# A corpus devoted to the cross-perception of mandarin Chinese vs. French social affects

Yan Lu [1] , Véronique Aubergé [1,2]

[1] Gipsa Lab, CNRS, France

[2] Laboratory of Informatics of Grenoble (LIG), CNRS, France

## 1. Problematics

This paper long term aims to study how social affects in mandarin Chinese speech are audio-visually perceived by native French, or native Chinese, either Chinese learners at different steps. The specificity of male vs. female speakers and listeners is specially focused.

The deep link between language and culture is particularly illustrated by the social affects. The social affects expressed in speech can be discriminated from emotions since they can be supposed to be voluntarily controlled by the subjects, learned specifically inside a given culture, contrary to emotions that are involuntarily controlled by the subjects [Aubergé, 2002]. The social affects are the information expressed in speech with audio-visual prosody about:

- attitude, intention or point of view of the speaker about the enunciation , in particular one specific attitude which is not to (want/can) express any attitude – that is often named trivially, surely wrongly, "neutral";
- some characteristics of the social situation during the interaction (in particular related to the social hierarchy that can be, following the culture, identified by politeness or authority);
- some parameters of the context of the socio-cultural interaction/enunciation : typically intimacy, infant directed speech, seduction etc..

Since they are culture dependent, the values of the social affects can vary between languages and cultures, even if it is expected that universals are numerous, and secondly for a same value of social affect, the prosodic morphology can greatly vary between languages.  Since Fonagy (1991), the heavy weight of the attitudes to access the meaning is more and more shown (Aubergé et al 1997, Wichmann, 2002; Campbell, 2005; de Moraes et al, 2010). Specifically in language learning the prosodic "false friends" effect and more generally misunderstanding, have been pointed in pairs of languages (French vs. Japanese vs. English by Shochi et al, 2008; Vietnamese vs. French by Mac et al., 2010).

China and France are geographically, linguistically and culturally far. To improve the French learning methods of Chinese learners, it has to study how are cross-perceived the mandarin Chinese language (one of the languages of China, the reference language commonly used in whole China) and the French language, since it seems that it has not yet be studied for the French learning to Chinese students.

Since they are socio-culture dependent, the values of the social affects can vary inside a same language with the social role, education, age and sex. In this study, it is supposed that the sexual identity in particular, both innate and stereotyped, can influence or structure the attitudes. In the case of Japanese for example, the whole prosody is built and learned by the subjects following the male vs female attribute (Sadanobu, 2004).

## 2. Literacy overview

After Darwin (1872), the research of the emotions and the social affects was amplified by physiologists and psychologists. Their researches were successful and significant, but the research field of emotion and attitude is not limited to the physiology and the psychology, and so many studies have been carried out by didacticians, linguists and phoneticians. Some of them were interested in the pragmatic aspects of the attitudes of a given language for didactic purpose (Callamand, 1987) ; some have conducted cross-cultural comparative studies of the attitudinal prosody of different languages(Mejvaldová, 2000 ; Shochi, 2008) ; some worked on the identification of emotions and attitudes with native and non native subjects (Abelin & Allwood, 2000, 2002 ; Shochi, 2008). Among these researches, we found some studies in which the authors have specially considered the specificity of sex. For example, Shochi et al., (2008), Mac et al (2010) showed each other in their work the difference of the production of such an attitude between a native male speaker and a native female speaker and the different results of the perception of this attitude; Markham (cited by Cornaire, 1998: 78-79) noted in her study that we are conditioned to listen more closely to a male speaker.

## 3. Methodology and results

### 3.1 The corpus structure

This paper presents specifically the design of a large audio-visual corpus aimed to study how French perceive Chinese social affects. That is consequently to study the common case of Chinese learners of French who were not specifically teached to produce the French prosody of social affects, and who would keep Chinese prosody on the French phrases that they produce.

It is thus necessary to measure the cross-perception of a set of phrases which contain all the possible cues possibly implied in the variation of social affects perception.

The function of social affects is implemented in the prosodic parameters together with all the prosodic functions: modalisation, segmentation/hierarchisation (syntactic function), focalisation. That is to only study the social affects function, and it is necessary to freeze the other functions, or on the contrary to vary it in all positions, in order to verify the same effect of a same social affect on all the variations for the other functions.

Consequently, the corpus is built to vary:

- the global length of phrases from 1 to 9 syllables;

- the syntactic structure from mono-word to complex structures, and this is for each global length;

- the length location of the syntactic frontier;

- the value of the tone, as much as possible for each syllable – Chinese has 4 tones, French is a non tonal language

Thus, are obtained 8 mono-words, 20 words of two syllables, 60 words of three syllables, 26 phrases of four syllables and 12 phrases of nine syllables.

That is 152 sentences.

Table 1 show the example of the phrases of 3 syllables

| N° | Ton | Chinois | translation word for word | English | structure |
|---|---|---|---|---|---|
| 29 | 2-1-0 | 抬桌子 | lift(1) * table(2) | lift the table | verb+noun |
| 30 | 2-2-0 | 胡萝卜 | carrot | carrot | noun |
| 31 | 2-3-0 | 量尺寸 | measure(1) size(2) | measure the size | verb+noun |
| 32 | 2-4-0 | 拿钥匙 | take(1) key(2) | take the key | verb+noun |

* 1 syllable in Chinese

### 3.2. The social affects values

After Aubergé et al 1997, Shochi 2008, Mac et al 2011, we retained:

- as « neutral » attitudes, that is as modalities:  declaration, interrogation

- as attitudes:  irritation, neutral surprise, positive surprise, negative surprise, evidence, doubt, sarcastic irony, scorn, seduction, admiration, resignation, confidence, disappointment

- as social parameters : authority, politeness

- as social context cues : infant-directed speech, intimacy

Each of the 152 sentences must be produced for each of the 19 values of social affects, that is 2888 sentences for the corpus.

### 3.2 The corpus recording

Two speakers, one male and one female, have been selected for the corpus recording. The speakers are native Chinese from the same area in China : Shaanxi province. They are language teacher who are used to produce « ecological » utterances; in order that the students could use this prototype in real life during their language practice.

The recording was carried out in sound proof room with High quality sound and video signals.

These data is under perceptive validation by Chinese native listeners, balanced with male vs. female subjects.  The following experiment will carry out the perception by French listeners, one native group in Chinese, other groups of different levels in Chinese language learning.

## 4. References

Abelin, A., & Allwood, J. (2000). Cross linguistic interpretation of emotional prosody. ISCA workshop on Speech and Emotion, Newcastle, Northern Ireland, 110–113.
Abelin, A., & Allwood, J. (2002). Cross linguistic interpretation of emotional prosody. Gothenburg papers in theoretical linguistics, Göteborg, Sweden.Callamand, D. (1987),

*Grammaire vivante du français : français langue étrangère*, Paris, Larousse, Diffusion CLE International.

Aubergé V. (2002). A Gestalt Morphology of Prosody Directed by Functions, Proc. of Speech Prosody 2002, Aix en Provence, France, 151-154.

Aubergé, V., Grépillat, T., Rilliard, A. 1997. Can we perceive attitudes before the end of sentences? The gating paradigm for prosodic contours. 5th Eurospeech, 871-874.

Campbell, N. (2004). Perception of Affect in Speech –towards an Automatic Processing of Paralinguistic Information in Pro. 8th ICSLP, Jeju, Korea, 881-884.

Cornaire C., (1998), *La compréhension orale*, Paris, CLE International, pp 78-79.

Fonagy, Y. (1991). La Vive Voix, Paris, Payot

Mac DK , Aubergé V , Rilliard A. Castelli E (2010), How prosodic attitudes can be recognized and confused: Vietnamese multimodal social affects , SLTU

Mejvaldova, J. (2000). Expressions prosodiques de certaines attitudes en thèque et en français: étude comparative, Paris, Université Paris 7 – Denis Diderot.

Moraes J.A., Rilliard A., Alberto B. & Shochi T. (2010). Production and perception of attitudinal meaning in Brazilian Portuguese. Proceedings of Speech Prosody 2010, Chicago, Illinois.

Sadanobu, T. (2004) A natural history of Japanese pressed voice. J. of the Phonetic Society of Japan 8 (1) 29-44.

Shochi, T. (2008), Prosodie des affects socioculturels en japonais, français et anglais : à la recherche des vrais et faux-amis pour le parcours de l'apprenant, PhD thesis, Grenoble University.

Wichmann, A. (2002) "Looking for attitudes in corpora",Language and Computers, From the COLT's mouth ... and others'. Edited by Leiv Egil Breivik and Angela Hasselgren , pp. 247-261(15)