Pronunciationforteachers.com/ Key Concepts

Speech and its Non-verbal Counterparts

Some researchers and teachers may consider speech to be a unimodal (i.e., speaking and listening only) phenomenon; however, when more than verbal cues are present, it is important to consider the bigger picture of a speech event. Findings from perception and neuroimaging studies have demonstrated that speech is a multimodal phenomenon (e.g., Hardison, 2012; Rosenblum, 2010; Skipper, van Wassenhove, Nusbaum, & Small, 2007). The visual affiliates of a speaker's verbal message include lip movements, brow movements, eye contact, facial expressions, manual gestures, body position, etc. Although such features of nonverbal communication serve unique functions (Gullberg, 2011), they also contribute to comprehension for a variety of perceivers, including the hearing impaired (e.g., Goldin-Meadow & Alibali, 2013), non-impaired listeners (e.g., Summerfield, 1979), and second-language (L2) learners of English (e.g., Hardison, 2018; Sueyoshi & Hardison, 2005), French (Inceoglu, 2016), and Japanese (Mori & Hayashi, 2006).

Research has shown a coordination of eyebrow raises and eyeblinks with pitch for English speakers (e.g., Flecha-García, 2010), and a correlation between head movement, voice amplitude, and fundamental frequency (the acoustic correlate of pitch) for Japanese speakers (Munhall, Jones, Callan, Kuratate, & Vatikiotis-Bateson, 2004). Krahmer and Swerts (2007) found that Dutch speakers who were instructed to use a visual beat with the hand, a head nod, or brow movement when producing a word showed corresponding acoustic effects such as longer word duration, regardless of the type of visual beat. Further, perceivers reported that the words with visual beats were more prominent, especially the words accompanied by manual beats.

Modern technology allows us to take a closer look at how components relate to one another in oral communication. Two recent studies used Praat (Boersma & Weenink, 2014) for phonetic analysis and annotation, and Anvil (Kipp, 2001, http://www.anvil-software.org/) for video annotation to provide an integrated time-aligned display of the relationship between the acoustic and visual elements of several speech events. In one study, three native Japanese-speaking teachers were videorecorded while teaching low-proficiency adult L1 English speakers in a classroom in Japan (Hardison & Reed, 2018). The content of the lessons included, but was not limited to, minimal pairs contrasting long and short vowels. Vowel duration serves a contrastive function in Japanese and is an important element for communication. The teachers, unaware of the focus of the study, exhibited two categories of gesture: head nods and manual gestures. Analysis revealed a significant relationship between the duration of the vowel each teacher produced in a word and the type of co-occurring gesture. More than 75% of the head nods for each teacher occurred with a long vowel; manual gestures tended to occur with a short vowel. Further analysis revealed that when the long vowels were produced, the peak of the syllable and pitch along with the peak or apex of the head nod were temporally aligned. To explore whether

the visual elements influenced the accuracy with which L2 Japanese learners could identify vowel duration, the above stimuli were presented to another group of learners (L1 English) who were unfamiliar with the teachers. Results revealed greater identification accuracy of vowel duration the in auditory-visual presentation (vs. auditory- or visual-only) presentation in which they could both see and hear the teachers.

In the second study, videorecordings of native and advanced nonnative English-speaking instructors showed temporal alignment of head movement, eyebrow raise, pitch peak, and manual gesture (Hardison, in press). However, within emphatic intonation phrases, a second rhythmic pattern emerged such that the interval between the apices of gestural beats (e.g., back-and-forth hand movements) increased in duration surrounding pitch-accented or stressed vowels, which were also longer in duration compared to other vowels. These types of gestural sequences serve as visual highlighters of specific lexical information (Dimitrova, Chu, Wang, Özyürek, & Hagoort, 2016). From a theoretical perspective, some researchers have suggested that the entrainment of gestural and speech rhythms may result from their joint production (e.g., Treffner, Peter, & Kleidon, 2008). From a pedagogical perspective, the study also found that implementing a multistep technology-assisted program to enhance the verbal and nonverbal communication skills of international teaching assistants improved their confidence. Based on course evaluations, their teaching effectiveness and ability to communicate important points also improved. Nonverbal communication is particularly useful for instructors in conveying ideas, especially emphasis, and is associated with students' perceptions of positive rapport (e.g., McCroskey, Richmond, Sallinen, Fayer, & Barraclough, 1995).

The presence of nonverbal cues may impact the perception of L2 interactive competence. In a recent study, Crowther (2018) found that native English-speaking raters' scores for L2 English learners' accentedness (degree of difference between L2 learners' speech and the target variety) and comprehensibility (ease of understanding) did not predict their assessments of the learners' performance on an interactive task in contrast to the findings for monologic tasks. Because the assessment procedure for the interactive task included a video of the interaction, raters may have been influenced by elements of nonverbal communication such as gesture, facial expression, and body language. Appropriate nonverbal behaviors can have a positive impact on the perception of an L2 speaker's proficiency; in contrast, the perception of proficiency may be lowered for an L2 speaker who uses inappropriate or no nonverbal communication strategies (Neu, 1990). In addition, assessments of interactive competence may be influenced by important elements such as turn-taking, topic initiation, and elaboration. Therefore, language teachers, eager to assist learners in enhancing the effectiveness of their performance in L2 interactive situations should consider the totality of the speech event by addressing both verbal and nonverbal elements of oral communication such as the appropriate use of gestures, eye contact, and facial expression.

References

Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer [Computer program]. Retrieved from http://www.fon.hum.uva.nl/praat/

Crowther, D. J. (2018). *Linguistic measures of second language speech: Moving from monologic to interactive speech* (Doctoral dissertation).

Dimitrova, D., Chu, M., Wang, L., Özyürek, A., & Hagoort, P. (2016). Beat that word: How listeners integrate beat gesture and focus on multimodal speech discourse. *Journal of Cognitive Neuroscience, 28,* 1255–1269. doi:10.1162/jocn_a_00963

Flecha-García, M. L. (2010). Eyebrow raises in dialogue and their relation to discourse structure, utterance function and pitch accents in English. *Speech Communication, 52,* 542–554. doi:10.1016/j.specom.2009.12.003

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology, 64*, 257–283. doi:10.1146/annurev-psych-113011-143802

Gullberg, M. (2011). Multilingual multimodality: Communicative difficulties and their solutions in second language use. In J. Streeck, C. Goodwin, & C. LeBaron (Eds.), *Embodied interaction: Language and body in the material world* (pp. 137–151). Cambridge: Cambridge University Press.

Hardison, D. M. (in press). Visualizing the acoustic and gestural beats of emphasis in multimodal discourse: Theoretical and pedagogical implications. *Journal of Second Language Pronunciation*.

Hardison, D. M. (2018). Effects of contextual and visual cues on spoken language processing: Enhancing L2 perceptual salience through focused training. In S. M. Gass, P. Spinner, & J. Behney (Eds.), *Salience in second language acquisition* (pp. 201-220). New York: Routledge.

Hardison, D. M. (2012). Second-language speech perception: A cross-disciplinary perspective on challenges and accomplishments. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 349–363). London: Routledge.

Hardison. D. M., & Reed, J. T. (2018). *Relationships among gesture type, pitch, and segmental duration in the speech of teachers of Japanese and their influence on learner perception.* Manuscript in preparation.

Inceoglu, S. (2016). Effects of perceptual training on second language vowel perception and production. *Applied Psycholinguistics, 37*, 1175–1199. doi:10.1017/S0142716415000533

Kipp, M. (2001). *Anvil–A generic annotation tool for multimodal dialogue. In *Proceedings of the 7th European Conference on Speech Communication and Technology* (pp. 1367–1370). Aalborg, Denmark: Eurospeech.

Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language, 57,* 396–414. doi:10.1016/j.jml.2007.06.005

McCroskey, J. C., Richmond, V. P., Sallinen, A., Fayer, J. M., & Barraclough, R. A. (1995). A cross-cultural and multi-behavioral analysis of the relationship between nonverbal immediacy and teacher evaluation. *Communication Education, 44,* 281–291.

Mori, J., & Hayashi, M. (2006). The achievement of intersubjectivity through embodied completions: A study of interactions between first and second language speakers. *Applied Linguistics, 27,* 195–219. doi:10.1093/applin/aml014

Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility: Head movement improves auditory speech perception. *Psychological Science, 15,* 133–137. doi:10.1111/j.0963-7214.2004.01502010.x

Neu, J. (1990). Assessing the role of nonverbal communication in the acquisition of communicative competence in L2. In R. C. Scarcella, E. S. Andersen, & S. D. Krashen (Eds.), *Developing communicative competence in a second language* (pp. 121–138). Boston: Heinle and Heinle.

Rosenblum, L. (2010). *See what I'm saying: The extraordinary powers of our five senses*. New York: W. W. Norton & Co.

Skipper, J. I., van Wassenhove, V., Nusbaum, H. W., & Small, S. L. (2007). Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex, 17*, 2387–2399. doi:10.1093/cercor/bhl147

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second-language listening comprehension. *Language Learning, 55,* 661–699. doi:10.1111/j.0023-8333.2005.00320.x

Summerfield, Q. (1979). Use of visual information for phonetic perception. *Phonetica, 36,* 314–331.

Treffner, P., Peter, M., & Kleidon, M. (2008) Gestures and phases: The dynamics of speech-hand communication. *Ecological Psychology, 20,* 32–64. doi: 10.1080/10407410701766643