophysics*, *19*, 394-398.

Samuel, A. G. (1997). Lexical activation produces potent phonemic percepts. , *32*, 97-127.

Samuel, A. G., & Pitt, M. A. (2003). Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language*, *48*(2), 416–434.

Selfridge, O. G. (1955). Pattern recognition and modern computers. In *Proceedings of the March 1-3, 1955, Western Joint Computer Conference* (p. 91–93). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/1455292.1455310

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, *1*, 59-85.