Neural bases of accented speech perception

Patti Adank[1,2]*, Helen E. Nuttall[1], Briony Banks[2] and Daniel Kennedy-Higgins[1]

[1] Division of Psychology and Language Sciences, Department of Speech, Hearing, and Phonetic Sciences, University College London, London, UK, [2] School of Psychological Sciences, University of Manchester, Manchester, UK

The recognition of unfamiliar regional and foreign accents represents a challenging task for the speech perception system (Floccia et al., 2006; Adank et al., 2009). Despite the frequency with which we encounter such accents, the neural mechanisms supporting successful perception of accented speech are poorly understood. Nonetheless, candidate neural substrates involved in processing speech in challenging listening conditions, including accented speech, are beginning to be identified. This review will outline neural bases associated with perception of accented speech in the light of current models of speech perception, and compare these data to brain areas associated with processing other speech distortions. We will subsequently evaluate competing models of speech processing with regards to neural processing of accented speech. See Cristia et al. (2012) for an in-depth overview of *behavioral* aspects of accent processing.

Keywords: cognitive neuroscience, speech perception, accented speech, fMRI, speech in noise, noise-vocoded speech, time-compressed speech

Processing Accent Variation at Pre- and Post-lexical Levels

Models outlining the neural organization of speech perception (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009) propose that the locus of processing intelligible speech is the temporal lobe within the ventral stream of speech processing. Rauschecker & Scott suggest that intelligibility processing has its center of gravity in left anterior STS (Superior Temporal Sulcus), while Hickok & Poeppel propose that processing intelligible speech is bilaterally organized and located both anteriorly and posteriorly to Heschl's Gyrus. However, both models are based on intelligible speech perception and do not make explicit predictions about the cortical substrates that subserve speech perception under challenging listening conditions (cf. Adank, 2012a) for a discussion on processing of intelligible speech).

A handful of fMRI studies address how the brain processes accent variation. Listening to difficult foreign phonemic contrasts (e.g., /l/-/r/ contrasts for Japanese listeners) has been associated with increased activation in auditory processing/speech production areas, including left Inferior Frontal Gyrus (IFG), left insula, bilateral ventral Premotor Cortex, right Pre- and Post-Central Gyrus, left anterior Superior Temporal Sulcus and Gyrus (STS/STG), left Planum Temporale (PT), left superior temporal parietal area (Stp), left Supramarginal Gyrus (SMG), and cerebellum bilaterally (Callan et al., 2004, 2014). It is noteworthy that the neural bases associated with listening to foreign languages overlap with those reported for unfamiliar accent processing, including bilateral STG/STS/MTG, and left IFG (Perani et al., 1996; Perani and Abutalebi, 2005; Hesling et al., 2012).

For sentence processing (**Table 1**, **Figure 1**), listening to an unfamiliar accent involves a network of frontal (left IFG, both Operculi/Insulas, Superior Frontal Gyrus), temporal (left Middle Temporal Gyrus [MTG], right STG), and medial regions (Supplementary Motor Area [SMA])

TABLE 1 | Reported brain regions in studies investigating processing of accented, time-compressed, or noise-vocoded speech, plus speech with added background noise vs. undistorted words or sentences.

Distortion	Study	Contrast	MNI	Location	Original location*
Unfamiliar accent	Adank et al., 2012b	Sentences unfamiliar > sentences familiar accent	−54, −40, 4	L MTG	L Post. STG/SMG
			−60, −34, 8	L MTG	L Post. STG/PT
			−60, −26, −4	L MTG	L Post. MTG
			60, −32, 2	R STG	R Post. STG/SMG
			−50, 12, 24	L POp	L POp/PG
			−46, 16, 12	L POp	L POp/PTr
			54, −26, −2	R STG	R Post. STG/MTG/SMG
			54, 4, −16	R RO	R Ant. STG/TP/MTG
			38, 18, 26	R PTr	R Central Opercular Cortex
	Adank et al., 2012a	Sentences in unfamiliar > sentences in familiar accent	−60, −12, −6	L MTG	L STG/STS
	Adank et al., 2013	Sentences in unfamiliar accent > unintelligible sentences	−62, −32, 4	L MTG	L STS
			−58, −4, −8	L FO	L STG
			−60, −16, −8	L MTG	L MTG
			−50, 18, 24	L PTr	L IFG PTr)
			−46, 28, −4	L POrb	L IFG POrb)
			−36, 22, −4	L Insula	L Insula
			56, −20, −6	R STG	R STG
			60, 2, −12	R STG	R STG
			−2, 10, 60	L SMA	L SMA
	Yi et al., 2014	Sentences in foreign accent > sentences in native accent	4, 24, 34	R MCC	R Paracingulate Gyrus
			34, −52, 62	R SPL	R Motor cortex, SPL, somatosensory cortex
			−40, 14, 8	L Insula	L Insula
			20, −2, 60	R SFG	R SFG
			32, 20, −6	*No location given*	R Insula
			−52, 10, 10	L POp	L IFG
			−26, 24, 0	L Insula	L Insula
			42, 14, 8	R IFG	R Insula
Time−compressed speech	Adank and Devlin, 2010	Time−compressed > normal−speed sentences	−60, −14, 0	L MTG	L Ant. STG/STS
			−58, −46, 4	L MTG	L Post. STG/STS
			64, −14, 0	R STG	R Ant. STG/STS
			56, −32, 4	R STG	R Post. STG/STS
			0, 12, 60	SMA	Pre−SMA
			0, 22, 44	SMA	Cingulate sulcus
			−36, 24, −4	L Insula	L FO
			36, 25, 2	R Insula	R FO
	Peelle et al., 2004	Time−compressed > normal−speed sentences	−28.38, −66.82, 47.33	L SPL	L Posterior parietal BA19/39/40)
			−28.54, −76.78, 32.63	L MOG	L Inferior parietal BA19/39)
			−54.12, −38.58, −16.66	L STG	L Inferior temporal BA20)
			−15.43, −62.52, 46.69	L SPL	L Posterior parietal BA7)
			14.07, −23.17, −4.77	R Thalamus	R Thalamus
			13.99, −7.32, −7.46	R Thalamus	R Subthalamic nucleus
			1.02, −38.08, −14.28	R Cerebellar Vermis	R Cerebellum
	Poldrack et al., 2001	Compression−related increases during sentence processing	−28, 54, 16	L MFG	L MFG
			34, 26, −4	R Insula	R IFG/Insula
			4, 32, 20	R ACC	R ACC
			18, 4, 8	*No location given*	Striatum
			66, −40, 8	R MTG	R STG

(Continued)

TABLE 1 | Continued

Distortion	Study	Contrast	MNI	Location	Original location*
Noise−vocoded speech	Erb et al., 2013	Noise−vocoded > clear sentences	−6, 26, 40	L SMedG	L SMA/ACC
			−30, 20, −5	L Insula	L Ant. Insula
			33, 23, −3	R Insula	R Ant. Insula
			−9, 11, 7	L Caudate Nucleus	L Caudate Nucleus
			12, 17, 10	R Caudate Nucleus	R Caudate Nucleus
	Zekveld et al., 2014	Noise−vocoded > clear sentences	−4, 8, 60	L SMeDG	L SFG
			−64, −40, 10	L STG	L STG
			−48, −42, 2	L MTG	L MTG
			−44, −38, 8	L STG	L MTG
Background noise	Adank et al., 2012a	Sentences in background noise > sentences in quiet	32, 28, 10	*No location given*	R IFG/FO
			−32, 24, 8	L Insula	L FO/IFG/Insula
			6, 14, 28	*No location given*	R Cingulate Gyrus
			−24, 40, −2	*No location given*	L Parahippocampal Gyrus
			−12, 10, −2	L Putamen	L Caudate
			12, 20, 36	R MCC	R Paracingulate/Cingulate
			30, 40, 24	R MFG	R Frontal Pole
			8, 22, 18	*No location given*	R Cingulate Gyrus
	Peelle et al., 2010	Sentences in continuous scanning EPI Sequence > sentences in quiet EPI sequence	−36, −74, 44	L IPL	L Inferior parietal cortex
			−40, −66, 44	L AG	L Angular gyrus
			−48, −60, 48	L IPL	L Inferior parietal cortex
			−56, −46, 8	L MTG	L Post. MTG
			−66, −44, 0	L MTG	L Post MTG
			−68, −14, 2	L STG	L Ant. STS
			−68, 2, −8	*No location given*	L Ant. STS
			−60, 4, −14	L STG	L Ant. STS

*Note that the list of papers is not exhaustive. Coordinates in Talairach space were converted to MNI space using the tal2icbm_spm algorithm www.brainmap.org/ale. Anatomical locations determined using the Anatomy ToolBox (Eickhoff et al., 2005, 2006, 2007) in SPM8 Wellcome Imaging Department, University College London, London, UK). *Original location as reported in the study. AG, Angular Gyrus; FFG, Fusiform Gyrus; FO, Frontal Operculum; IFG, Inferior Frontal Gyrus; IOG, Inferior Occipital Gyrus; IPL, Inferior Parietal Lobule; MCC, Middle Cingulate Cortex; MFG, Middle Frontal Gyrus; MTG, Middle Temporal Gyrus; PG, Precentral Gyrus; POp, Pars Opercularis; PT, Planum Temporale; PTr, Pars Triangularis; POrb, Par Orbitalis: RO, Rolandic Operculum; SMA, Supplementary Motor Area; SMedG, Superior Medial Gyrus; SMG, Supramarginal Gyrus; STG, Superior Temporal Gyrus; STG, Superior Temporal Planum; STS, Superior Temporal Sulcus; TP, Temporal Pole.*
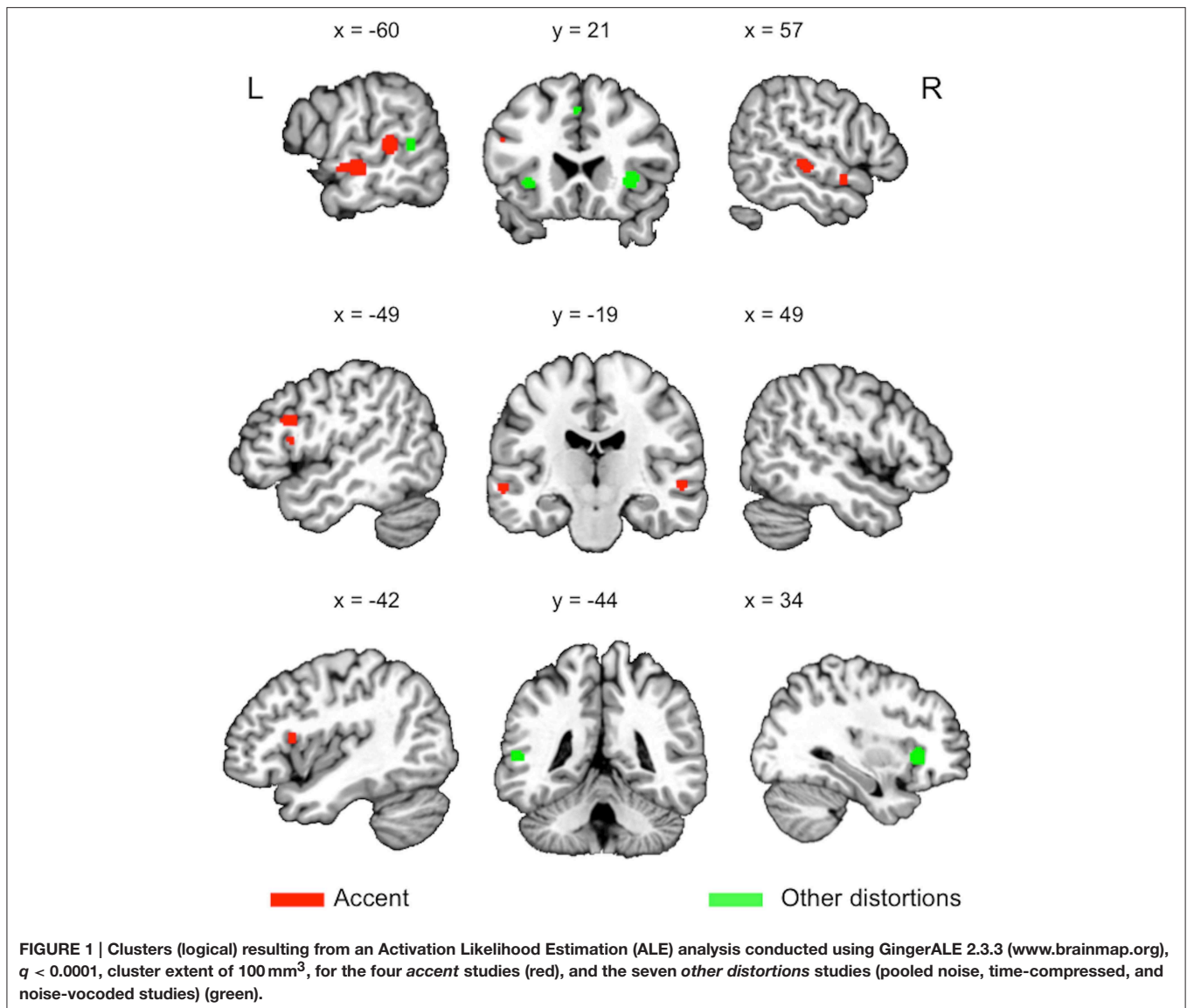
(Adank, 2012b; Adank et al., 2012b, 2013; Yi et al., 2014). It is unclear how the accent processing network maps onto the networks in Rauschecker and Scott (2009) and Hickok and Poeppel (2007). The coordinates for accent processing in the left temporal lobe are located anteriorly and posteriorly to Hickok and Poeppel's proposed STG area for spectrotemporal analysis, while the coordinates in left IFG are located inside Hickok and Poeppel's left inferior frontal area assigned to the dorsal stream's articulatory network. In contrast, the temporal coordinates in **Table 1** fit well with Rauschecker & Scott's antero-ventral and postero-dorsal areas placed anteriorly and posteriorly to left primary auditory cortex, respectively, and the left IFG coordinates fall within their antero-ventral left inferior frontal area.

Accented Speech vs. Other Challenging Listening Conditions

As is the case with other types of distorted speech, understanding accented speech is associated with increased listening effort (Van Engen and Peelle, 2014). However, accent variation is of a conceptually different nature than variation in the acoustic signal resulting from an extrinsic source such as noise, i.e., phonetic realizations that differ from the listener's native realization of speech sounds. Furthermore, in contrast to speech-intrinsic variation, noise compromises the auditory system's representation of speech from ear to brain. Accented speech also differs from distortions such as noise-vocoded or time-compressed speech as the variation does not affect the acoustic integrity of the acoustic signal, as only specific phonemic and suprasegmental characteristics vary.

Processing speech in noise involves areas also activated for speech in an unfamiliar accent (**Table 1**): left insula (Adank et al., 2012a), left MTG (Peelle et al., 2010), left Pars Opercularis (POp), bilateral Pars Triangularis (PTr). Comprehension of time-compressed sentences activates left MTG (Poldrack et al., 2001; Adank and Devlin, 2010), right STG (Peelle et al., 2004; Adank and Devlin, 2010), SMA and left Insula (Adank and Devlin, 2010), while noise-vocoded speech activates left Insula (Erb et al., 2013), and left MTG/STG (Zekveld et al., 2014). However,

FIGURE 1 | Clusters (logical) resulting from an Activation Likelihood Estimation (ALE) analysis conducted using GingerALE 2.3.3 (www.brainmap.org), $q < 0.0001$, **cluster extent of 100 mm^3, for the four** *accent* **studies (red), and the seven** *other distortions* **studies (pooled noise, time-compressed, and noise-vocoded studies) (green).**

it is clear from **Figure 1** that processing accented speech also activates areas outside the network activated for processing speech in noise, time-compressed speech, and noise-vocoded speech.

Another problem in identifying networks governing accent processing is that perceiving variation in an unfamiliar accent (i.e., in an accent that differs from one's own accent and that the listener has had little or no exposure to) is confounded with cognitive load. Note that such confounds also exist for other distortions of the speech signal, such as background noise. Listeners process speech in an unfamiliar accent slower and less efficiently (Floccia et al., 2006). It is thus unclear to which extent the network supporting accented speech perception is shared with the network associated with increased task/cognitive load processing. Notably, an increase in task difficulty/working memory load relates to increases in BOLD-activation in left insula (Wild et al., 2012), and in left MTG, SMA, left PTr,

and right STG (Wild et al., 2012), and could therefore explain activations in these regions related to processing accented speech. Directly comparing the neural processing of familiar/unfamiliar accents may help distinguishing between the two networks.

Accounts of Accented and Distorted Speech Processing

The current debate regarding how listeners understand others in challenging listening conditions focuses on the location and nature of neural substrates recruited for effective speech comprehension. The three accounts discussed below offer specific predictions regarding the neural networks involved in processing accented speech.

First, *auditory-only accounts* (Obleser and Eisner, 2009) hold that speech perception includes a prelexical abstraction process

in which variation in the acoustic signal is "stripped away" to allow the perception system access to abstract linguistic representations. The abstraction process is placed at locations predominantly in the temporal (STS and STG) lobes. This account predicts that processing of accented speech takes place predominantly in the ventral stream, with minimal involvement of the dorsal stream.

Second, *motor recruitment accounts* suggest that auditory areas in the ventral stream and speech production areas in the dorsal stream are required to process unfamiliar speech signals (Wilson and Knoblich, 2005; Pickering and Garrod, 2013). These accounts assume that listening to speech results in the automatic activation of articulatory motor plans required for producing speech (Watkins et al., 2003). These motor plans provide forward models with information of articulatory mechanics, to be used when the incoming signal is ambiguous/unclear. Accented speech contains variation that can lead to ambiguities, and these accounts thus predict that perception of accented speech involves active involvement of speech production processes.

Third, *executive recruitment accounts* propose that activation of (pre-) motor areas during perception of distorted speech signals is not related to actual articulatory processing, but reflects the recruitment of general cognitive processes, such as increased attention, or decision processes (Rodríguez-Fornells et al., 2009; Venezia et al., 2012). Indeed, behavioral data suggest that recruitment of executive functions for processing accented speech (Adank and Janse, 2010; Janse and Adank, 2012; Banks et al., 2015) also predicts activation of frontal regions including left frontal operculum and anterior insula and precentral gyrus, as these regions have also been associated with executive functions such as working memory (Moisala et al., 2015).

The results in **Table 1** contrast with predictions made by the auditory-only account (Obleser and Eisner, 2009), as areas associated with processing accent variation in **Table 1** refer to a more widespread network than predicted. Instead, the network in **Table 1** converges with the latter two accounts, as activation is located across ventral and prefrontal areas in the dorsal stream. We propose that these three accounts are synthesized into a single mixed account for processing of accented speech that brings together neural substrates associated with increased involvement of auditory and phonological processing (e.g., bilateral posterior STG), (pre-)motor recruitment for sensorimotor mapping (e.g., SMA), and substrates associated with increased reliance on cognitive control processes (e.g., IFG, insula, and frontal operculum).

Concluding Remarks

The neural mechanisms responsible for processing accent variation in speech are not clearly outlined, but constitute a topic of active investigation in the field of speech perception. However, to progress our understanding in this area, future studies should meet several aims to overcome previous design limitations.

First, experiments should be designed so that contributions from processing accented speech and effortful processing can be teased apart (Venezia et al., 2012). Second, studies should aim to distinguish between brain activity related to processing accent variation and other distortions, such as background noise. Adank et al. (2012a) contrasted sentences in a familiar accent embedded in background noise with sentences in an unfamiliar accent, to disentangle areas associated with processing accent-related variation from those associated with processing speech in background noise: Left posterior temporal areas in STG (extending to PT) and right STG (extending into insula) were more activated for accented speech than speech in noise, while bilateral FO/insula were more activated for speech in noise compared to accented speech, indicating that the neural architecture for processing accented speech and speech in background noise is not generic. Third, different accents vary in how much they deviate from the listener's own accent. Greater deviation between accents is associated with greater processing cost, but the neural response associated with variations in distance between accents has not been explored using fMRI. A recent study using Transcranial Magnetic Stimulation (TMS) showed a causal role for lip and tongue motor cortex in perceived speaker and listener distance processing (Bartoli et al., 2013). Another study used EEG to show that regional and foreign accents might be processed differently: processing sentences in an unfamiliar foreign accent reduces the size of the N400 compared to unfamiliar native accents (Goslin et al., 2012). It may be fruitful to use a wider variety of neuroscience techniques, including (combinations of) fMRI, EEG, MEG, and TMS, to investigate how the brain successfully accomplishes accented speech perception. Third, as processing effort, or cognitive load, is inevitably confounded with processing unfamiliar variation in accented speech, experiments should be designed to identify neural substrates associated with processing accent variation and those associated with increased cognitive load. One possibility would be to examine task difficulty and accent processing in a fully crossed factorial design to single out areas that show increased BOLD-activation for accented speech and for task difficulty. Finally, the contribution of production resources to processing accented speech should be examined, to explicitly test predictions from motor and executive recruitment accounts (e.g., Du et al., 2014).

References

Adank, P. (2012a). Design choices in imaging speech comprehension: an Activation Likelihood Estimation (ALE) meta-Analysis. *Neuroimage* 63, 1601–1613. doi: 10.1016/j.neuroimage.2012.07.027

Adank, P. (2012b). The neural bases of difficult speech comprehension and speech production and their overlap: two Activation Likelihood Estimation (ALE) meta-analyses. *Brain Lang.* 122, 42–54. doi: 10.1016/j.bandl.2012.04.014

Adank, P., Davis, M., and Hagoort, P. (2012a). Neural dissociation in processing noise and accent in spoken language comprehension. *Neuropsychologia* 50, 77–84. doi: 10.1016/j.neuropsychologia.2011.10.024

Adank, P., and Devlin, J. T. (2010). On-line plasticity in spoken sentence comprehension: adapting to time-compressed speech. *Neuroimage* 49, 1124–1132. doi: 10.1016/j.neuroimage.2009.07.032

Adank, P., Evans, B. G., Stuart-Smith, J., and Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 520–529. doi: 10.1037/a0013552

Adank, P., and Janse, E. (2010). Comprehension of a novel accent by young and older listeners. *Psychol. Aging* 25, 736–740. doi: 10.1037/a0020054

Adank, P., Noordzij, M. L., and Hagoort, P. (2012b). The role of Planum Temporale in processing accent variation in spoken language comprehension. *Hum. Brain Mapp.* 33, 360–372. doi: 10.1002/hbm.21218

Adank, P., Rueschemeyer, S. A., and Bekkering, H. (2013). The role of accent imitation in sensorimotor integration during processing of intelligible speech. *Front. Hum. Neurosci.* 4:634. doi: 10.3389/fnhum.2013.00634

Banks, B., Gowen, E., Munro, K., and Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *J. Acoust. Soc. Am.* 137, 2015–2024. doi: 10.1121/1.4916265

Bartoli, E., D'Ausilio, A., Berry, J., Badino, L., Bever, T., and Fadiga, L. (2013). Listener–speaker perceived distance predicts the degree of motor contribution to speech perception. *Cereb. Cortex* 25, 281–288. doi: 10.1093/cercor/bht257

Callan, D. E., Callan, A. M., and Jones, J. A. (2014). Speech motor brain regions are differentially recruited during perception of native and foreign-accented phonemes for first and second language listeners. *Front. Neurosci.* 8:275. doi: 10.3389/fnins.2014.00275

Callan, D. E., Jones, J. A., Callan, A. M., and Akahane-Yamada, R. (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory – auditory/orosensory internal models. *Neuroimage* 22, 1182–1194. doi: 10.1016/j.neuroimage.2004.03.006

Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A. R., and Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Front. Psychol.* 3:479. doi: 10.3389/fpsyg.2012.00479

Du, Y., Buchsbaum, B., Grady, C. L., and Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7126–7131. doi: 10.1073/pnas.1318738111

Eickhoff, S. B., Heim, S., Zilles, K., and Amunts, K. (2006). Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. *Neuroimage* 32, 570–582. doi: 10.1016/j.neuroimage.2006.04.204

Eickhoff, S. B., Paus, T., Caspers, S., Grosbras, M. H., Evans, A., Zilles, K., et al. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage* 36, 511–521. doi: 10.1016/j.neuroimage.2007.03.060

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335. doi: 10.1016/j.neuroimage.2004.12.034

Erb, J., Henry, M. J., Eisner, F., and Obleser, J. (2013). The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *J. Neurosci.* 33, 10688–10697. doi: 10.1523/JNEUROSCI.4596-12.2013

Floccia, C., Goslin, J., Girard, F., and Konopczynski, G. (2006). Does a regional accent perturb speech processing? *J. Exp. Psychol. Hum. Percept. Perform.* 32, 1276–1293. doi: 10.1037/0096-1523.32.5.1276

Goslin, J., Duffy, H., and Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. *Brain Lang.* 122, 92–102. doi: 10.1016/j.bandl.2012.04.017

Hesling, I., Dilharreguy, B., Bordessoules, M., and Allard, M. (2012). The neural processing of second language comprehension modulated by the degree of proficiency: a listening connected speech FMRI study. *Open Neuroimag. J.* 6, 1–11. doi: 10.2174/1874440001206010044

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113

Janse, E., and Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Q. J. Exp. Psychol.* 65, 1563–1585. doi: 10.1080/17470218.2012.658822

Moisala, M., Salmela, V., Salo, E., Carlson, S., Vuontela, V., Salonen, O., et al. (2015). Brain activity during divided and selective attention to auditory and visual sentence comprehension tasks. *Front. Hum. Neurosci.* 9:86. doi: 10.3389/fnhum.2015.00086

Obleser, J., and Eisner, F. (2009). Pre-lexical abstraction of speech in the auditory cortex. *Trends Cogn. Sci.* 13, 14–19. doi: 10.1016/j.tics.2008.09.005

Peelle, J. E., Eason, R. J., Schmitter, S., Schwarzbauer, C., and Davis, M. H. (2010). Evaluating an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory processing. *Neuroimage* 52, 1410–1419. doi: 10.1016/j.neuroimage.2010.05.015

Peelle, J. E., McMillan, C., Moore, P., Grossman, M., and Wingfield, A. (2004). Dissociable patterns of brain activity during comprehension of rapid and syntactically complex speech: evidence from fMRI. *Brain Lang.* 91, 315–325. doi: 10.1016/j.bandl.2004.05.007

Perani, D., and Abutalebi, J. (2005). The neural basis of first and second language processing. *Curr. Opin. Neurobiol.* 15, 202–206. doi: 10.1016/j.conb.2005.03.007

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., et al. (1996). Brain processing of native and foreign languages. *Neuroreport* 15–17, 2439–2444. doi: 10.1097/00001756-199611040-00007

Pickering, M. J., and Garrod, S. (2013). An integrated theory of language production and comprehension. *Behav. Brain Sci.* 36, 329–347. doi: 10.1017/S0140525X12001495

Poldrack, R. A., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenich, M., et al. (2001). Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI. *J. Cogn. Neurosci.* 13, 687–697. doi: 10.1162/089892901750363235

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. doi: 10.1038/nn.2331

Rodríguez-Fornells, A., Cunillera, T., Mestres-Missé, A., and de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 3711–3735. doi: 10.1098/rstb.2009.0130

Van Engen, K. J., and Peelle, J. E. (2014). Listening effort and accented speech. *Front. Hum. Neurosci.* 8:577. doi: 10.3389/fnhum.2014.00577

Venezia, J. H., Saberi, K., Chubb, C., and Hickok, G. (2012). Response bias modulates the speech motor system during syllable discrimination. *Front. Psychol.* 3:157. doi: 10.3389/fpsyg.2012.00157

Watkins, K. E., Strafella, A. P., and Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41, 989–994. doi: 10.1016/S0028-3932(02)00316-0

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., and Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *J. Neurosci.* 32, 14010–14021. doi: 10.1523/JNEUROSCI.1528-12.2012

Wilson, M., and Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychol. Bull.* 131, 460–473. doi: 10.1037/0033-2909.131.3.460

Yi, H., Smiljanic, R., and Chandrasekaran, B. (2014). The neural processing of foreign-accented speech and its relationship to listener bias. *Front. Hum. Neurosci.* 8:768. doi: 10.3389/fnhum.2014.00768

Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., and Kramer, S. E. (2014). The eye as a window to the listening brain: neural correlates of pupil size as a measure of cognitive listening load. *Neuroimage* 101, 76–86. doi: 10.1016/j.neuroimage.2014.06.069