The Coordination between Beat Gestures and Prosodic Prominence in Chinese EFL Learners' Academic Oral Presentation

Yating Cao[1,2], *Hua Chen*[2]

[1]Nanjing University; University of Warwick

[2]Nanjing University

yatingcao@smail.nju.edu.cn, chenhua@nju.edu.cn

Human communication is multimodal in nature. Gestures, as one of the most important non-verbal cues, often accompany people's speech production, providing another window into people's minds (Kendon, 2004; McNeill, 1992). Pitch accent, a form of prosodic specification, highlights specific segments (syllables or words) within a sentence to convey focus or information status (Cutler, 1984). Beat gesture, according to McNeill (1992), refers to "a simple flick of the hand or fingers up and down; or back and forth". This type of gesture has been found to be the most frequently used gesture in daily conversation, particularly taking up 94.6% of gesture types in academic-style conversations (Shattuck-Hafnagel & Ren, 2018). Similar to pitch accent, beat gesture is believed to emphasize information status, acting as a "gestural yellow highlight" (McNeill, 2006). There is increasing evidence that beat gesture and pitch accent are closely coordinated. For instance, Yasinnik et al. (2004) found a significant coupling between beat gesture and pitch accent, with 86% of beat gestures produced by an Australian English speaker aligned with a pitch accent. Shattuck-Hafnagel & Ren (2018) analyzed 1,334 beat gestures in a 20-min academic speech, revealing a close coordination between beat gesture and the prosodic structure of spoken utterances.

While the coordination between prosodic and gestural prominence is well established in prior research, most studies have focused on non-tonal languages, such as English (Loehr, 2012), Catalan (Esteve-Gibert et al., 2014), and Dutch (De Ruiter, 1998). It would be interesting to ask how prosody and gesture coordinate in non-stress tone languages. Furthermore, it should be noted that research on prosody-gesture coordination has primarily concentrated on native speakers, with limited exploration of non-native speakers. Graziano et al. (2020) investigated the phonological coordination of gesture and speech in English and French monolinguals and bilinguals during a cartoon retelling task. Results showed that there was no significant difference concerning the time-aligned coordination of gesture and speech in these two groups, which is possibly explained by the relatively high level of proficiency in the two languages. In light of the findings, it would be important to further test participants at different proficiency levels to better inform the nature of gesture-speech links and theories of language development as well (Gullberg, 2022; Shattuck-Hafnagel, 2019). Interestingly, Mandarin Chinese is a tonal language characterized by a syllable-timed rhythmic pattern; hence, Chinese EFL learners might accentuate every word within a sentence without clear communicative intentions as a result of the L1 transfer (Chen & Wang, 2015). In this sense, investigating how this group of learners coordinates prosodic and gestural prominence in their L2 is of great importance and warrants further investigation.

To respond to these gaps, the present study took a preliminary attempt to analyze the academic oral presentation of Chinese EFL learners and addressed the following questions: (1) How does beat gesture align with prosodic prominence in Chinese EFL learners' speech production? (2) Is there any difference between proficiency levels?

The data for this study were obtained from the Multimodal Corpus of Academic English Speech by Chinese Learners (Chen et al., 2022). Participants in this corpus were required to deliver a two-minute academic oral presentation in English. They were randomly assigned to either the informative speech with the topic "How to prepare for an oral presentation?" or the persuasive speech with the topic "To read selectively or extensively? What is your understanding? Support your ideas with details." All the speech samples have been rated by 15 linguistic experts from various linguistic dimensions (i.e., phonetics, syntax, vocabulary, and discourse) via focused analytical scoring, and all the results have been validated through multiple methods. Based on the total score, all the samples could be ranked and categorized into different proficiency levels (i.e., advanced, intermediate, and low). Given the exploratory nature of the present study, a modest sample size of 12 speech samples was selected, evenly distributed between the advanced and intermediate groups. Additionally, all speakers discussed the same topic and achieved the gender balance.

The data annotation consisted of two parts – prosody and gesture. To eliminate potential cross-modal effects, these two annotations were independently completed. For prosodic annotation, the speech was initially transcribed orthographically and labeled for its intonational structure – mainly for pitch accents, using ToBI as the prosodic annotation system in Praat. And all the gesture annotation was accomplished in ELAN and three main tiers were coded, including orthographic transcription, manual gesture phase, and prosodic information. The first author conducted the annotation, and a second coder annotated 25% of randomly-selected data to establish inter-rater reliability and the Cohen's Kappa value (k=0.82) indicated the high agreement.

The results indicated that intermediate-level Chinese EFL learners employed beat gestures irrespective of the placement of pitch accents and the information status of their accompanying speech. Beat gestures were predominantly used to serve a meta-linguistic function. In contrast, proficient learners produced significantly fewer beat gestures, and when present, these gestures were more likely to co-occur with pitch accents, emphasizing crucial discourse information to fulfill a meta-pragmatic function. These findings offer empirical evidence from L2 learners, contributing to the existing literature on multimodal prosody and providing pedagogical implications for English teaching.

Index Terms: prosodic prominence, beat gesture, speech-gesture coordination, proficiency level

References

Chen, H., Wang, X. 2015. A study of English phrasal accentuation in Chinese EFL learners' read speech. *Foreign Languages and Their Teaching*, *3*, 13-18+40.

Chen, H., Cao, Y., Ma, D. 2022. MCAESCL-A Multimodal Corpus of Academic English Speech by Chinese Learners. *Technology Enhanced Foreign Language Education*, *3*, 45- 51+109.

Cutler, A. 1984. Stress and accent in language production and understanding. In D. Gibbon & H. Richter (Eds.), I*ntonation, Accent and Rhythm*. De Gruyter, 77-90.

De Ruiter, J. P. 1998. *Gesture and speech production*. Unpublished Doctoral Dissertation. Nijmegen: Radboud University.

Esteve-Gibert, N., Pons, F., Bosch, L., & Prieto, P. 2014. Are gesture and prosodic prominences always coordinated? Evidence from perception and production. *Proceedings of the International Conference on Speech Prosody 2014* (Dublin, Ireland), 222-226.

Graziano, M., Nicoladis, E., & Marentette, P. 2020. How referential gestures align with speech: Evidence from monolingual and bilingual speakers. *Language Learning*, *70*(1), 266–304.

Gullberg, M. 2022. The relationship between gestures and speaking in L2 learning. In Derwing, T. M., J. M. Munro & R. I. Thomson (Eds.), *The Routledge handbook of second language acquisition and speaking*. New York: Routledge, 386-398.

Kendon, A. 2004. *Gesture: Visible Action as Utterance*. Cambridge University Press.

Loehr, D. P. 2012. Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology*, *3*(1), 71–89.

McNeill, D. 1992. *Hand and Mind: What Gestures Reveal about Thought.* Chicago: University of Chicago Press.

McNeill, D. 2006. Gesture: A psycholinguistic approach. *The Encyclopedia of Language and Linguistics*, 58–66.

Shattuck-Hufnagel, S. 2019. Toward an (even) more comprehensive model of speech production planning. *Language, Cognition and Neuroscience*, *34*(9), 1202–1213.

Shattuck-Hufnagel, S., & Ren, A. 2018. The prosodic characteristics of non-referential co-speech gestures in a sample of academic-lecture-style speech. *Frontiers in Psychology, 9.*

Yasinnika, Y., Renwick, M., & Shattuck-Hufnagel, S. 2004. The timing of speech-accompanying gestures with respect to prosody. *Proceedings of the International Conference: From Sound to Sense*, 97–102.