

Accepted Manuscript

Audio-visual expressions of attitude: How many different attitudes can perceivers decode?

Hansjörg Mixdorff , Angelika Hönemann , Albert Rilliard , Tan Lee ,
Matthew K.H. Ma

PII: S0167-6393(17)30040-7
DOI: [10.1016/j.specom.2017.08.009](https://doi.org/10.1016/j.specom.2017.08.009)
Reference: SPECOM 2486



To appear in: *Speech Communication*

Received date: 21 January 2017
Revised date: 12 July 2017
Accepted date: 28 August 2017

Please cite this article as: Hansjörg Mixdorff , Angelika Hönemann , Albert Rilliard , Tan Lee ,
Matthew K.H. Ma , Audio-visual expressions of attitude: How many different attitudes can perceivers
decode?, *Speech Communication* (2017), doi: [10.1016/j.specom.2017.08.009](https://doi.org/10.1016/j.specom.2017.08.009)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Audio-visual expressions of attitude: How many different attitudes can perceivers decode?

Hansjörg Mixdorff^{*a}, Angelika Hönemann^{a,b}, Albert Rilliard^c
Tan Lee^d, Matthew K.H. Ma^d

^aBeuth University of Applied Science Berlin, ^bUniversity of Bielefeld, Bielefeld, Germany,
^cLIMSI, CNRS, Université Paris-Saclay, Orsay, France & Federal University of Rio de Janeiro,
CNPq, Brazil, ^dChinese University of Hong Kong, Hong Kong

Received XXXXXX; received in revised form XXXXXXXX; accepted XXXXXX
Available online XXXXXX

Abstract

Based on the paradigm by Rilliard et al. we collected audio-visual expressions of attitudes such as arrogance, irony, sincerity and politeness in German. In the experimental design subjects are immersed in sixteen different communicative situations in which they are supposed to portray a certain attitude in a short dialog. Attitudes can be propositional, that is, reactions to a factual situation and/or social, that is, with respect to the relationship with the collocutor. Furthermore, attitudes can be of positive or negative valence or neutral. Undeniably there is a large repertory of subtle differences in the way certain talkers express certain attitudes. The important question is, however, whether collocutors either from the same language or a different one can actually decode these attitudes reliably. On that account we carried out three perceptual experiments in which we presented our recordings of the portrayed attitudes audio-visually, audio-only and video-only. In the first study, German perceivers rated the expressions given the intended attitude, in the second study, they had to choose the most suitable in a choice of five attitudes, and in the third study raters were able to assign freely the term best matching each attitudinal expression. This last experiment was recently replicated by native speakers of Cantonese in Hong Kong. The current article reviews and reevaluates the results from the first three experiments with the German subjects under the premise that perceivers actually have a more limited set of attitudinal registers which they can reliably draw on. This means that expressions can be sorted into a much smaller number of categories than the projected sixteen. In addition we compare and contrast these resulting clusters with the new data from the Cantonese speaking group. Our results indicate indeed a small number of readily decoded attitudes forming four clusters depending on the experiment design - which are also distinct acoustically. Clusters from the statistical analysis are very similar for the German and the Cantonese perceivers and overlap with basic emotions. This result suggests that expressions of attitudes with low identification rates are more complex to decode and require more pragmatic information, that is, more contextual and possibly idiosyncratic information to be interpreted correctly.

@2017 Elsevier B.V. All rights reserved.

Keywords: social attitudes; multi-modality; speech perception; speech production; expressive speech; prosodic analysis

1. Introduction

Human communication besides and beneath the sheer transfer of linguistic information always has a social goal. Information about the mental state, emotions, mood or attitudes of the speaker and listener is

exchanged during a dialog. The affective state is influenced, for instance, by the situation or role of the dialog partners. In order to aid interactive grounding, comprehension and floor management, speakers use verbal and visual means to convey meta-information about the relevance, novelty and importance of what

* Corresponding author.

Download English Version:

<https://daneshyari.com/en/article/6960890>

Download Persian Version:

<https://daneshyari.com/article/6960890>

[Daneshyari.com](https://daneshyari.com)