

Visual Speech Perception of Arabic Emphatics

Authors : Maha Saliba Foster

Abstract : Speech perception has been recognized as a bi-sensory process involving the auditory and visual channels. Compared to the auditory modality, the contribution of the visual signal to speech perception is not very well understood. Studying how the visual modality affects speech recognition can have pedagogical implications in second language learning, as well as clinical application in speech therapy. The current investigation explores the potential effect of speech visual cues on the perception of Arabic emphatics (AEs). The corpus consists of 36 minimal pairs each containing two contrasting consonants, an AE versus a non-emphatic (NE). Movies of four Lebanese speakers were edited to allow perceivers to have partial view of facial regions: lips only, lips-cheeks, lips-chin, lips-cheeks-chin, lips-cheeks-chin-neck. In the absence of any auditory information and relying solely on visual speech, perceivers were above chance at correctly identifying AEs or NEs across vowel contexts; moreover, the models were able to predict the probability of perceivers' accuracy in identifying some of the COIs produced by certain speakers; additionally, results showed an overlap between the measurements selected by the computer and those selected by human perceivers. The lack of significant face effect on the perception of AEs seems to point to the lips, present in all of the videos, as the most important and often sufficient facial feature for emphasis recognition. Future investigations will aim at refining the analyses of visual cues used by perceivers by using Principal Component Analysis and including time evolution of facial feature measurements.

Keywords : Arabic emphatics, machine learning, speech perception, visual speech perception

Conference Title : ICSR2020 : International Conference on Scientific Research and Development

Conference Location : Chicago, United States

Conference Dates : December 12-13, 2020