How do audible smiles and frowns affect speech comprehension?

Anne van Leeuwen MA, prof. Dr. Hugo Quené and prof. Dr. Jos van Berkum
Utrecht Institute of Linguistics OTS, Utrecht University

We often smile (and frown) while we talk. Listeners to such affective speech have to integrate the affective and the linguistic cues in the speech signal. Following up on earlier work (see Quené et al., 2012), we investigated whether and how affective phonetics (i.e. vocal expressions such as smiling) and affective semantics (sentence-level meaning) interact during spoken language comprehension of sentences and how *perspective* modifies these interactions. We explored this by presenting phonetically and semantically manipulated spoken Dutch sentences to listeners while collecting behavioral and neural (ERP) measurements.

The target materials consisted of utterances that contained a positive or negative content word. Additionally, perspective was taken into account so that sentences were in first person ('ik') or in third person ('hij' or 'ze'). Utterances were phonetically manipulated using Praat's LPC analysis and resynthesis. Between analysis and resynthesis, the formant frequencies were manipulated (upwards or downwards shift: 10%) to imitate the spectral effects of facial expressions while talking (Ohala, 1980; Quené et al., 2012). This resulted in affective congruent realizations (positive – smiling, negative – frowning), or affective incongruent realizations (negative – smiling, positive – frowning). As a control measure we constructed neutral utterances that were as similar as possible to the target materials, but without being affective in nature. These sentences had a low frequent or high frequent content word and they carried neutral articulation. The same LPC analysis and resynthesis were used, but without altering the formants. This resulted in phonetically neutral synthetic realizations (high frequent – neutral, low frequent neutral). See examples below.

Table 1: Example sentences

☺ = smile manipulation
☹ = frown manipulation
😐 = neutral

EXPERIMENTAL DESIGN[1]	**Perspective**	
	1st person	**3rd person**
Congruent	Ik word erg *vrolijk* van hem ☺ Ik word erg *somber* van hem ☹	Hij/ze wordt erg *vrolijk* van hem ☺ Hij/ze wordt erg *somber* van hem ☹
Incongruent	Ik word erg *somber* van hem ☺ Ik word erg *vrolijk* van hem ☹	Hij/ze wordt erg *vrolijk* van hem ☹ Hij/ze wordt erg *somber* van hem ☺
Low frequent	Ik word erg *bedrijvig* hierdoor 😐	Hij/ze wordt erg *bedrijvig* hierdoor 😐
High frequent	Ik word erg *wakker* van koffie 😐	Hij/ze wordt erg *wakker* van koffie 😐

In the EEG study, participants just listened to the utterances while we measured their brain response 200 ms before until 1000 ms after the onset of the critical word. In the behavioral study, two different groups of participants were asked to jugde wether the utterance (truncated at the offset of the critical word) was

[1] Gloss:
- Ik/hij/ze word(t) erg vrolijk/somber van hem
  'I/he/she become(s) very cheerful/sad because of him'
- Ik/hij/ze word(t) erg bedrijvig/wakker …
  'I/he/she become(s) very active/alert …

positive/negative (in terms of meaning), or smiling/frowning (in terms of articulation).

The general predictions were that incongruent sentences were responded to more slowly and eliciting a greater N400 component than congruent sentences. We were especially interested in the effect of perspective on affective sentences: would listening to first person perspective result in a qualitative and/or quantitative different response than the responses to third person perspective? We hypothesized that perspective would modify the effects found in the affective sentences, but not in the control sentences because in the first case first person perspective conveys direct information about the affective state of the speaker, while in the latter case there is no affective information in the speech signal.

Results will be discussed in the light of affective processing and how smiling and frowning influence the way listeners process speech. We will also discuss perspective and how perspective modifies responses to affective stimuli.

References

(1) Ditman, T., Brunye, T. T., Mahoney, C. R., & Taylor, H. A. (2010). Simulating an enactment effect: Pronouns guide action simulation during narrative comprehension. *Cognition, 115*(1), 172-178.
(2) Drahota, A., Costall, A., & Reddy, V. (2008). The vocal communication of different kinds of smile. *Speech Communication, 50*(4), 278-287.
(3) Ohala, J. J. 1980. The acoustic origin of the smile. *J. Acoust. Soc. Am.* 68.S33.
(4) Papeo, L., Corradi-Dell'Acqua, C., & Rumiati, R. I. (2011). "She" is not like "I": The tie between language and action is in our imagination. *Journal of Cognitive Neuroscience, 23*(12), 3939-3948.
(5) Quené, H., Semin, G. R., & Foroni, F. (2012). Audible smiles and frowns affect speech comprehension. *Speech Communication.*
(6) Schroder, M. (2006). Expressing degree of activation in synthetic speech. *Ieee Transactions on Audio Speech and Language Processing, 14*(4), 1128-1136.
(7) Tartter, V. C., & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *Journal of the Acoustical Society of America, 96*(4), 2101-2107.