MORE THAN AFFECT: HUMAN FACIAL EXPRESSIONS PROVIDE ICONIC AND PRAGMATIC FUNCTIONS

JONAS NÖLLE[*1], YICHEN WU[1], PABLO ARIAS[2,3], OLIVER G. B. GARROD[1], PHILIPPE G. SCHYNS[1] and RACHAEL E. JACK[1]

[*]Corresponding Author: jonas.noelle@glasgow.ac.uk
[1]School of Psychology & Neuroscience, University of Glasgow, UK
[2]Cognitive Science, Lund University, Sweden
[3] Sound Perception and Design, IRCAM, Paris, France

Human facial expressions are a powerful tool for social communication. In addition to the visually salient appearance of the human face (Tomasello et al., 2007; Vick et al., 2007), it has a high capacity to generate many different facial expressions that can each represent different nuanced social messages. Though used daily for a variety of social interactions, facial expressions have mainly been studied as displays of affect in humans (Ekman, 1994) and nonhuman primates (Van Hooff, 1972; Waller & Micheletta, 2013) and less so as a pragmatic tool for communication (cf. Crivelli & Fridlund, 2018). Thus, while multimodal accounts of communication continue to gain attention (e.g., Perniss, 2018; Holler & Levinson, 2019), including in discussions of language origins (Zlatev et al., 2017; Fröhlich et al., 2019), the contribution of facial expressions to multimodal communication remains understudied. One main challenge to this endeavor is the sheer number and complexity of facial expressions (Jack et al., 2018).

To unravel this complexity, we used a powerful data-driven method (Yu et al., 2012) that agnostically generates facial movements and objectively measures their effect on social perception (see Jack & Schyns, 2017 for a review). Here, we examined the semiotic potential of facial expressions combined with speech as part of multimodal signaling, addressing two questions: whether facial expressions can 1) express meanings iconically (e.g., representing size) and/or 2) serve as pragmatic markers, akin to speech prosody, described below.

Experiment 1. We tested whether facial expressions can communicate iconic meanings, similar to manual gestures, e.g., pinching to express ‘tiny’ numbers

referred to in speech (Woodin et al., 2020). We presented participants (20 native English speakers, 10 females) with a scenario (e.g., *"There are 108 animals in the field"*) followed by a series of multimodal stimuli—i.e., speakers each displaying a facial expression and commenting on the scenario one of three vague quantifiers *("Of these, several/many/few were cows"*). Participants viewed each stimulus and estimated the number communicated by the speaker. Facial expressions displayed on each trial comprise combinations of eye/eyebrow movements (called Action Units—AUs; Ekman & Friesen, 1978) that do not affect speech, agnostically generated by a generative model of real human facial movements (Yu et al., 2012). We generated different voices using recordings of two native English speakers, normalized and transformed via pitch and spectral envelope shifts to create naturalistic but tightly controlled speaker identities (Arias et al., 2021). We generated synchronized lip movements using a neural net (Cudeiro et al., 2019) and aligned the onset of each facial expression with the spoken quantifier. Finally, we displayed each facial expression on a randomly generated face identity using a generative model of human face shape/complexion (Zhan et al., 2019). Analysis of the statistical relationship between the facial movements presented on each trial and each participant's responses (Ince et al., 2017) revealed that specific facial movements modulate the quantities estimated by the participants.

Experiment 2. Using the same data-driven approach, we examined the pragmatic function of facial expressions to mark confidence or doubt, also known as 'Feeling of Knowing,' akin to prosodic cues when answering questions (Jiang & Pell, 2017). Here, the scenario comprised a question (e.g., "*Is she a good leader?"*) followed by a multimodal stimulus (i.e., speakers displaying a facial expression) responding 'yes' or 'no.' Participants rated the perceived confidence of the speaker's answer using a 5-point scale ('very doubtful' to 'very confident'). Analysis using ordinal logistic regression revealed that specific facial movements influence the perception of confidence (AU43 eyes closed, AU1-2 brow raiser) and doubt (AU4 brow lowerer, AU7 lid tightener).

In sum, we used a data-driven psychophysical approach combined a with generative model of the human face to precisely characterize the facial movements that serve pragmatic functions in multimodal communication. Our results suggest that facial movements originally evolved to control sensory input (Susskind et al., 2008) can be exapted to ground more abstract social signals or iconic representations, with direct implications for central theories of multimodal communication and language origins (e.g., Holler & Levinson, 2019).

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. [759796]) to REJ; Wellcome Trust [Senior Investigator Award; 107802], Multidisciplinary University Research Initiative/Engineering and Physical Sciences Research Council [172046-01] awarded to PGS; the Swedish Research Council [2014-1371] awarded to PA. We would also like to thank Bodo Winter for discussions about the design, Philip McAleer and Gaby Mahrholz for help with sound recordings, as well as Niamh Mary McLaughlin and Jude Kiltie, who kindly lent us their voices.

References

Arias, P., Rachman, L., Liuni, M., & Aucouturier, J.-J. (2021). Beyond Correlation: Acoustic Transformation Methods for the Experimental Study of Emotional Voice and Speech. *Emotion Review*, *13*(1), 12–24. https://doi.org/10.1177/1754073920934544

Crivelli, C., & Fridlund, A. J. (2018). Facial Displays Are Tools for Social Influence. *Trends in Cognitive Sciences*, *22*(5), 388–399. https://doi.org/10.1016/j.tics.2018.02.006

Cudeiro, D., Bolkart, T., Laidlaw, C., Ranjan, A., & Black, M. J. (2019). *Capture, Learning, and Synthesis of 3D Speaking Styles*. 10101–10111. https://openaccess.thecvf.com/content_CVPR_2019/html/Cudeiro_Capture_Learning_and_Synthesis_of_3D_Speaking_Styles_CVPR_2019_paper.html

Ekman, P. (1994). All emotions are basic. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 65–58). Oxford University Press.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: Investigator's guide*. Consulting Psychologists Press.

Fröhlich, M., Sievers, C., Townsend, S. W., Gruber, T., & Schaik, C. P. van. (2019). Multimodal communication and language origins: Integrating gestures and vocalizations. *Biological Reviews*, *94*(5), 1809–1829. https://doi.org/10.1111/brv.12535

Holler, J., & Levinson, S. C. (2019). Multimodal Language Processing in Human Communication. *Trends in Cognitive Sciences*, *23*(8), 639–652. https://doi.org/10.1016/j.tics.2019.05.006

Ince, R. A. A., Giordano, B. L., Kayser, C., Rousselet, G. A., Gross, J., & Schyns, P. G. (2017). A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula. *Human Brain Mapping*, *38*(3), 1541–1573. https://doi.org/10.1002/hbm.23471

Jack, R. E., Crivelli, C., & Wheatley, T. (2018). Data-Driven Methods to Diversify Knowledge of Human Psychology. *Trends in Cognitive Sciences*, *22*(1), 1–5. https://doi.org/10.1016/j.tics.2017.10.002

Jack, R. E., & Schyns, P. G. (2017). Toward a Social Psychophysics of Face Communication. *Annual Review of Psychology*, *68*(1), 269–297. https://doi.org/10.1146/annurev-psych-010416-044242

Jiang, X., & Pell, M. D. (2017). The sound of confidence and doubt. *Speech Communication*, *88*, 106–126. https://doi.org/10.1016/j.specom.2017.01.011

Perniss, P. (2018). Why We Should Study Multimodal Language. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01109

Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W., & Anderson, A. K. (2008). Expressing fear enhances sensory acquisition. *Nature Neuroscience*, *11*(7), 843–850. https://doi.org/10.1038/nn.2138

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution*, *52*(3), 314–320. https://doi.org/10.1016/j.jhevol.2006.10.001

Van Hooff, J. (1972). A comparative approach to the phylogeny of laughter and smiling. *Non-Verbal Communication*, 209–241.

Vick, S.-J., Waller, B. M., Parr, L. A., Smith Pasqualini, M. C., & Bard, K. A. (2007). A Cross-species Comparison of Facial Morphology and Movement in Humans and Chimpanzees Using the Facial Action Coding System (FACS). *Journal of Nonverbal Behavior*, *31*(1), 1–20. https://doi.org/10.1007/s10919-006-0017-z

Waller, B. M., & Micheletta, J. (2013). Facial Expression in Nonhuman Animals. *Emotion Review*, *5*(1), 54–59. https://doi.org/10.1177/1754073912451503

Woodin, G., Winter, B., Perlman, M., Littlemore, J., & Matlock, T. (2020). 'Tiny numbers' are actually tiny: Evidence from gestures in the TV News Archive. *PLOS ONE*, *15*(11), e0242142. https://doi.org/10.1371/journal.pone.0242142

Yu, H., Garrod, O. G. B., & Schyns, P. G. (2012). Perception-driven facial expression synthesis. *Computers & Graphics*, *36*(3), 152–162. https://doi.org/10.1016/j.cag.2011.12.002

Zhan, J., Garrod, O. G. B., van Rijsbergen, N., & Schyns, P. G. (2019). Modelling face memory reveals task-generalizable representations. *Nature Human Behaviour*, *3*(8), 817–826. https://doi.org/10.1038/s41562-019-0625-3

Zlatev, J., Wacewicz, S., Żywiczyński, P., & van de Weijer, J. (2017). Multimodal-first or pantomime-first?: Communicating events through pantomime with and without vocalization. *Interaction Studies*, *18*(3), 465–488. https://doi.org/10.1075/is.18.3.08zla