Published in final edited form as:

Lang Cogn Neurosci. 2016 ; 31(7): 869–875. doi:10.1080/23273798.2016.1192656.

Using effective connectivity analyses to understand processing architecture: Response to commentaries by Samuel, Spivey and McQueen, Eisner and Norris

David W. Gow Jr.^{1,2,3,4,*} and Bruna B. Olson⁵

¹Neuropsychology Laboratory, Massachusetts General Hospital, 100 Cambridge St., rm 2030, Boston, MA 02114

²Department of Psychology, Salem State University, 352 Lafayette St., Salem, MA 01970

³Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, 149 Thirteenth St., S2301, Charlestown, MA.02129

⁴Harvard-MIT Division of Health Sciences and Technology, 77 Massachusetts Ave., E25-519, Cambridge, MA 02139

⁵University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655

Keywords

speech perception; semantic context effect; sentential context effect; modularity; top-down processing; effective connectivity; localization; task effect; interactive processing

In Gow and Olson (2015b), we used a powerful new effective connectivity technique to tackle long-standing issues about the role of interactive processes in speech perception. As commentaries by Spivy (in press), Samuel (in press) and McQueen et al. (in press) note, our empirical approach addresses fundamental inferential limitations of commonly used behavioural and BOLD imaging techniques for determining the functional architecture of cognitive processes. The consensus breaks down on the question of whether our technique resolves the modularity debate in speech perception and spoken word recognition. **Spivey** and **Samuel** describe these results as decisive on the issue of whether some higher level representations influence phonetic processing, while **McQueen et al.** argue that the same results have no bearing on the interactivity issue.

Our key finding was that strong behavioural evidence of sentential influences on the explicit categorization of ambiguous phonemes is accompanied by a pattern of increased directed and mediated influence of brain regions implicated in lexical representation on regions implicated in acoustic-phonetic processing during the 100-500 ms following the onset of ambiguous items. We interpret this result as compelling evidence that there are interactive

*Corresponding author: David W. Gow Jr., Neuropsychology Laboratory, Massachusetts General Hospital, Rm. 2030, 100 Cambridge St., Boston, MA 02114, Tel: 617-726-6143, gow@helix.mgh.harvard.edu, bruna.olson@umassmed.edu

The authors and their institutions have no conflicts of interest related to this work.

processes in speech perception. The commentaries point out that this interpretation depends on a clear understanding of task and stimulus demands, the accurate interpretation of the functional significance of critical brain regions, and a strong linking hypothesis to relate effective connectivity patterns to interactive processing.

Complexity and Task Demands

Spivey highlighted the “painfully realistic” complexity of these results. We feel his pain. The 32 region of interest processing network that our data-driven techniques identify is of typical size compared to networks we have found in analyses of data collected in phoneme identification, picture verification and lexical decision tasks (cf. Gow & Nied, 2014; Gow & Olson, 2015a; Gow & Segawa, 2009; Gow, Segawa, Ahlfors, & Lin, 2008). To our knowledge, no implemented psycholinguistic models of spoken word recognition even approach this level of complexity.

This complexity is both daunting and reassuring. It is daunting in that it suggests that our models need to be significantly more complex, and as **Spivey** notes, will probably depend on more powerful modelling methods including dynamical systems approaches to understand how the rich pattern of interaction we see supports language understanding. This complexity is reassuring though, because it suggests a context for understanding conflicting results concerning the modular versus interactive nature of speech processing. Like **Samuel**, we believe that behavioral paradigms that exploit the indirect effects of interactive processing including work on exploring selective adaptation modulated by the Ganong effect (Samuel, 2001) and compensation for coarticulation by perceptually restored phonemes (Elman & McClelland, 1988; Magnuson, McMurray, Tanenhaus, & Aslin, 2003a; Magnuson, McMurray, Tanenhaus, & Aslin, 2003b; Samuel & Pitt, 2003) can provide direct evidence of interactive processing in speech perception. Unfortunately, these indirect demonstrations tend to produce small effect sizes, and may be sensitive to a variety of task parameters (Magnuson et al., 2003b; McQueen, Jesse, & Norris, 2009; Pitt & McQueen, 1998; Samuel & Pitt, 2003). This presents a problem. How is it possible to design an experiment that addresses diverse known (and potential unknown) task factors while retaining sufficient design power to capture a subtle behavioural effect? We believe that our effective connectivity approach takes an important step towards resolving this problem by providing concurrent measures of all of the processing dynamics that support performance.

Consider **McQueen et al.**'s objection that the retrospective phoneme verification task does not reflect “normal processing” because it introduces additional task-specific processes. We agree that it entails some task-specific processes, but suggest that these are empirically separable from the underlying components of normal processing. In addition to the memory processes involved in normal incremental sentence processing (Caplan & Waters, 2013), subjects must maintain a representation of the sentence for a short period of time (less than 2 seconds from sentence onset), and at probe presentation interpret the visual letter prompt and compare it with a phonemically segmented representation of the sentence. Given the short time that information needed to be held, and the concurrent processing demands of auditory sentence processing, it is unlikely that subjects engaged in concurrent rehearsal during the analysis window (Barrouillet & Camos, 2012). We designed our retrospective

phoneme verification task with the intent to delay most metalinguistic processing (e.g. target encoding, segmentation and search) until after the analysis window was over. McQueen et al. propose an alternative processing model in which subjects monitor for the four potential probe letters during sentence processing. This would require subjects to hold all four probes in memory, make continuous phoneme judgements, and hold the results of those running judgements in a buffer until the probe is presented. We tried to minimize the likelihood of this strategy by using four different targets, randomizing target position, and employing an equal number of target-present and target-absent trials to minimize guessing or response bias. We cannot rule out McQueen et al.'s model, but it seems unlikely. However, the robust effect of sentence context on performance strongly suggests that any strategic processes were superimposed on processes associated with a fundamental language process — the interpretation of meaningful sentences.

The critical question is whether the primary result, increased influence by the supramarginal gyrus (SMG) on activation in the posterior superior temporal gyrus (pSTG) for experimental versus baseline trials, was the result of top-down influences on speech processing or task-induced demands. The top-down interpretation is consistent with the predictions of our neurofunctional model of lexically mediated speech perception (Gow, 2012), as well as a series of effective connectivity results we have seen in previous explorations of lexical effects on speech processing (Gow & Nied, 2014; Gow & Olson, 2015a; Gow & Segawa, 2009; Gow et al., 2008).

Our claims for top-down influences on speech perception are unaffected if subjects were engaging in continuous phoneme monitoring. If this were the case, both experimental and baseline trials would both contain potential phoneme targets in the same position. This would introduce comparable loads on attention and memory as subjects were called upon to mark and remember target phonemes. Under the model of phoneme monitoring offered by Norris, McQueen and Cutler (2000), lexical and phonemic analyses would be integrated at post-perceptual phoneme decision stage. McQueen et al. do not suggest that the pSTG plays a role in phonemic decision-making independent of perceptual analysis. BOLD imaging studies suggest that phonemic decision making processes may be localized in the left pars triangularis, and that response conflict associated with difficult discriminations is mediated by the anterior cingulate (Blumstein, Myers, & Rissman, 2005; Hasson, Skipper, Nusbaum, & Small, 2007; Myers, Blumstein, Walsh, & Eliassen, 2009). The fact that neither of those regions emerged as a region of interest during our analysis window suggests that any potential phonemic decision processes took place after the interval we investigated.

The only systematic difference between baseline and experimental trials was that experimental trials contained a word with a phonetically ambiguous onset. While most participants reported they were unaware of this manipulation, it is possible that odd sounding onsets cued subjects to remember a likely target. The paradigm's online processing demands during sentence presentation preclude simultaneous with articulatory rehearsal (Barrouillet & Camos, 2012). However, subjects could maintain a representation of the target phoneme using an attention-based mechanism (cf. Baddeley, 2000a; Barrouillet & Camos, 2012; Vergauwe & Cowan, 2015). BOLD imaging studies show that this mechanism relies on interactions between anterior prefrontal and inferior parietal brain areas (Gruber,

2001; Gruber & von Cramon, 2003; Trost & Gruber, 2012). While such interactions could potentially indirectly influence SMG interactions with pSTG, we found no evidence of increased interaction between anterior prefrontal and inferior parietal areas in experimental versus baseline trials. Two portions of the right superior frontal gyrus, a region believed to play a role in voluntary attentional control (cf. Hopfinger, Buonocore, & Mangun, 2000), showed increased influence on pSTG activation in experimental trials. This may be consistent with strategic focus on acoustic-phonetic representation in experimental trials.

In summary, the retrospective phoneme verification task enlists a large network of brain regions devoted to a combination of task-specific and more naturalistic spoken sentence comprehension processes. The primary neural result, increased influence of the SMG on pSTG for experimental trials, is consistent with our neurofunctional model of lexical processing and previous experimental studies of potential lexical influences on speech perception. A review of models of phoneme decision-making, memory and attention does not reveal any task-induced processes that can account for our primary neural result.

Sentential versus Lexical Effects

Samuel is in agreement that the current results demonstrate top-down influences on acoustic-phonetic processing, but questions whether they reflect sentential (structural) or lexical (associative) effects. His analysis is based on: (1) our focus on the direct effects of lexical brain regions that might mediate sentential effects rather than a well-identified brain region whose activation indexes sentential representation, (2) our use of sentence contexts that often contained strong lexical associates of the intended interpretation of ambiguous target words, and (3) a review of behavioural evidence for dissociation between lexical and sentential context effects. We find this argument persuasive on all counts. Indeed, as **Spivey** points out, there is a large behavioural literature suggesting that speech perception is influenced by diverse linguistic and contextual factors. While our design may have inadvertently highlighted the role of associative processes, we recognize that even simple sentence level stimuli potentially engage many forms of constraint on speech perception. We hope to systematically explore these mechanisms in future work.

Interpreting Localized Activation

All three commentaries raise the critical issue of functional localization. This is indeed a crux of our approach. While the overall patterns of effective connectivity we see across studies strongly suggests that interactivity is a fundamental property of neural networks, the ability to infer an explicit description of task-specific processing architecture depends on the accurate characterization of the functional significance of activation in key brain regions (Gow & Caplan, 2012a, 2012b). **McQueen et al.** and **Spivey** both observe that the individual brain regions we focus on have been implicated in a range of linguistic and cognitive processes. For McQueen et al., this severely limits our ability to make strong claims about their roles in processing. In contrast, Spivey takes the position shared by a number of researchers (Anderson, 2015; Friston et al., 1996; Gow & Caplan, 2012a) that network level analyses provide an important tool for contextualizing activation and reconciling seemingly divergent localization findings.

McQueen et al. raise concerns about the functional characterization of activation in pSTG and supramarginal gyrus (SMG). In both cases, our functional interpretation of activation is based on a strong convergence of evidence from multiple techniques across a large number of studies. Direct cortical surface recordings show that pSTG encodes contrastive phonetic features (Mesgarani, Cheung, Johnson, & Chang, 2014; Steinschneider et al., 2011). MEG and EEG results show that it is sensitive to acoustic-phonetic manipulations (Gage, Poeppel, Roberts, & Hickok, 1998; Naatanen, Paavilainen, Rinne, & Alho, 2007; Phillips et al., 2000). Furthermore, BOLD imaging results comparing activation by speech and closely matched nonspeech auditory analogs reliably implicate the bilateral posterior superior temporal gyrus in speech sound processing (see reviews by Hickok & Poeppel, 2007; Turkeltaub & Coslett, 2010). A similar convergence of results motivates our dual lexicon model (Gow, 2012), which suggests that a dorsal lexicon in the SMG mediates the two-way mapping between acoustic-phonetic and articulatory representation. This model reflects evidence including: (1) patterns of lexically-conditioned disruption of reproduction conduction aphasia following SMG damage (Romani, Galluzzi, & Olson, 2011; Shallice, Rumati, & Zadini, 2000), (2) systematic variation between SMG gray matter density and vocabulary size (Lee et al., 2007; Richardson, Thomas, Filippi, Harth, & Price, 2010), and (3) modulated SMG activation as a function of word learning (Cornelissen et al., 2004) and manipulations of lexical properties including lexical neighbourhood density and competitor environment (Peramunage, Blumstein, Myers, Goldrick, & Baese-Berk, 2011; Prabhakaran, Blumstein, Myers, Hutchison, & Britton, 2006; Righi, Blumstein, Mertus, & Worden, 2009).

McQueen et al. note that pSTG activation has been implicated in a range of non-acoustic-phonetic processes including sensorimotor integration of speech, auditory imagery, and verbal working memory. It is possible that pSTG participation in these tasks reflects co-localized functions with no acoustic-phonetic component. However, it seems to us that it is more parsimonious to suggest that all of these tasks draw on a common representation of speech sounds. The sensorimotor integration of speech requires a sensory representation. Similarly, evidence for behavioural interference and facilitation effects between audition and auditory imagery (Farah & Smith, 1983; Okada & Matsuoka, 1992) and evidence for widespread overlap in patterns of BOLD activation for heard and closely matched imagined auditory stimuli (Yoo, Lee, & Choi, 2001; Zatorre & Halpern, 2005) suggest imagery and auditory processing rely on some common processes and representations that are partially localized in the superior temporal gyrus (see review by Hubbard, 2010). We see the same partial overlap between speech processing and working memory networks. In our model (Gow, 2012; Gow & Caplan, 2012b), acoustic-phonetic representations in STG are normalized through a process of interactions with higher-level representations. These normalized representations directly address the lack of invariance problem at the heart of speech perception, and provide a strong candidate representation for the phonological similarity effect, which is a core working memory phenomenon (cf. Baddeley, 1986; Baddeley, 2000b; Larsen, Baddeley, & Andrade, 2000). The argument that a single brain region may provide common representations serving multiple functions by participating in multiple networks provides a straightforward account of the results that McQueen et al. cite in their discussion of the supramarginal gyrus. Rauschecker and Scott (2009) examined nonhuman primates and suggested that the SMG plays essentially the same role that we

suggest it does -- as a mediator between representations of sound and articulation. All of the other references cited by McQueen et al. ((Jacquemot & Scott, 2006; Obleser & Eisner, 2009; Turkeltaub & Coslett, 2010) draw their conclusions about the role of SMG based on studies that used words or wordlike phonological patterns as stimuli. It is not surprising that the processing of such stimuli for any task would involve some form of lexical representation.

What are brain regions talking about?

McQueen et al. argue that even if our task reflected normal speech processing and the SMG and pSTG play the functional roles we ascribe to them, our results would not settle the argument over interactive processing because we do not know what kind of “interaction” our Granger effects reflect. There are several ways to characterize this interaction. At a formal level, Granger causation is a measure of information flow (technically, the transfer entropy of Gaussian variables) between stochastic processes (Barnett, 2009). At a very primitive physiological level, any interaction between brain regions can be described as the exchange of excitatory and inhibitory input that influence patterns of distributed neural activity.

McQueen et al. are primarily concerned with a different level of analysis. They want to know whether these Granger effects demonstrate that lexical representations influence the form of phonetic representation. Posterior STG activation reflects acoustic-phonetic representation, but we our neural analyses do not reveal the form of these representations. Our behavioural results show that context affects the explicit categorization of phonetically ambiguous segments, but they do not uniquely tie categorization judgments to phonetic representation. In all implemented computational models of speech recognition that we know of including TRACE, MERGE and Shortlist B (McClelland & Elman, 1986; Norris & McQueen, 2008; Norris et al., 2000), interaction between representations or processes produces changes in either the relative strength of competing localized representations (selection), or the pattern of distributed representations (modulation).

McQueen et al. suggest two alternative interpretations of our Granger effects. The first is that top-down feedback from SMG to pSTG reflects off-line lexically-mediated perceptual learning (Norris, McQueen, & Cutler, 2003). There is no evidence that participants in our study showed perceptual learning. Phonetically ambiguous segments only occurred in contexts in which lexical context supported both interpretations of ambiguous phonetic cues. Moreover, individual subjects classified the same ambiguous item (e.g. ^lɹʌsk) differently on first and second presentations almost 90% of the time. The second interpretation they suggest is that SMG to pSTG influences reflect attentional processes. As we have already noted, this interpretation is inconsistent with current understanding of the neural substrates of auditory attention.

The premise of McQueen et al.'s objections is that top-down processes cannot improve perceptual processes (Dennett, 1992; Norris et al., 2000). If the point of processing is to identify words and there is sufficient bottom-up support to identify a strong lexical candidate, they wonder what can be gained by modifying lower level speech representations through feedback. General systems theory and engineering provide several answers to this

question. TRACE (McClelland & Elman, 1986) relies on a form of positive feedback that is both well-understood and widely used as a mechanism to produce or explain increases in signal gain and the acceleration of natural and artificial processes (Aström & Murray, 2008). While Frauenfelder (Frauenfelder & Peeters, 1998) found no evidence for these advantages in TRACE simulations using a small restricted lexicon, subsequent simulations using a large, more realistic lexicon showed that TRACE feedback can speed word recognition and make it more robust when the signal is degraded (Magnuson, Strauss, & Harris, 2005).

Feedback loops are also widely used in electrical engineering to ensure that information is transmitted accurately – a process known as parity checking. We suggested that feedback-driven parity checking may prevent representational drift as the same speech representation drives parallel distributed analyses at higher (e.g. lexical, semantic, syntactic) levels of representation. **Samuel** reframes this analysis in terms of resonance. Resonance retains the idea that feedback plays a role in coordinating and mutually reinforcing multiple levels of analysis to stabilize representations, but also introduces a temporal component. In resonance systems, representations develop in parallel rather than in sequence. Samuel points to behavioural evidence suggesting that this process may last roughly one second in speech perception (Connine, Blasko, & Hall, 1991; Swinney, 1979). Under this interpretation, the top-down influences that we see developing around 300 ms after the onset of ambiguous phonemes is both consistent with estimates of when listeners show electrophysiological sensitivity to lexical information (Hagoort, 2008), and well within the bounds of this resonance window.

We hope that data-driven effective connectivity analyses can play a significant role in resolving long-standing questions about the functional architecture of a variety of cognitive and linguistic processes. While our primary goal has been to strengthen inferences about how different types of representation interact to support performance, we realize that the approach draws attention to the limits of the current understanding of functional localization, the challenges of characterizing functional interactions, and the sheer complexity of processing underlying even the simplest of tasks. While daunting, this study demonstrates that these challenges can be navigated by examining key brain interactions in the context of broader network function.

Acknowledgements

We thank David Caplan and James Magnuson for their thoughtful comments on this manuscript. This work was originally supported by the National Institute of Deafness and Communicative Disorders (R01 DC003108) and benefited from support from the MIND Institute and the NCRR Regional Resource Grant (41RR14075) for the development of technology and analysis tools at the Athinoula A. Martinos Center for Biomedical Imaging, as well as internal support from the Massachusetts General Hospital Department of Neurology.

References

- Anderson ML. Precis of After phrenology: Neural reuse and the interactive brain. *The Behavioral and Brain Sciences*. 2015; :1–22. DOI: 10.1017/S0140525X15000631
- Aström, KJ.; Murray, RM. *Feedback systems: An introduction for scientists and engineers*. Princeton University Press; Princeton, NJ: 2008.
- Baddeley, A. *Working memory*. Oxford University Press; Oxford: 1986.

- Baddeley A. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*. 2000a; 4(11):417–423. doi: [http://dx.doi.org/10.1016/S1364-6613\(00\)01538-2](http://dx.doi.org/10.1016/S1364-6613(00)01538-2). [PubMed: 11058819]
- Baddeley AD. The phonological loop and the irrelevant speech effect: some comments on Neath (2000). *Psychonomic Bulletin & Review*. 2000b; 7(3):544–549. DOI: 10.3758/BF03214369 [PubMed: 11082863]
- Barnett L. Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables. *Physical Review Letters*. 2009; 103(23)doi: 10.1103/PhysRevLett.103.238701
- Barrouillet P, Camos V. As time goes by: Temporal constraints in working memory. *Current Directions in Psychological Science*. 2012; 21(6):413–419. DOI: 10.1177/0963721412459513
- Blumstein SE, Myers EB, Rissman J. The perception of voice onset time: an fMRI investigation of phonetic category structure. *Journal of Cognitive Neuroscience*. 2005; 17(9):1353–1366. DOI: 10.1162/0898929054985473 [PubMed: 16197689]
- Caplan D, Waters G. Memory mechanisms supporting syntactic comprehension. *Psychonomic Bulletin & Review*. 2013; 20(2):243–268. DOI: 10.3758/s13423-012-0369-9 [PubMed: 23319178]
- Connine CM, Blasko DG, Hall M. Effects of subsequent context in auditory word recognition: Temporal and linguistic constraints. *Journal of Memory and Language*. 1991; 30:234–250. DOI: 10.1016/0749-596X(91)90005-5
- Cornelissen K, Laine M, Renvall K, Saarinen T, Martin N, Salmelin R. Learning new names for new objects: cortical effects as measured by magnetoencephalography. *Brain and Language*. 2004; 89(3):617–622. DOI: 10.1016/j.bandl.2003.12.007 [PubMed: 15120553]
- Dennett, DC. 'Filling-in' versus finding out. In: Pick, HL.; Van den Broek, P.; Knill, DC., editors. *Cognition, conception, and methodological issues*. American Psychological Association; Washington, D.C: 1992. p. 33–49.
- Elman JL, McClelland JL. Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*. 1988; 27(1):143–165. DOI: 10.1016/0749-596X(88)90071-X
- Farah MJ, Smith AF. Perceptual interference and facilitation with auditory imagery. *Perception & Psychophysics*. 1983; 33(5):475–478. DOI: 10.3758/BF03202899 [PubMed: 6877994]
- Frauenfelder, UH.; Peeters, G. Simulating the timecourse of spoken word recognition: An analysis of lexical competition in TRACE. In: Grainger, J.; Jacobs, AM., editors. *Localist connectionist approaches to human cognition*. Lawrence Erlbaum; Mahwah, NJ: 1998. p. 101–146.
- Friston KJ, Price CJ, Fletcher P, Moore C, Frackowiak RS, Dolan RJ. The trouble with cognitive subtraction. *NeuroImage*. 1996; 4(2):97–104. DOI: 10.1006/nimg.1996.0033 [PubMed: 9345501]
- Gage N, Poeppel D, Roberts TP, Hickok G. Auditory evoked M100 reflects onset acoustics of speech sounds. *Brain Research*. 1998; 814(1-2):236–239. DOI: 10.1016/S0006-8993(98)01058-0 [PubMed: 9838140]
- Gow DW. The cortical organization of lexical knowledge: A dual lexicon model of spoken language processing. *Brain and Language*. 2012; 121(3):273–288. DOI: 10.1016/j.bandl.2012.03.005 [PubMed: 22498237]
- Gow DW Jr. Caplan DN. New levels of language processing complexity and organization revealed by granger causation. *Frontiers in Psychology*. 2012a; 3:506.doi: 10.3389/fpsyg.2012.00506 [PubMed: 23293611]
- Gow DW Jr. Caplan DN. New levels of language processing complexity and organization revealed by granger causation. *Frontiers in Psychology*. 2012b; 3:506.doi: 10.3389/fpsyg.2012.00506 [PubMed: 23293611]
- Gow DW, Nied AC. Rules from words: Phonotactic biases in speech perception. *PloS One*. 2014; 9(1): 1–12. DOI: 10.1371/journal.pone.0086212
- Gow DW, Olson BB. 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*. 2015a; 82:41–55. DOI: 10.1016/j.jml.2015.03.004 [PubMed: 25883413]
- Gow DW, Olson BB. Sentential influences on acoustic-phonetic processing: A Granger causality analysis of multimodal imaging data. *Language, Cognition and Neuroscience*. 2015b; doi: 10.1080/23273798.2015.1029498

- Gow DW, Segawa JA. Articulatory mediation of speech perception: a causal analysis of multi-modal imaging data. *Cognition*. 2009; 110(2):222–236. DOI: 10.1016/j.cognition.2008.11.011 [PubMed: 19110238]
- Gow DW, Segawa JA, Ahlfors SP, Lin F-H. Lexical influences on speech perception: A Granger causality analysis of MEG and EEG source estimates. *NeuroImage*. 2008; 43(3):614–623. DOI: 10.1016/j.neuroimage.2008.07.027 [PubMed: 18703146]
- Gruber O. Effects of domain-specific interference on brain activation associated with verbal working memory task performance. *Cerebral Cortex*. 2001; 11(11):1047–1055. DOI: 10.1093/cercor/11.11.1047 [PubMed: 11590114]
- Gruber O, von Cramon DY. The functional neuroanatomy of human working memory revisited. Evidence from 3-T fMRI studies using classical domain-specific interference tasks. *NeuroImage*. 2003; 19(3):797–809. DOI: 10.1016/S1053-8119(03)00089-2 [PubMed: 12880808]
- Hagoort P. The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 2008; 363(1493):1055–1069. DOI: 10.1098/rstb.2007.2159
- Hasson U, Skipper JJ, Nusbaum HC, Small SL. Abstract coding of audiovisual speech: beyond sensory representation. *Neuron*. 2007; 56(6):1116–1126. DOI: 10.1016/j.neuron.2007.09.037 [PubMed: 18093531]
- Hickok G, Poeppel D. The cortical organization of speech processing. *Nature Reviews Neuroscience*. 2007; 8:393–402. DOI: 10.1038/nrn2113 [PubMed: 17431404]
- Hopfinger JB, Buonocore MH, Mangun GR. The neural mechanisms of top-down attentional control. *Nature Neuroscience*. 2000; 3(3):284–291. DOI: 10.1038/72999 [PubMed: 10700262]
- Hubbard TL. Auditory imagery: empirical findings. *Psychological Bulletin*. 2010; 136(2):302–329. DOI: 10.1037/a0018436 [PubMed: 20192565]
- Jacquemot C, Scott SK. What is the relationship between phonological short-term memory and speech processing? *Trends in Cognitive Sciences*. 2006; 10(11):480–486. DOI: 10.1016/j.tics.2006.09.002 [PubMed: 16997610]
- Larsen JD, Baddeley A, Andrade J. Phonological similarity and the irrelevant speech effect: implications for models of short-term verbal memory. *Memory*. 2000; 8(3):145–157. DOI: 10.1080/096582100387579 [PubMed: 10889899]
- Lee H, Devlin JT, Shakeshaft C, Stewart LH, Brennan A, Glensman J, Price CJ. Anatomical traces of vocabulary acquisition in the adolescent brain. *The Journal of Neuroscience : the official journal of the Society for Neuroscience*. 2007; 27(5):1184–1189. DOI: 10.1523/JNEUROSCI.4442-06.2007 [PubMed: 17267574]
- Magnuson J, McMurray B, Tanenhaus MK, Aslin RS. Lexical effects on compensation for coarticulation: the ghost of Christmas past. *Cognitive Science*. 2003a; 27(2):285–298. DOI: 10.1016/s0364-0213(03)00004-1
- Magnuson JS, McMurray B, Tanenhaus MK, Aslin RS. Lexical effects on compensation for coarticulation: a tale of two systems? *Cognitive Science*. 2003b; 27(5):801–805. DOI: 10.1016/s0364-0213(03)00067-3
- Magnuson, JS.; Strauss, T.; Harris, HH. Interaction in spoken word recognition models: Feedback helps; Paper presented at the Annual Meeting of the Cognitive Science Society; Stresa, Italy. 2005.
- McClelland JL, Elman JL. The TRACE model of speech perception. *Cognitive Psychology*. 1986; 18(1):1–86. 00100285(86)900150. [PubMed: 3753912]
- McQueen J, Eisner F, Norris D. When brain regions talk to each other during speech processing, what are they talking about? Commentary on Gow and Olson. *Language, Cognition and Neuroscience*. 2015; (in press). doi: 10.1080/23273798.2015.1029498
- McQueen JM, Jesse A, Norris D. No lexical, Æiprelexical feedback during speech perception or: Is it time to stop playing those Christmas tapes? *Journal of Memory and Language*. 2009; 61(1):1–18. DOI: 10.1016/j.jml.2009.03.002
- Mesgarani N, Cheung C, Johnson K, Chang EF. Phonetic feature encoding in human superior temporal gyrus. *Science*. 2014; 343(6174):1006–1010. DOI: 10.1126/science.1245994 [PubMed: 24482117]

- Myers EB, Blumstein SE, Walsh E, Eliassen J. Inferior frontal regions underlie the perception of phonetic category invariance. *Psychological Science*. 2009; 20(7):895–903. DOI: 10.1111/j.1467-9280.2009.02380.x [PubMed: 19515116]
- Naatanen R, Paavilainen P, Rinne T, Alho K. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology : official journal of the International Federation of Clinical Neurophysiology*. 2007; 118(12):2544–2590. DOI: 10.1016/j.clinph.2007.04.026 [PubMed: 17931964]
- Norris D, McQueen JM. Shortlist B: a Bayesian model of continuous speech recognition. *Psychological Review*. 2008; 115(2):357–395. DOI: 10.1037/0033-295X.115.2.357 [PubMed: 18426294]
- Norris D, McQueen JM, Cutler A. Merging information in speech recognition: feedback is never necessary. *The Behavioral and Brain Sciences*. 2000; 23(3):299–325. DOI: 10.1017/S0140525X00003241 [PubMed: 11301575]
- Norris D, McQueen JM, Cutler A. Perceptual learning in speech. *Cognitive Psychology*. 2003; 47(2):204–238. DOI: 10.1016/S0010-0285(03)00006-9 [PubMed: 12948518]
- Obleser J, Eisner F. Pre-lexical abstraction of speech in the auditory cortex. *Trends in cognitive sciences*. 2009; 13(1):14–19. DOI: 10.1016/j.tics.2008.09.005 [PubMed: 19070534]
- Okada H, Matsuoka K. Effects of auditory imagery on the detection of a pure tone in white noise: experimental evidence of the auditory Perky effect. *Perceptual and Motor Skills*. 1992; 74(2):443–448. DOI: 10.2466/pms.1992.74.2.443 [PubMed: 1594403]
- Peramunage D, Blumstein SE, Myers EB, Goldrick M, Baese-Berk M. Phonological neighborhood effects in spoken word production: an fMRI study. *Journal of Cognitive Neuroscience*. 2011; 23(3):593–603. DOI: 10.1162/jocn.2010.21489 [PubMed: 20350185]
- Phillips C, Pellathy T, Marantz A, Yellin E, Wexler K, Poeppel D, Roberts T. Auditory cortex accesses phonological categories: an MEG mismatch study. *Journal of Cognitive Neuroscience*. 2000; 12(6):1038–1055. DOI: 10.1162/08989290051137567 [PubMed: 11177423]
- Pitt MA, McQueen J. Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*. 1998; 39(3):347–380. DOI: 10.1006/jmla.1998.2571
- Prabhakaran R, Blumstein SE, Myers EB, Hutchison E, Britton B. An event-related fMRI investigation of phonological, lexical competition. *Neuropsychologia*. 2006; 44(12):2209–2221. DOI: 10.1016/j.neuropsychologia.2006.05.025 [PubMed: 16842827]
- Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience*. 2009; 12(6):718–724. DOI: 10.1038/nn.2331 [PubMed: 19471271]
- Richardson FM, Thomas MS, Filippi R, Harth H, Price CJ. Contrasting effects of vocabulary knowledge on temporal and parietal brain structure across lifespan. *Journal of Cognitive Neuroscience*. 2010; 22(5):943–954. DOI: 10.1162/jocn.2009.21238 [PubMed: 19366285]
- Righi G, Blumstein SE, Mertus J, Worden MS. Neural systems underlying lexical competition: An eye tracking and fMRI study. *Journal of Cognitive Neuroscience*. 2009; 22(2):213–224. DOI: 10.1162/jocn.2009.21200
- Romani C, Galluzzi C, Olson A. Phonological-lexical activation: A lexical component or an output buffer? Evidence from aphasic errors. *Cortex; a journal devoted to the study of the nervous system and behavior*. 2011; 47(2):217–235. DOI: 10.1016/j.cortex.2009.11.004 [PubMed: 20163794]
- Samuel AG. Knowing a word affects the fundamental perception of the sounds within it. *Psychological Science*. 2001; 12(4):348–351. DOI: 10.1111/1467-9280.00364 [PubMed: 11476105]
- Samuel AG. Commentary on "sentential influences on acoustic-phonetic processing: A Granger causality analysis of multimodal imaging data". *Language, Cognition and Neuroscience*. (in press).
- Samuel AG, Pitt MA. Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language*. 2003; 48(2):416–434. DOI: 10.1016/S0749-596X(02)00514-4
- Shallice T, Rumati RI, Zadini A. The selective impairment of the phonological output buffer. *Cognitive neuropsychology*. 2000; 17(6):517–546. DOI: 10.1080/02643290050110638 [PubMed: 20945193]

- Spivey MJ. Semantics influences speech perception: Commentary on Gow and Olson. *Language, Cognition and Neuroscience*. 2016; (in press). doi: 10.1080/23273798.2016.1140788
- Steinschneider M, Nourski KV, Kawasaki H, Oya H, Brugge JF, Howard MA 3rd. Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. *Cerebral Cortex*. 2011; 21(10):2332–2347. DOI: 10.1093/cercor/bhr014 [PubMed: 21368087]
- Swinney D. Lexical access during sentence comprehension: (Re-consideration of sentence context effects. *Journal of Verbal Learning and Verbal Behavior*. 1979; 18:645–659. DOI: 10.1016/S0022-5371(79)90355-4
- Trost S, Gruber O. Evidence for a double dissociation of articulatory rehearsal and non-articulatory maintenance of phonological information in human verbal working memory. *Neuropsychobiology*. 2012; 65(3):133–140. DOI: 10.1159/000332335 [PubMed: 22378145]
- Turkeltaub PE, Coslett HB. Localization of sublexical speech perception components. *Brain and Language*. 2010; 114(1):1–15. DOI: 10.1016/j.bandl.2010.03.008 [PubMed: 20413149]
- Vergauwe E, Cowan N. Attending to items in working memory: evidence that refreshing and memory search are closely related. *Psychonomic Bulletin and Review*. 2015; 22(4):1001–1006. DOI: 10.3758/s13423-014-0755-6 [PubMed: 25361821]
- Yoo SS, Lee CU, Choi BG. Human brain mapping of auditory imagery: event-related functional MRI study. *Neuroreport*. 2001; 12(14):3045–3049. DOI: 10.1097/00001756-200110080-00013 [PubMed: 11568634]
- Zatorre RJ, Halpern AR. Mental concerts: musical imagery and auditory cortex. *Neuron*. 2005; 47(1): 9–12. DOI: 10.1016/j.neuron.2005.06.013 [PubMed: 15996544]