# Perception and production of acoustical contrasts for pharyngeal fricatives in non-native learners: a training effect supported by embodied based training.

*Marie Joe Kfoury*[1,2], *Axelle Calcus*[1], *Brigitte Charlier* [1,2]

[1] Université Libre Bruxelles, Center for Research in Cognition and Neurosciences (CRCN), Brussels, Belgium

[2] Centre Comprendre et Parler, Brussels, Belgium

Marie.kfoury@ulb.be, axelle.calcus@ulb.be , brigitte.charlier@ulb.be

Perceiving and producing nonnative phonemes has been viewed as one of the most challenging aspects in L2 acquisition for adult speakers (Lee et al., 2015). A growing number of studies have focused on the role of multimodal perception training in phonological learning (Hazan et al., 2005; Hoetjes & van Maastricht, 2020; Kelly et al., 2017; Xi et al., 2020a). To our knowledge, only one study has focused on the role of embodied training relying on the production of hand gestures on learning new phonological contrast at the segmental level (Li et al., 2021). In this study, participants who performed hand gestures depicting the air burst of Mandarin aspirated plosives during the training significantly improved their production, more so than participants who were trained with no gesture (Li et al., 2021). These promising results show that multimodal training involving hand gestures helps with the pronunciation of new phonological contrasts. However, to examine the potential effect of training with body movements depicting phonetic properties on L2 segmental acquisition, along with assessing the importance of the congruency between body movements and the target phonetic features, it is crucial to add incongruent body movements to the training. Moreover, it should be noted that our study does not solely focus on hand gestures; it encompasses the investigation of the impact of body movements. Therefore, the present study aims at evaluating the potential benefit of using congruent vs. incongruent body movements during an embodied training on both perception and production of Arabic pharyngealized speech units, considering the holistic nature of body movements in the L2 acquisition process.

The body-based training aimed to expose participants to body movements that were thought to be either "congruent" or "incongruent" with the articulatory information of the Arabic pharyngealized consonants. For example, a wringing movement of the hands was used to depict the constriction that happens at the root of the tongue for the syllable [aʕa], while a bending backward of the head with a straight and steady descending movement of the hands was used to emulate the place of articulation (back of the tongue) and the voicelessness of the syllable [aħa]. However, in the case of incongruent body movements, the movements of the hands did not match any specific articulatory features of the sounds, such as a rotating motion of the forearm ending in one of the forearms stretching upward which was very different to the sudden fricative sound for the syllable [aʕa] (see figure 1). Forty-five French speakers were trained to pronounce syllables containing voiced pharyngeal fricative and voiceless pharyngeal fricative. Participants were randomly allocated to one of three training types: (i) congruent movements, (ii) incongruent movements or (iii) no movements. In all three training types, participants were asked to watch the video of someone performing congruent (i), incongruent (ii) or no (iii) movements. In both congruent and incongruent training groups, participants were then asked to imitate the movements they saw on the video. Participants were asked to perform an ABX discrimination task and repetition task, to assess speech perception and production respectively, in a pre-test (i.e., before watching the training videos), an immediate post-test (i.e., just after watching the videos), and a delayed post-test (one week after the test). For the repetition task analysis, five Lebanese native speakers were recruited to blindly rate participants' recordings for syllable pronunciation accuracy.

We predict that producing body movements that are congruent (i) with the target phonetic properties during the training would be more beneficial than producing body movements that are incongruent (ii) with the target phonetic properties they represent, or than not producing body movements (iii). Our study will be one of the first to explore the additional role of an embodied training involving the production of incongruent gestures- rather than simply comparing the use of congruent gesture with no gesture condition (Li et al., 2021).

In the speech perception analysis, a generalized mixed model was employed, utilizing a binary dependent variable coded as 1 for correct responses and 0 for incorrect responses. Independent variables, such as condition (CongM, INCongM, NM), timing (Pre, Post, Home), and contrast type (/ʕ/, /h/), were taken into account, while individual variability was accounted for through the inclusion of a random effect for Participant. Similarly, in the speech production analysis, a generalized mixed model with a Poisson family specification was utilized to handle the dependent variable, consisting of scores ranging from 1 to 10 assigned by Lebanese judges. Independent variables, including condition (CongM, INCongM, NM), timing (Pre, Post, Home), and contrast type (/ʕ/, /h/), were incorporated into the model, along with a random effect for Participant to address individual differences.

Results indicate that after the embodied-based training, using both incongruent and congruent movements, there is a significant improvement in perception of the voiced pharyngeal fricative /ʕ/, while no significant improvement is observed for the unvoiced version /h/ (Figure 2). Similarly, in terms of production, the training leads to a significant enhancement in producing the voiced pharyngeal fricative /ʕ/ for both incongruent and congruent movement groups. However, only the group using incongruent movements shows a significant improvement in producing the unvoiced version /h/. These effects are sustained over a one-week period. Interestingly, at pre-test, participants were poorer at discriminating and producing the voiced than unvoiced pharyngeal fricative.

In addition to the aforementioned results, it is important to note that these findings will be discussed in comparison to the results of a second study we conducted using the same materials. In this study, instead of asking participants to perform the movements, they were instructed to observe the movements on the screen without physically engaging in them. This allows for an examination of whether the effects observed with embodied training are replicable or if perceptual training alone yields similar benefits. Interestingly, the results of this second study indicate that embodied training may have a more beneficial effect than perceptual training alone.

**Index Terms**: Embodied training; second language acquisition; Arabic pharyngealized speech units.

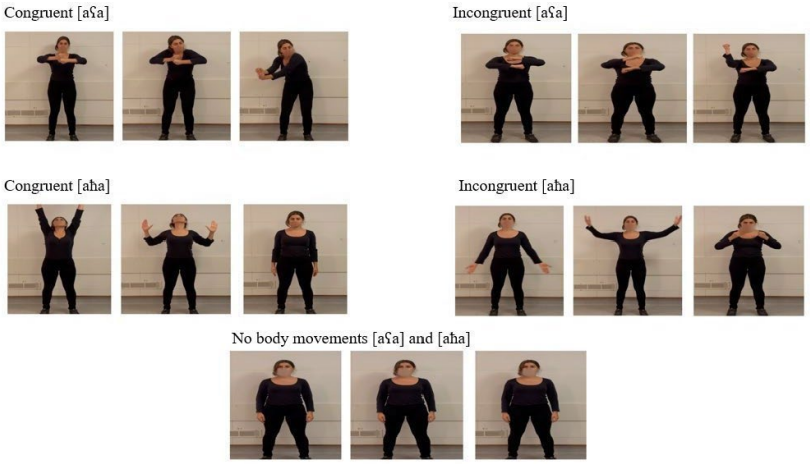


Figure 1: *Body movements across condition (Congruent Body Movement, Incongruent Body Movement, No Body Movement) and contrast type [aʕa], [aħa]).*

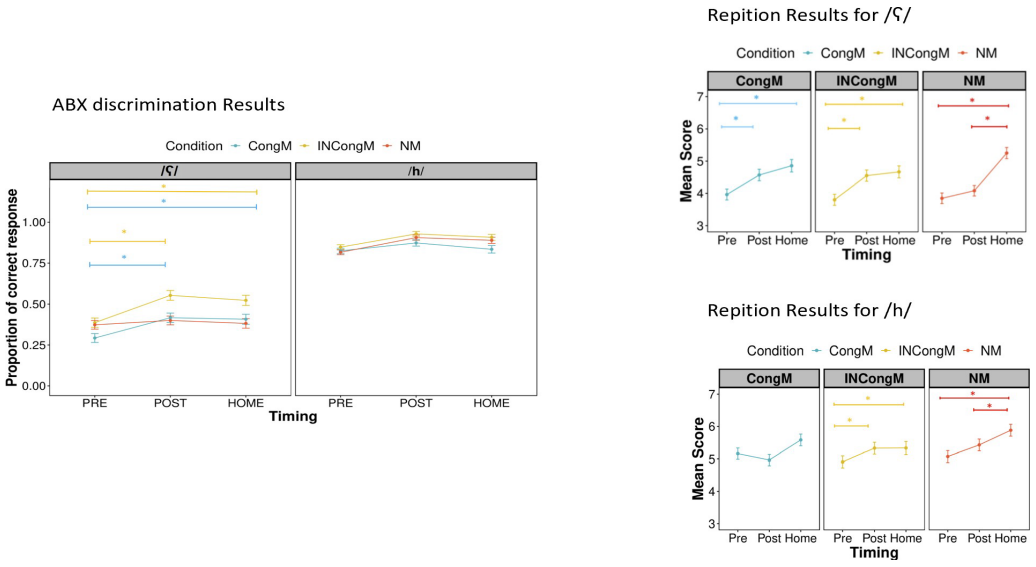


Figure 2: *Proportion of correct response for the ABX discrimination task; on the right side: the judge mean score on a scale from1 to 10 (1 being the lowest and 10 the highest) for the repetition task across conditions (Congruent BM = Body Movements, Incongruent BM and No BM), test sessions (PRE-test, POST-test, and post-test-HOME) for the two contrasts [aʕa], [aħa].*

# References


Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, *47*(3), 360–378. https://doi.org/10.1016/j.specom.2005.04.007

Hoetjes, M., & van Maastricht, L. (2020). Using Gesture to Facilitate L2 Phoneme Acquisition: The Importance of Gesture and Phoneme Complexity. *Frontiers in Psychology*, *11*(November), 1–16. https://doi.org/10.3389/fpsyg.2020.575032

Kelly, S., Bailey, A., & Hirata, Y. (2017). Metaphoric gestures facilitate perception of intonation more than length in auditory judgments of non-native phonemic contrasts. *Collabra: Psychology*, *3*(1), 1–11. https://doi.org/10.1525/collabra.76

Li, P., Xi, X., Baills, F., & Prieto, P. (2021). Training non-native aspirated plosives with hand gestures: learners' gesture performance matters. *Language, Cognition and Neuroscience*, *36*(10), 1313–1328. https://doi.org/10.1080/23273798.2021.1937663

Xi, X., Li, P., Baills, F., & Prieto, P. (2020). Hand gestures facilitate the acquisition of novel phonemic contrasts when they appropriatelymimic target phonetic features. *Journal of Speech, Language, and Hearing Research*, *63*(11), 3571–3585. https://doi.org/10.1044/2020_JSLHR-20-00084

Lee, J., Jang, J., & Plonsky, L. (2015). The Effectiveness of Second Language Pronunciation Instruction: A Meta-Analysis. *Applied Linguistics*, *36*(3), 345–366. https://doi.org/10.1093/applin/amu040