



May 29, 2018

Dear Editors of Cognition,

I am hereby submitting for your consideration a manuscript entitled “How optimal is word-referent identification under multimodal uncertainty?” coauthored by Michael C. Frank and myself. The paper studies how adults recognize a spoken word when they have access to uncertain cues from speech and the visual referent. The case of multimodal uncertainty is crucial to advance our knowledge about language understanding in real life---unlike in some simplified laboratory settings, word recognition “in the wild” is a challenging task, and listeners often have to combine noisy cues from various modalities. We focused on the referential context because, first, it is under-studied in the adult literature, and second, because it allows us to relate the findings to some puzzling questions about early word learning (e.g., Stager and Werker, 1997).

We addressed this question using a novel research strategy which compares human performance to an Ideal Observer Model. Crucially, this strategy allowed us not only to determine when humans followed the optimal predictions, but also to understand when and how exactly they deviated from optimality. This research strategy could be used well beyond the current work in order to study similar question in word recognition (e.g., sound symbolism). Moreover, it offers a first step towards a formal framework that accounts for the interaction of speech and visual referents in early word learning.

I note that: (1) the results reported in the manuscript have been in part presented to the Cognitive Science Society as Fourtassi, A. & Frank, M. C. (2017). Word identification under multimodal uncertainty. Proceedings of the 39th Annual Conference of the Cognitive Science Society, but have not been published in any peer reviewed publication, nor are they under consideration for publication elsewhere; (2) both authors have agreed to this submission; and (3) none of the authors have any financial interest or a conflict of interest regarding this work and this submission.

The manuscript in its entirety is 46 pages long, and it includes 1 table, 7 figures, and 1 appendix. The model, sample sizes and analytic plans were preregistered with the Open

Science Framework. All data and analytic code are available on github.com. The links to these repositories were provided in the main text under the appropriate sections.

Sincerely,

Abdellah Fourtassi, Ph.D.
Department of Psychology
Stanford University