



Listening effort and accented speech

Kristin J. Van Engen^{1*} and Jonathan E. Peelle²

¹ Department of Psychology, Washington University in St. Louis, St. Louis, MO, USA

² Department of Otolaryngology, Washington University in St. Louis, St. Louis, MO, USA

*Correspondence: kvanengen@wustl.edu

Edited by:

Patti Adank, University College London, UK

Reviewed by:

Caroline Floccia, University of Plymouth, UK

Keywords: listening effort, speech comprehension, accent, speech perception, speech perception in noise

Understanding spoken language requires mapping acoustic input onto stored phonological and lexical representations. Speech tokens, however, are notoriously variable: they fluctuate within speakers, across speakers, and in different acoustic environments. As listeners, we must therefore perceive speech in a manner flexible enough to accommodate acoustic signals that imperfectly match our expectations. When these mismatches are small, comprehension can proceed with minimal effort; when acoustic variations are more substantial, additional cognitive resources are required to process the signal. A schematic model of speech comprehension is shown in **Figure 1**, emphasizing that different degrees of acoustic mismatch will require varying levels of cognitive recruitment. Recent research increasingly supports a critical role for executive processes—such as verbal working memory and cognitive control—in understanding degraded speech (Wingfield et al., 2005; Eckert et al., 2008; Rönnberg et al., 2013). However, to date, the literature has focused on sources of increased acoustic challenge that originate in the listener (hearing loss) or in the listening environment (background noise). Largely unexplored are the cognitive effects of accented speech (i.e., speech produced by a speaker who does not share a native language or dialect with the listener), a ubiquitous source of variability in speech intelligibility. Here we argue that accented speech must also be considered within a framework of listening effort.

LISTENING EFFORT

Recent years have seen an increasing focus on the cognitive effects of acoustic

challenge during speech comprehension (Mattys et al., 2012). A common theme is that when speech is acoustically degraded, it deviates from what listeners are used to (i.e., stored phonological and lexical representations), resulting in a mismatch between expectation and percept (Sohoglu et al., 2012, 2014). As a result, listeners must recruit additional cognitive resources to make sense of degraded speech (Rönnberg et al., 2013). Given that listeners' cognitive resources are limited, at some point the allocation of cognitive resources to resolve acoustic challenge will begin to impinge upon other types of behavior. Indeed, even mild hearing loss has been shown to impact syntactic processing (Wingfield et al., 2006), running memory for speech (McCoy et al., 2005), and subsequent memory for short stories (Piquado et al., 2012). Further support for the connection between acoustic and cognitive processing comes from the fact that behavioral challenges are exacerbated in older adults due to age-related cognitive decline (Wingfield et al., 2005).

If increased executive processing is required to deal with acoustic challenge, the effects should not only be apparent in listeners with hearing loss, but in listeners with good hearing in cases of external auditory interference. Consistent with this view, acoustic distortion reduces the episodic recall of word pairs (Heinrich and Schneider, 2011) or word lists (Rabbitt, 1968; Cousins et al., 2014). Conversely, increasing speech clarity through the use of listener-oriented speech facilitates recognition memory for spoken sentences (Van Engen et al., 2012). Thus, listening effort appears to be a general consequence of challenging speech signals, in which

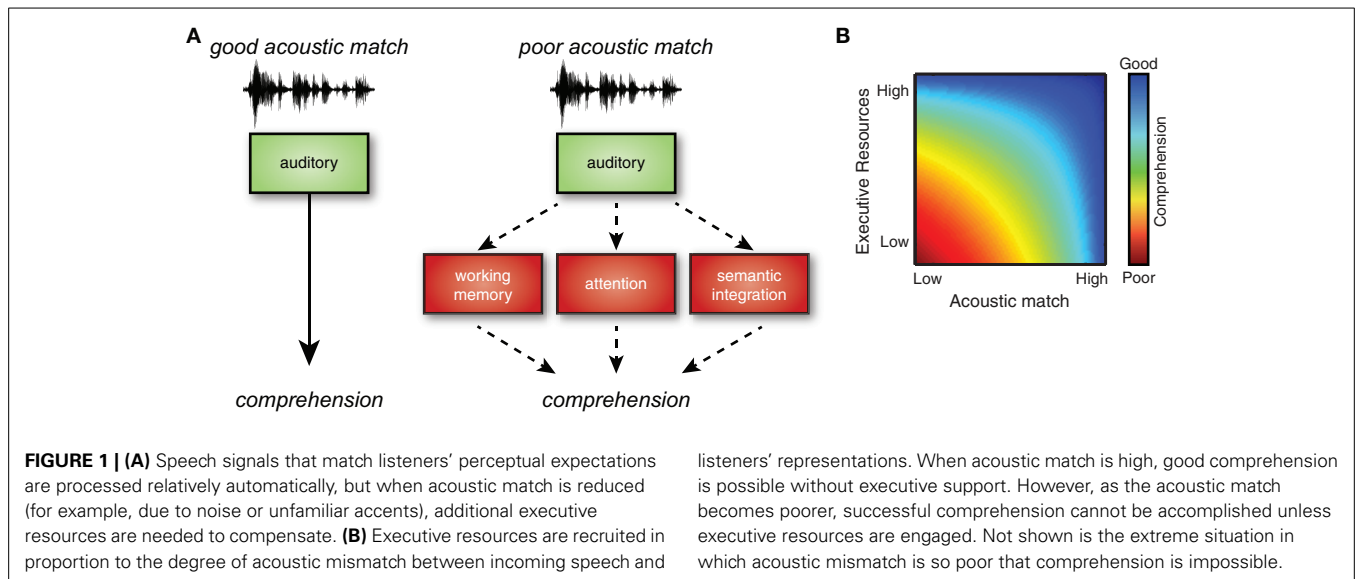
acoustic mismatch can arise due to either internal factors such as hearing loss, or external factors such as background noise.

Functional neuroimaging studies have begun to link these additional executive resources to specific neural systems by identifying increased neural activity resulting from acoustic challenge during speech comprehension (Davis and Johnsrude, 2003; Eckert et al., 2008, 2009; Adank, 2012; Hervais-Adelman et al., 2012; Obleser et al., 2012; Erb et al., 2013). These increases in neural activity frequently involve areas not seen during “normal” speech comprehension—such as frontal operculum, anterior cingulate, and premotor cortex—consistent with listeners' recruitment of additional executive resources to cope with acoustic challenge. Evidence that these increases in brain activity are task-relevant comes from the fact that they vary as a function of attention (Wild et al., 2012), and modulate behavioral performance on subsequent trials (Vaden et al., 2013).

Taken together, then, there is clear evidence that when speech is acoustically degraded, listeners must rely on additional cognitive resources, supported by an extensive network of brain regions. This general principle has been shown in listeners with hearing loss and in good-hearing listeners presented with acoustically degraded materials. In the next section we consider how these findings may play out in the context of understanding accented speech.

LISTENING EFFORT AND ACCENTED SPEECH

If acoustic deviation from stored phonological/lexical representations is indeed the



primary cause of increased listening effort, then speech produced in an unfamiliar accent (whether a regional accent or a foreign accent) should similarly affect not only speech intelligibility, but also the efficiency and accuracy of linguistic processing, and memory for what has been heard. Furthermore, accented speech would also be expected to involve the recruitment of compensatory executive resources.

Foreign-accented speech, for example, is characterized by systematic segmental and/or suprasegmental deviations from native language norms. Naturally, these mismatches can lead to a reduction in the intelligibility of the speech (Gass and Varonis, 1984; Munro and Derwing, 1995; Bent and Bradlow, 2003; Burda et al., 2003; Ferguson et al., 2010; Gordon-Salant et al., 2010a,b). However, even when foreign-accented speech is fully intelligible to listeners (i.e., they can correctly repeat or transcribe it), processing it requires more effort than processing native accents: listeners report that accented speech is more difficult to understand (Munro and Derwing, 1995; Schmid and Yeni-Komshian, 1999), and it is processed more slowly (Munro and Derwing, 1995; Floccia et al., 2009) and comprehended less well than native-accented speech (Anderson-Hsieh and Koehler, 1988; Major et al., 2002). Similar effects have been observed for unfamiliar regional accents: Adank et al. (2009) have shown, for example, that listeners' response times and error rates on

a semantic verification task (i.e., responding to simple true/false questions spoken with different accents) are higher for speech produced in an unfamiliar regional accent. (For a review of the costs associated with processing accented speech across the lifespan, see Cristia et al., 2012.)

The behavioral consequences of listening to accented speech, therefore, include reductions in intelligibility, comprehensibility, and processing speed—all effects that mirror those seen under conditions involving acoustic degradation. To date, there are few functional neuroimaging studies investigating whether increased brain activity is also seen in response to accented speech, although published accounts suggest this is indeed the case (Adank et al., 2012). In general, we would expect that when listening to accented speech, people will recruit comparable executive resources as when listening to other forms of degradation. This would be consistent with increased activity in regions of premotor cortex, inferior frontal gyrus, and the cingulo-opercular network.

That being said, it is important to acknowledge that mismatches between incoming signals and stored representations can arise by different mechanisms. For degraded speech—including steady-state background noise, hearing impairment, or aided listening—listeners experience a loss of acoustic information. This loss is systematic insofar as it involves the inaudibility of a particular

portion of the acoustic signal. In accented speech, there are systematic mismatches between the incoming signal and listeners' expectations, but these arise through phonetic and phonological deviations rather than through signal loss. The degree to which the source of acoustic mismatch affects the type and degree of compensatory cognitive processing required for understanding speech remains an open question. It could be that degraded and accented speech require similar types of executive compensation, and thus both neural and behavioral consequences are largely similar. A second option is that although listeners show similar behavioral consequences to these two types of speech, they are obtained through the use of different underlying neural mechanisms. Finally, there may be differences in both the neural and behavioral consequences of degraded compared to accented speech, or between different types of accented speech. The available preliminary evidence suggests a possible dissociation at the neural level, with different patterns of recruitment for speech in noise compared to accented speech (Adank et al., 2012), and in regional compared to foreign accents (Goslin et al., 2012). However, additional data are needed, and the results may also depend on the level of spoken language processing being tested (Peelle, 2012), task demands, and other factors that determine cognitive challenge for listeners.

ADDITIONAL CONTRIBUTIONS TO LISTENING EFFORT

There are undoubtedly a number of additional influences on the perception of accented speech which may not be relevant for acoustically degraded speech. These include familiarity with an accent (Gass and Varonis, 2006), cultural expectations (Hay and Drager, 2010), and intrinsic listener motivation (Evans and Iverson, 2007). Acoustic familiarity may be specifically related to speech, or simply reflect the experience of a particular listener (Holt, 2006). Together, this confluence of factors can interact with acoustic mismatch to determine the degree of perceptual effort experienced by listeners.

WHY DO THE COGNITIVE CONSEQUENCES OF ACCENTED SPEECH MATTER?

If understanding accented speech indeed requires additional cognitive support, then listeners are likely to have greater difficulty not only understanding their accented interlocutors (i.e., reduced intelligibility), but also comprehending and remembering what they have said, and possibly in managing other information or tasks while listening to accented speech. Given the ubiquity of accented speakers (both foreign and regional) in contemporary society, the practical implications of these problems are wide-ranging. Consider, for example, classrooms with foreign-accented teachers or medical settings where patients and medical personnel who *do* share a language may nevertheless *not* speak with similar accents. In such situations the compensatory cognitive processing that can often (though not always) maintain high intelligibility between speakers and listeners may still come at a cost to listeners' ability to encode critical information. Within the context of a broader framework for effortful listening, it is clear that such challenges will be further exacerbated in the frequently-encountered case of acoustic degradation (such as from background noise or hearing loss), where mismatches between incoming speech and listeners expectations can arise from *both* loss of acoustic information and from distortion due to accent. It has been observed, for example, that noisy or reverberant listening environments disproportionately reduce the intelligibility

of foreign-accented speech as compared to native-accented speech (Van Wijngaarden et al., 2002; Rogers et al., 2006).

An important point is that effortful listening is not an all-or-none phenomenon; rather, the level of cognitive compensation required will depend on the degree of acoustic mismatch in any given listening situation. A relatively mild accent, for example, or one that is highly familiar to a particular listener, can be well understood and require little to no additional effort. Furthermore, we know that listeners can rapidly adapt to both foreign-accented speech (Clarke and Garrett, 2004; Bradlow and Bent, 2008; Sidaras et al., 2009; Baese-Berk et al., 2013) and speech produced in unfamiliar regional accents (Clopper and Bradlow, 2008; Maye et al., 2008; Adank and Janse, 2010). Assuming that understanding accented speech is cognitively challenging due to mismatches between signals and listener expectations, as suggested by the general model of effortful listening presented here, it follows that such perceptual adaptation to an accent will decrease listening effort, and thereby *increase* functional cognitive capacity: Adaptation effectively reduces the mismatch between incoming speech and listener expectations, thus lowering the demand for compensatory executive processes (Figure 1). Auditory training with accented speech may therefore not only be useful for improving intelligibility, but also for increasing listeners' cognitive capacity¹.

CONCLUSIONS

When speech does not conform to listeners' expectations, additional cognitive processes are required to facilitate comprehension. In the case of acoustic degradation, it is increasingly accepted that this type of effortful listening can interfere with subsequent attention, language, and memory processes. Here we have argued that accented speech shares critical characteristics with acoustically degraded speech, and that considering the cognitive consequences of acoustic mismatch is critical

¹It is possible that improvements over the course of perceptual learning also rely to some degree on executive processes. So although the most straightforward prediction is that perceptual adaption will reduce the demand on executive resources, the degree to which this actually happens is an empirical question.

in understanding how listeners deal with accented speech.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the Dana Foundation and the National Institute on Aging of the National Institutes of Health under award number R01AG038490.

REFERENCES

- Adank, P. (2012). The neural bases of difficult speech comprehension and speech production: two Activation Likelihood Estimation (ALE) meta-analyses. *Brain Lang.* 122, 42–54. doi: 10.1016/j.bandl.2012.04.014
- Adank, P., Davis, M. H., and Hagoort, P. (2012). 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., 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. Human.* 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
- Anderson-Hsieh, J., and Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Lang. Learn.* 38, 561–613. doi: 10.1111/j.1467-1770.1988.tb00167.x
- Baese-Berk, M. M., Bradlow, A. R., and Wright, B. A. (2013). Accent-independent adaptation to foreign accented speech. *J. Acoust. Soc. Am.* 133, EL174–EL180. doi: 10.1121/1.4789864
- Bent, T., and Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *J. Acoust. Soc. Am.* 114, 1600–1610. doi: 10.1121/1.1603234
- Bradlow, A. R., and Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition* 106, 707–729. doi: 10.1016/j.cognition.2007.04.005
- Burda, A. N., Scherz, J. A., Hageman, C. F., and Edwards, H. T. (2003). Age and understanding speakers with Spanish or Taiwanese accents. *Percept. Mot. Skills* 97, 11–20. doi: 10.2466/pms.2003.97.1.11
- Clarke, C. M., and Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *J. Acoust. Soc. Am.* 101, 2299–2310. doi: 10.1121/1.1815131
- Clopper, C. G., and Bradlow, A. R. (2008). Perception of dialect variation in noise: intelligibility and classification. *Lang. Speech* 51, 175–198. doi: 10.1177/0023830908098539
- Cousins, K. A. Q., Dar, H., Wingfield, A., and Miller, P. (2014). Acoustic masking disrupts time-dependent mechanisms of memory encoding in word-list recall. *Mem. Cogn.* 42, 622–638. doi: 10.3758/s13421-013-0377-7
- Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., and Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Front. Psychol.* 3:479. doi: 10.3389/fpsyg.2012.00479

- Davis, M. H., and Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *J. Neurosci.* 23, 3423–3431.
- Eckert, M. A., Menon, V., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., et al. (2009). At the heart of the ventral attention system: the right anterior insula. *Hum. Brain Mapp.* 30, 2530–2541. doi: 10.1002/hbm.20688
- Eckert, M. A., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., and Dubno, J. R. (2008). Age-related effects on word recognition: reliance on cognitive control systems with structural declines in speech-responsive cortex. *JARO* 9, 252–259. doi: 10.1007/s10162-008-0113-3
- 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
- Evans, B. G., and Iverson, P. (2007). Plasticity in vowel perception and production: a study of accent change in young adults. *J. Acoust. Soc. Am.* 121, 3814–3826. doi: 10.1121/1.2722209
- Ferguson, S. H., Jongman, A., Sereno, J. A., and Keum, K. A. (2010). Intelligibility of foreign-accented speech for older adults with and without hearing loss. *J. Am. Acad. Audiol.* 21, 153–162. doi: 10.3766/jaaa.21.3.3
- Floccia, C., Butler, J., Goslin, J., and Ellis, L. (2009). Regional and foreign accent processing in English: can listeners adapt? *J. Psycholinguist. Res.* 38, 379–412. doi: 10.1007/s10936-008-9097-8
- Gass, S., and Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of non-native speech. *Lang. Learn.* 34, 65–87. doi: 10.1111/j.1467-1770.1984.tb00996.x
- Gass, S., and Varonis, E. M. (2006). The effect of familiarity on the comprehensibility of non-native speech. *Lang. Learn.* 34, 65–87. doi: 10.1111/j.1467-1770.1984.tb00996.x
- Gordon-Salant, S., Yeni-Komshian, G. H., and Fitzgibbons, P. J. (2010a). Recognition of accented English in quiet by younger normal-hearing listeners and older listeners with normal-hearing and hearing loss. *J. Acoust. Soc. Am.* 128, 444–455. doi: 10.1121/1.3397409
- Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., and Schurman, J. (2010b). Short-term adaptation to accented English by younger and older adults. *J. Acoust. Soc. Am.* 128, EL200–EL204. doi: 10.1121/1.3486199
- 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
- Hay, J., and Drager, K. (2010). Stuffed toys and speech perception. *Linguistics* 48, 865–892. doi: 10.1515/ling.2010.027
- Heinrich, A., and Schneider, B. A. (2011). Elucidating the effects of ageing on remembering perceptually distorted word pairs. *Q. J. Exp. Psychol.* 64, 186–205. doi: 10.1080/17470218.2010.492621
- Hervais-Adelman, A., Carlyon, R. P., Johnsrude, I. S., and Davis, M. H. (2012). Brain regions recruited for the effortful comprehension of noise-vocoded words. *Lang. Cogn. Process.* 27, 1145–1166. doi: 10.1080/01690965.2012.662280
- Holt, L. L. (2006). The mean matters: effects of statistically defined nonspeech spectral distributions on speech categorization. *J. Acoust. Soc. Am.* 120, 2801–2817. doi: 10.1121/1.2354071
- Major, R. C., Fitzmaurice, S. F., Bunta, F., and Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: implications for ESL assessment. *TESOL Q.* 36, 173–190. doi: 10.2307/3588329
- Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: a review. *Lang. Cogn. Process.* 27, 953–978. doi: 10.1080/01690965.2012.705006
- Maye, J., Aslin, R. N., and Tanenhaus, M. K. (2008). The weckud wetch of the wast: lexical adaptation to a novel accent. *Cogn. Sci.* 32, 543–562. doi: 10.1080/03640210802035357
- McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R., and Wingfield, A. (2005). Hearing loss and perceptual effort: downstream effects on older adults' memory for speech. *Q. J. Exp. Psychol.* 58, 22–33. doi: 10.1080/02724980443000151
- Munro, M. J., and Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Lang. Learn.* 45, 73–97. doi: 10.1111/j.1467-1770.1995.tb00963.x
- Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., and Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *J. Neurosci.* 32, 12376–12383. doi: 10.1523/JNEUROSCI.4908-11.2012
- Peelle, J. E. (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
- Piquado, T., Benichov, J. I., Brownell, H., and Wingfield, A. (2012). The hidden effect of hearing acuity on speech recall, and compensatory effects of self-paced listening. *Int. J. Audiol.* 51, 576–583. doi: 10.3109/14992027.2012.684403
- Rabbitt, P. M. A. (1968). Channel capacity, intelligibility and immediate memory. *Q. J. Exp. Psychol.* 20, 241–248. doi: 10.1080/14640746808400158
- Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., and Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Appl. Psycholinguist.* 27, 465–485. doi: 10.1017/S014271640606036X
- Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031
- Schmid, P. M., and Yeni-Komshian, G. H. (1999). The effects of speaker accent and target predictability on perception of mispronunciations. *J. Speech Lang. Hear. Res.* 42, 56–64.
- Sidasar, S. K., Alexander, J. E. D., and Nygaard, L. C. (2009). Perceptual learning of systematic variation in Spanish-accented speech. *J. Acoust. Soc. Am.* 125, 3306–3316. doi: 10.1121/1.3101452
- Sohoglu, E., Peelle, J. E., Carlyon, R. P., and Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *J. Neurosci.* 32, 8443–8453. doi: 10.1523/JNEUROSCI.5069-11.2012
- Sohoglu, E., Peelle, J. E., Carlyon, R. P., and Davis, M. H. (2014). Top-down influences of written text on perceived clarity of degraded speech. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 186–199. doi: 10.1037/a0033206
- Vaden, K. I. J., Kuchinsky, S. E., Cute, S. L., Ahlstrom, J. B., Dubno, J. R., and Eckert, M. A. (2013). The cingulo-opercular network provides word-recognition benefit. *J. Neurosci.* 33, 18979–18986. doi: 10.1523/JNEUROSCI.1417-13.2013
- Van Engen, K. J., Chandrasekaran, B., and Smiljanic, R. (2012). Effects of speech clarity on recognition memory for spoken sentences. *PLoS ONE* 7:e43753. doi: 10.1371/journal.pone.0043753
- Van Wijngaarden, S., Steeneken, H., and Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *J. Acoust. Soc. Am.* 111, 1906–1916. doi: 10.1121/1.1456928
- Wild, C. J., Yusuf, A., Wilson, D., 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
- Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A., and Cox, C. (2006). Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. *J. Am. Acad. Audiol.* 17, 487–497. doi: 10.3766/jaaa.17.7.4
- Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in older adulthood: what it is and how it interacts with cognitive performance. *Curr. Dir. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 21 May 2014; paper pending published: 08 June 2014; accepted: 14 July 2014; published online: 05 August 2014.

Citation: Van Engen KJ and Peelle JE (2014) Listening effort and accented speech. *Front. Hum. Neurosci.* 8:577. doi: 10.3389/fnhum.2014.00577

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Van Engen and Peelle. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.