



# The hemispheric lateralization of speech processing depends on what “speech” is: a hierarchical perspective

Jonathan E. Peelle\*

Department of Otolaryngology, Washington University in St. Louis, St. Louis, MO, USA  
\*Correspondence: peellej@ent.wustl.edu

**Edited by:**

Russell A. Poldrack, University of Texas, USA

**Reviewed by:**

Russell A. Poldrack, University of Texas, USA

A recurring question in neuroimaging studies of spoken language is whether speech is processed largely bilaterally, or whether the left hemisphere plays a more dominant role (cf., Hickok and Poeppel, 2007; Rauschecker and Scott, 2009). Although questions regarding underlying mechanisms are certainly of interest, the discussion unfortunately gets sidetracked due to the imprecise use of the word “speech”: by being more explicit about the type of cognitive and linguistic processing to which we are referring it may be possible to reconcile many of the disagreements present in the literature.

## LEVELS OF PROCESSING DURING CONNECTED SPEECH COMPREHENSION

A relatively uncontroversial starting point is to acknowledge that understanding a spoken sentence requires a listener to analyze a complex acoustic signal along a number of levels, listed schematically in **Figure 1**. Phonemes must be distinguished, words identified, and grammatical structure taken into account so that meaning can be extracted. These processes operate in an interactive parallel fashion, and as such are difficult to fully disentangle. Such interdependence also means that as researchers we often use “speech” as a term of convenience to mean:

1. Amplitude-modulated noise or spectral transitions, as might be similar to aspects of spoken language;
2. Phonemes (“b”), syllables (“ba”), or pseudowords (“bab”);
3. Words (“bag”);
4. Phrases (“the bag”);
5. Sentences (“The bag of carrots fell to the floor”) or narratives.

<sup>1</sup>Including me.

<sup>2</sup>In fact, the term “lateralization” is also used to variously mean (a) one hemisphere performing a task and the other not being involved, or (b) both hemispheres being engaged in a task, but one hemisphere is doing more of the work or being slightly more efficient, potentially compounding the confusion.

Naturally, because different types of spoken language require different cognitive mechanisms—spanning sublexical, lexical, and suprallexical units—using an unqualified term such as “speech” can lead to confusion about the processes being discussed. Although this point might seem obvious, a quick review of the speech literature demonstrates that many authors<sup>1</sup> have at one time or another assumed their definition of “speech” was obvious enough that they need not give it, leaving readers to form their own opinions.

Below I will briefly review literature in relation to the neural bases for two types of spoken language processing: unconnected speech (isolated phonemes and single words) and connected speech (sentences or narratives). The goal is to illustrate that, within the context of a hierarchical neuroanatomical framework, there are aspects of “speech” processing that are both bilateral and lateralized.

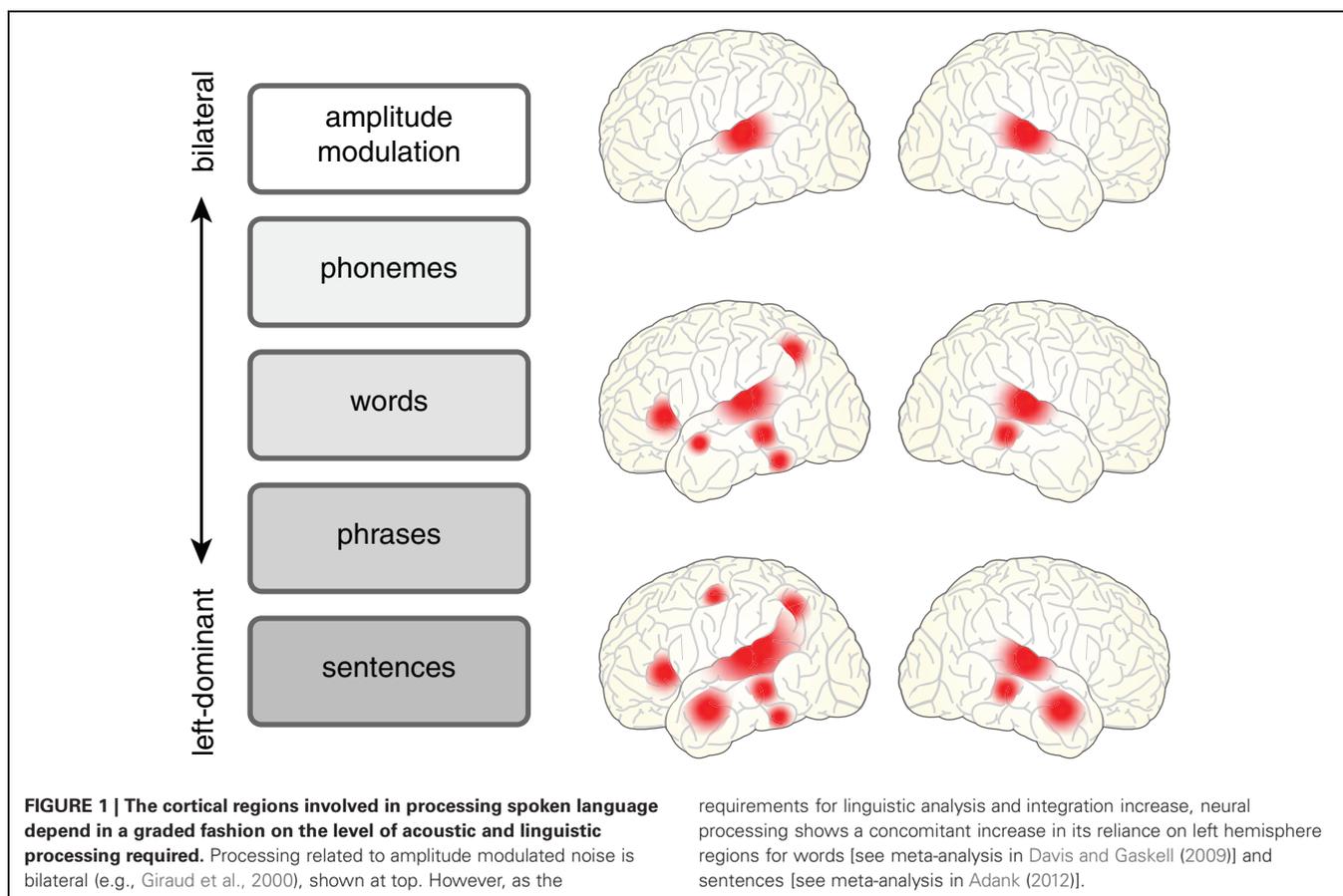
## UNCONNECTED SPEECH IS PROCESSED LARGELY BILATERALLY IN TEMPORAL CORTEX

The first cortical way station for acoustic input to the brain is primary auditory cortex: not surprisingly, acoustic stimuli activate this region robustly in both hemispheres, whether they consist of pure tones (Belin et al., 1999; Binder et al., 2000) or amplitude-modulated noise (Giraud et al., 2000; Hart et al., 2003; Overath et al., 2012). Although there is speculation regarding hemispheric differences in specialization for these low level signals (Poeppel, 2003; Giraud et al., 2007; Obleser et al., 2008; McGettigan and Scott, 2012), for the current discussion, it is

sufficient to note that both left and right auditory cortices respond robustly to most auditory stimuli, and that proposed differences in hemispheric preference relate to a modulation of this overall effect<sup>2</sup>.

Beyond low-level acoustic stimulation, phonemic processing requires both an appropriate amount of spectral detail and the relationship to a pre-existing acoustic category (i.e., the phoneme). The processing of isolated syllables results in activity along the superior temporal sulcus and middle temporal gyrus, typically on the left but not the right (Liebenthal et al., 2005; Heinrich et al., 2008; Agnew et al., 2011; DeWitt and Rauschecker, 2012). Although this may suggest a left hemisphere specialization for phonemes, listening to words (which, of course, include phonemes) reliably shows strong activity in bilateral middle and superior temporal gyrus (Price et al., 1992; Binder et al., 2000, 2008). In addition, stroke patients with damage to left temporal cortex are generally able to perform reasonably well on word-to-picture matching tasks (Gainotti et al., 1982); the same is true of healthy controls undergoing a Wada procedure (Hickok et al., 2008). Together these findings suggest that the right hemisphere is able to support at least some degree of phonemic and lexical processing.

That being said, there are also regions that show increased activity for words in the left hemisphere but not the right, particularly when pseudowords are used as a baseline (Davis and Gaskell, 2009). Both pseudowords and real words rely on stored representations of speech sounds (they share phonemes), but real words



also involve consolidated lexical and/or conceptual information (Gagnepain et al., 2012). Left-hemisphere activations likely reflect the contribution of lexical and semantic memory processes that are accessed in an obligatory manner during spoken word recognition. Within the framework outlined in **Figure 1**, spoken words thus lie between very low level auditory processing (which is essentially bilateral) and the processing of sentences and narratives (which, as I will discuss below, is more strongly left lateralized).

Processing of phonemes and single words therefore appears to be mediated in large part by both left and right temporal cortex, although some indications of lateralization may be apparent.

### CONNECTED SPEECH RELIES ON A LEFT-LATERALIZED FRONTOTEMPORAL NETWORK

In addition to recognizing single words, comprehending connected speech—such as meaningful sentences—depends on

integrative processes that help determine the syntactic and semantic relationship between words. These processes rely not only on phonemic and lexical information, but also on prosodic and rhythmic cues conveyed over the course of several seconds. In other words, a sentence is not simply a string of phoneme-containing items, but conveys a larger meaning through its organization (Vandenberghe et al., 2002; Humphries et al., 2006; Lerner et al., 2011; Peelle and Davis, 2012). In addition to providing content in and of itself, the syntactic, semantic, and rhythmic structure present in connected speech also supports listeners' predictions of upcoming acoustic information.

An early and influential PET study of connected speech by Scott et al. showed increased activity in the lateral aspect of left anterior temporal cortex for spoken sentences relative to unintelligible spectrally-rotated versions of these sentences (Scott et al., 2000). Subsequent studies, due in part to the use of

a greater number of participants, have typically found intelligibility effects bilaterally, often along much of the length of superior temporal cortex (Crinion et al., 2003; Friederici et al., 2010; Wild et al., 2012a). In addition, a large and growing number of neuroimaging experiments show left inferior frontal involvement for intelligible sentences, either compared to an unintelligible control condition (Rodd et al., 2005, 2010; Awad et al., 2007; Obleser et al., 2007; Okada et al., 2010; Peelle et al., 2010a; McGettigan et al., 2012; Wild et al., 2012b) or parametrically correlating with intelligibility level (Davis and Johnsrude, 2003; Obleser and Kotz, 2010; Davis et al., 2011). Regions of left inferior frontal cortex are also involved in processing syntactically complex speech (Peelle et al., 2010b; Tyler et al., 2010; Obleser et al., 2011) and in resolving semantic ambiguity (Rodd et al., 2005, 2010, 2012; Snijders et al., 2010). In most of these studies activity in right inferior frontal cortex is not significant, or is noticeably smaller in extent than

activity in the left hemisphere. These functional imaging studies are consistent with patient work demonstrating that participants with damage to left inferior frontal cortex have difficulty with sentence processing (e.g., Grossman et al., 2005; Peelle et al., 2007; Papoutsis et al., 2011; Tyler et al., 2011).

Processing connected speech thus relies more heavily on left hemisphere language regions, most obviously in inferior frontal cortex. The evidence outlined above suggests this is largely due to the increased linguistic demands associated with sentence processing compared to single words.

### THE IMPORTANCE OF STATISTICAL COMPARISONS FOR INFERENCES REGARDING LATERALITY

In many of the above papers (and in my interpretation of them) laterality was not statistically assessed, but inferred based on the presence or absence of an activation cluster in a particular brain region. That is, seeing a cluster of activation in left inferior frontal gyrus but not the right, and concluding that this particular task has a “left lateralized” pattern of neural activity. However, simply observing a response in one region, but not another, does not mean that these regions significantly differ in their activity (the “imager’s fallacy”; Henson, 2005). This is a well-known statistical principle, but one that can remain difficult to follow in the face of compelling graphical depictions of data (Nieuwenhuis et al., 2011).

Nevertheless, for true claims of differential hemispheric contributions to speech processing, the left and right hemisphere responses need to be directly compared. Unfortunately, for functional imaging studies hemispheric comparisons are not as straightforward as they seem, in part because our left and right hemispheres are not mirror images of each other. There are, however, a number of reasonable ways to approach this challenge, including:

1. Extracting data from regions of interest (ROIs), including independently defined functional regions (Kriegeskorte et al., 2009) or probabilistic cytoarchitecture (Eickhoff et al., 2005), and averaging over voxels to compare left and right hemisphere responses. Sometimes these ROIs end

up being large, which does not always support the specific hypotheses being tested, and not all regions may be available. However, this approach is relatively straightforward to implement and interpret.

2. Using a custom symmetric brain template for spatial normalization (Bozic et al., 2010). This may result in less veridical spatial registration, but enables voxel-by-voxel statistical tests of laterality by flipping images around the Y axis, avoiding the problem of ROI selection (and averaging).
3. Comparing left vs. right hemisphere responses using a multivariate classification approach (McGettigan et al., 2012). Multivariate approaches are robust to large ROIs, as their performance is typically driven by a smaller (more informative) subset of all voxels studied. Multivariate approaches may be somewhat more challenging to implement, however, and (depending on the size of the ROI used) may limit spatial specificity.

In the absence of these or similar statistical comparisons, any statements about lateralization of processing need to be made (and taken) lightly.

### CONCLUSIONS

I have not intended to make any novel claims about the neural organization of speech processing, merely to clarify what has already been shown: phonological and lexical information is processed largely bilaterally in temporal cortex, whereas connected speech relies on a left-hemisphere pathway that includes left inferior frontal gyrus. Importantly, the distinction between unconnected and connected speech is not dichotomous, but follows a gradient of laterality depending on the cognitive processes required: lateralization emerges largely as a result of increased linguistic processing.

So, is speech processed primarily bilaterally, or along a left-dominant pathway? It depends on what sort of “speech” we are talking about, and being more specific in our characterizations will do much to advance the discussion. Of more interest will be future studies that continue to identify the constellation of cognitive processes supported by these neuroanatomical networks.

### ACKNOWLEDGMENTS

I am grateful to Matt Davis for helpful comments on this manuscript. This work was supported by NIH grants AG038490 and AG041958.

### REFERENCES

- Adank, P. (2012). Design choices in imaging speech comprehension: an activation likelihood estimation (ALE) meta-analysis. *Neuroimage* 63, 1601–1613.
- Agnew, Z. K., McGettigan, C., and Scott, S. K. (2011). Discriminating between auditory and motor cortical responses to speech and nonspeech mouth sounds. *J. Cogn. Neurosci.* 23, 4038–4047.
- Awad, M., Warren, J. E., Scott, S. K., Turkheimer, F. E., and Wise, R. J. S. (2007). A common system for the comprehension and production of narrative speech. *J. Neurosci.* 27, 11455–11464.
- Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., and Pike, B. (1999). Event-related fMRI of the auditory cortex. *Neuroimage* 10, 417–429.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–528.
- Binder, J. R., Swanson, S. J., Hammeke, T. A., and Sabsevitz, D. S. (2008). A comparison of five fMRI protocols for mapping speech comprehension systems. *Epilepsia* 49, 1980–1997.
- Bozic, M., Tyler, L. K., Ives, D. T., Randall, B., and Marslen-Wilson, W. D. (2010). Bihemispheric foundations for human speech comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17439–17444.
- Crinion, J. T., Lambon Ralph, M. A., Warburton, E. A., Howard, D., and Wise, R. J. S. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain* 126, 1193–1201.
- Davis, M. H., and Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 3773–3800.
- Davis, M. H., Ford, M. A., Kherif, F., and Johnsrude, I. S. (2011). Does semantic context benefit speech understanding through “top-down” processes? Evidence from time-resolved sparse fMRI. *J. Cogn. Neurosci.* 23, 3914–3932.
- Davis, M. H., and Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *J. Neurosci.* 23, 3423–3431.
- DeWitt, I., and Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proc. Natl. Acad. Sci. U.S.A.* 109, E505–E514.
- Eickhoff, S., Stephan, K., Mohlberg, H., Grefkes, C., Fink, G., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335.
- Friederici, A. D., Kotz, S. A., Scott, S. K., and Obleser, J. (2010). Disentangling syntax and intelligibility in auditory language comprehension. *Hum. Brain Mapp.* 31, 448–457.
- Gagnepain, P., Henson, R. N., and Davis, M. H. (2012). Temporal predictive codes for spoken words in auditory cortex. *Curr. Biol.* 22, 615–621.

- Gainotti, G., Miceli, G., Silveri, M. C., and Villa, G. (1982). Some anatomo-clinical aspects of phonemic and semantic comprehension disorders in aphasia. *Acta Neurol. Scand.* 66, 652–665.
- Giraud, A.-L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S. J., and Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* 56, 1127–1134.
- Giraud, A.-L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., et al. (2000). Representation of the temporal envelope of sounds in the human brain. *J. Neurophysiol.* 84, 1588–1598.
- Grossman, M., Rhee, J., and Moore, P. (2005). Sentence processing in frontotemporal dementia. *Cortex* 41, 764–777.
- Hart, H. C., Palmer, A. R., and Hall, D. A. (2003). Amplitude and frequency-modulated stimuli activate common regions of human auditory cortex. *Cereb. Cortex* 13, 773–781.
- Heinrich, A., Carlyon, R. P., Davis, M. H., and Johnsrude, I. S. (2008). Illusory vowels resulting from perceptual continuity: a functional magnetic resonance imaging study. *J. Cogn. Neurosci.* 20, 1737–1752.
- Henson, R. (2005). What can functional neuroimaging tell the experimental psychologist? *Q. J. Exp. Psychol.* 58A, 193–233.
- Hickok, G., Okada, K., Barr, W., Pa, J., Rogalsky, C., Donnelly, K., et al. (2008). Bilateral capacity for speech sound processing in auditory comprehension: evidence from Wada procedures. *Brain Lang.* 107, 179–184.
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.
- Humphries, C., Binder, J. R., Medler, D. A., and Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *J. Cogn. Neurosci.* 18, 665–679.
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., and Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540.
- Lerner, Y., Honey, C. J., Silbert, L. J., and Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* 31, 2906–2915.
- Liebenthal, E., Binder, J. R., Spitzer, S. M., Posing, E. T., and Medler, D. A. (2005). Neural substrates of phonemic perception. *Cereb. Cortex* 15, 1621–1631.
- McGettigan, C., Evans, S., Agnew, Z., Shah, P., and Scott, S. K. (2012). An application of univariate and multivariate approaches in fMRI to quantifying the hemispheric lateralization of acoustic and linguistic processes. *J. Cogn. Neurosci.* 24, 636–652.
- McGettigan, C., and Scott, S. K. (2012). Cortical asymmetries in speech perception: what's wrong, what's right and what's left? *Trends Cogn. Sci.* 16, 269–276.
- Nieuwenhuis, S., Forstmann, B. U., and Wagenmakers, E.-J. (2011). Erroneous analysis of interactions in neuroscience: a problem of significance. *Nat. Neurosci.* 14, 1105–1107.
- Obleser, J., Eisner, F., and Kotz, S. A. (2008). Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *J. Neurosci.* 28, 8116–8124.
- Obleser, J., and Kotz, S. A. (2010). Expectancy constraints in degraded speech modulate the language comprehension network. *Cereb. Cortex* 20, 633–640.
- Obleser, J., Meyer, L., and Friederici, A. D. (2011). Dynamic assignment of neural resources in auditory comprehension of complex sentences. *Neuroimage* 56, 2310–2320.
- Obleser, J., Wise, R. J. S., Dresner, M. A., and Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *J. Neurosci.* 27, 2283–2289.
- Okada, K., Rong, F., Venezia, J., Matchin, W., Hsich, I.-H., Saberi, K., et al. (2010). Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. *Cereb. Cortex* 20, 2486–2495.
- Overath, T., Zhang, Y., Sanes, D. H., and Poeppel, D. (2012). Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: fMRI evidence. *J. Neurophysiol.* 107, 2042–2056.
- Papoutsis, M., Stamatakis, E. A., Griffiths, J., Marslen-Wilson, W. D., and Tyler, L. K. (2011). Is left fronto-temporal connectivity essential for syntax? Effective connectivity, tractography and performance in left-hemisphere damaged patients. *Neuroimage* 58, 656–664.
- Peelle, J. E., Cooke, A., Moore, P., Vesely, L., and Grossman, M. (2007). Syntactic and thematic components of sentence processing in progressive nonfluent aphasia and nonaphasic frontotemporal dementia. *J. Neurolinguistics* 20, 482–494.
- Peelle, J. E., and Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Front. Psychology* 3:320. doi: 10.3389/fpsyg.2012.00320
- Peelle, J. E., Eason, R. J., Schmitter, S., Schwarzbauer, C., and Davis, M. H. (2010a). Evaluating an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory processing. *Neuroimage* 52, 1410–1419.
- Peelle, J. E., Troiani, V., Wingfield, A., and Grossman, M. (2010b). Neural processing during older adults' comprehension of spoken sentences: age differences in resource allocation and connectivity. *Cereb. Cortex* 20, 773–782.
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Commun.* 41, 245–255.
- Price, C. J., Wise, R., Ramsay, S., Friston, K., Howard, D., Patterson, K., et al. (1992). Regional response differences within the human auditory cortex when listening to words. *Neurosci. Lett.* 146, 179–182.
- Rauschecker, J. P., and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724.
- Rodd, J. M., Davis, M. H., and Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cereb. Cortex* 15, 1261–1269.
- Rodd, J. M., Johnsrude, I. S., and Davis, M. H. (2012). Dissociating frontotemporal contributions to semantic ambiguity resolution in spoken sentences. *Cereb. Cortex* 22, 1761–1773.
- Rodd, J. M., Longe, O. A., Randall, B., and Tyler, L. K. (2010). The functional organisation of the fronto-temporal language system: evidence from syntactic and semantic ambiguity. *Neuropsychologia* 48, 1324–1335.
- Scott, S. K., Blank, C. C., Rosen, S., and Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123, 2400–2406.
- Snijders, T. M., Petersson, K. M., and Hagoort, P. (2010). Effective connectivity of cortical and subcortical regions during unification of sentence structure. *Neuroimage* 52, 1633–1644.
- Tyler, L. K., Marslen-Wilson, W. D., Randall, B., Wright, P., Devereux, B. J., Zhuang, J., et al. (2011). Left inferior frontal cortex and syntax: function, structure and behaviour in patients with left hemisphere damage. *Brain* 134, 415–431.
- Tyler, L. K., Shafto, M. A., Randall, B., Wright, P., Marslen-Wilson, W. D., and Stamatakis, E. A. (2010). Preserving syntactic processing across the adult life span: the modulation of the frontotemporal language system in the context of age-related atrophy. *Cereb. Cortex* 20, 352–364.
- Vandenberghe, R., Nobre, A. C., and Price, C. J. (2002). The response of left temporal cortex to sentences. *J. Cogn. Neurosci.* 14, 550–560.
- Wild, C. J., Davis, M. H., and Johnsrude, I. S. (2012a). Human auditory cortex is sensitive to the perceived clarity of speech. *Neuroimage* 60, 1490–1502.
- Wild, C. J., Yusuf, A., Wilson, D., Peelle, J. E., Davis, M. H., and Johnsrude, I. S. (2012b). Effortful listening: the processing of degraded speech depends critically on attention. *J. Neurosci.* 32, 14010–14021.

Received: 09 October 2012; accepted: 25 October 2012; published online: 16 November 2012.

Citation: Peelle JE (2012) The hemispheric lateralization of speech processing depends on what "speech" is: a hierarchical perspective. *Front. Hum. Neurosci.* 6:309. doi: 10.3389/fnhum.2012.00309

Copyright © 2012 Peelle. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.