It's not what you say, it's how you say it: A computational multimodal analysis of how linguistic and vocal features influence consumers' psychological arousal to advertising

Yee Jun Ow[a], Bahareh Amini[b], Ed Hunt[a], Jameson L Hayes[c], Edward C Malthouse[a]

[a]*Medill School of Journalism, Media, Integrated Marketing Communications, Northwestern University, Evanston, IL, USA*

[b]*Department of Advertising and Public Relations, The University of Alabama, Tuscaloosa, USA*

[c]*School of Journalism and Mass Communications, College of Information and Communications, Columbia, SC, USA*

Advertising research has focused primarily on message content, with less attention given to how messages are delivered. This study draws on the Elaboration Likelihood Model and the Stimulus–Organism–Response framework to examine how both linguistic and vocal features jointly influence psychological response during advertisement exposure. Using a dataset of 5,212 English-language television advertisements and electrodermal activity (EDA) collected from 106 at-home viewers, linguistic features were extracted using LIWC and vocal features were derived from iMotions acoustic voice analysis. A gradient boosting model examined the direct, relative, and interactive effects of these features on psychological response. Results show that vocal characteristics—particularly speaking rate, emotional neutrality, intonation, and loudness—exert the strongest influence on psychological arousal while linguistic categories like lifestyle and motivational language also contribute. Importantly, linguistic and vocal features interact, which indicates that persuasive effects depend on how message content and delivery work together.

Keywords: multimodal persuasion, vocal delivery, linguistic features, psychological arousal, electrodermal activity (EDA), voice analysis