

How prosodic attitudes can be false friends: Japanese vs. French social affects

Takaaki Shochi, Véronique Aubergé, Albert Rilliard

Institut de la Communication Parlée, UMR CNRS 5009, Grenoble, France

{Takaaki.Shochi, Veronique.Auberge, Albert.Rilliard}@icp.inpg.fr

Abstract

The attitudes of the speaker during a verbal interaction are affects linked to the speaker intentions, and are built by the language and the culture. They are a very large part of the affects expressed during an interaction, voluntary controlled. This paper describes several experiments which show that some attitudes own both to Japanese and French and are implemented in perceptively similar prosody, but that some Japanese attitudes don't exist and/or are wrongly decoded by French listeners. Results are presented for 12 attitudes and three levels of language (naive, beginner, intermediary). It must particularly be noted that French listeners, naive in Japanese, can very well recognize admiration, authority and irritation; that they don't discriminate Japanese question and declaration before the intermediary level, and that the extreme Japanese politeness is interpreted as impoliteness by French listeners, even when they can speak a good level of Japanese.

1. Introduction

The affects in speech are expressed following different cognitive processing levels, from involuntary controlled expressions to the intentionally, voluntary, control of the attitudes of the speaker. Attitudes are sometimes assimilated to emotions since both use to be expressed in the direct acoustic channel with prosodic encoding. But if the emotion expressions are carried by voice in parallel to the speech structures, the attitudes expressions are part of the language interaction building. Some affects values, like the surprise, are discussed to be classified either as attitudes vs. nor as emotions. Our position is to assume that the surprise can be an emotion if it is involuntary processed, and can be an attitude when it is voluntary processed, and then it needs to be learned. When a speaker does not produce any attitude on an utterance, it must be consider as a special attitude that is that this speaker has the intention not to express any information about his intentions, that is the "no attitude" attitude. It is why we consider that studying affects in the speech interaction is crucial for the expressive speech modelling, and needs to describe precisely what can be the attitudes of a speaker in a language and a culture and how it is morphologically encoded. The cross-cultural study is an helpful method to enter this problem.

Like for all the language specifications, the attitudes can be expected to be universal for some of them: universal values and for a part universal prosodic morphology (like authority, surprise). Because attitudes are constructed socially for and by the language, they can exist or not from one language to the other, and prosodic realization of one specific attitude in a specific language may not be recognized or may be ambiguous in the learner's language.

After presenting the Japanese corpus on which is based this paper, we summarize different perceptive experiments. The first is the validation of the Japanese attitudes by native Japanese listeners. Then we present (1) how French listeners naive – level 0 – in Japanese do interpret these Japanese attitudes (2) what French listeners, beginners in Japanese – level 1 – did learn from Japanese attitudes (3) what French listeners, quite fluent in Japanese – level 2 – cannot still identify from Japanese and which kind of ambiguities and confusions are implied by such acquisition.

2. Selection of 12 Japanese Attitudes

We have selected a set of 12 attitudes for Japanese, supposed to be representative from the literacy, especially from the Japanese teaching methods: doubt-incredulity, evidence, exclamation of surprise, authority, irritation, arrogance-impoliteness, sincerity-politeness, admiration, *kyoshuku*, simple-politeness, declaration and interrogation. Several of these attitudes are specific or specifically important to the Japanese culture, especially those linked to the Japanese politeness strategy; "simple politeness", sincerity-politeness and *kyoshuku* vs. arrogance-impoliteness. The attitude of "sincerity-politeness" appears when speaker is inferior facing to his interlocutor who is superior in Japanese society. The speaker expresses by this prosodic attitude that his intention is serious and sincere. The attitude of "*kyoshuku*" (no lexical entry to translate it into English) is a typical Japanese cultural attitude: even if such situations can happen in all cultures, the Japanese language has chosen to specially encode this situation as an "attitudineme". It appears when the speaker wants to express his contradictory opinion on a situation in which his social states is inferior to his interlocutor, not in the aim to disturb his superior but to help him, or when the speaker is willing to get a favour from his superior. It is described by T Sadanobu ([11] p. 34) as "*a mixture of suffering ashamedness and embarrassment, comes from the speaker's consciousness of the fact that his/her utterance of request imposes a burden to the hearer*".

3. The corpus

Our aim being to measure the perceptive behavior of French listeners, we need to get some references data on Japanese attitudes. These utterances must be free of lexico-syntactic information about attitudes (only prosodic attitudinal information are expected). The other prosodic variations must be balanced, in order to control and measure the possible bias of the Japanese intonation or lexical stress, that could be interpreted by French listeners like some attitudes cues. The first step is thus the recording of a very controlled corpus, constructed according to a few structure principles. The

construction of opposed minimal pairs allows observing the effect of the targeted factor only. On the basis of such controlled corpora, an acoustic analysis leads to a statistical model of prosodic variations, which can be used in order to synthesize the captured prosodic variations by analysis-resynthesis methods.

In order to validate this work, a compared analysis of the competences of natural prosody contained in the corpus vs. the performances of the synthetic prosody extracted by analysis will be realized in order to define a description of the observed prosody.

The corpus is based on seven sentences ranging from one to eight moras. The syntactic structure of the sentences used in the corpus is either a single word, or a simple verbal object structure. For the eight-moras utterances, the position of lexical accent may be on the first, second, and third mora or absent. In order to express some attitudes like “doubt” or “surprise”, the vowel [u] may be inserted at phrase final position, and lexical accent will be realized at the seventh mora too. The sentences were constructed in order to have no particular connotations in any region of Japan. Each sentence is produced with all the attitudinal functions.

One male Japanese native language teacher produced recorded all the attitudes for this corpus. We already mentioned that the attitudes are socially constructed for each language. Therefore foreign language learners have to learn prosodic attitudes corresponding to the target language and culture, and languages teachers are able to elicit different types of attitudes for didactic/pragmatic reasons.

The complete corpus contains 84 stimuli, i.e. seven utterances produced with twelve different attitudes. All stimuli are used for the perception test.

Table 1. Corpus of Japanese attitudes: 7 utterances of different length with different position of the lexical accent, which is marked with a star, each produced with the 12 different attitudes.

Nb mora	Utterance	Translation
1	Me	The eye
2	Na*ra	Nara
5 (3+2)	Na*rade neru	He sleeps in Nara
8 (4+4)	Na*goyade nomimas	He drinks in Nagoya
8 (4+4)	Nara*shide nomimas	He drinks in Nara Town
8 (4+4)	Matsuri*de nomimas	He drinks at the party
8 (4+4)	Naniwade nomimas	He drinks at Naniwa

4. Experimental protocol

The first step is the validation of the corpus attitudes by Japanese native listeners. 15 Japanese listeners (11 females and 4 males) who speak the Tokyo dialect, whom mean age is 29.5 had to choose attitudes in a closed choice of the 12 attitudes. On a Japanese interface, each label is largely described and illustrated by examples of situations in which such an attitude can happen.

French has been chosen to be culturally distant from the Japanese culture and language.

- A first experiment was held with 15 French listeners (10 females and 5 males) naïve in Japanese, that is who have never heard the Japanese language. They are the “level 0” group. The mean age is 25.4. Listeners don’t report to suffer of any listening disorder. The French interface proposes a French translation of each label, largely introduced by a long

explanation about each attitude, giving many examples of relevant situations. No subject expresses any trouble to “understand” the attitude label.

- A second experiment was held with 16 French listeners, learning the Japanese language, and evaluated to be in the same coherent level (labeled level 1 in our Japanese cursus at Grenoble University), that is a beginner level: they can speak and understand Japanese, but are still not fluent. They use the same test interface.

- A third experiment was held with 15 French listeners, evaluated to be in the same level 2, that is that they are quite fluent in Japanese. They use the same test interface.

All the subjects of these experiments listens to each stimulus one time only. For each stimulus, they were asked to answer the perceived attitude amongst the twelve possible attitudes. The presentation order of the stimuli was randomized in a different order for each subject.

5. Results

5.1. Validation with Japanese Listeners

According to a chi-square test, all attitudes were above the chance level. Then, we tested a possible effect of the stimuli length for the distributions of selected attitudes. The results show a significant difference of the length between two and five-moras sentences and also between the five and eight-moras one. But, there is no effect of the lexical accent.

In order to determine which attitude listeners recognized over chance, a criterion was used: the mean identification rate must be over twice the theoretical chance level.





According to this criterion, seven attitudes (i.e. “*arrogance-impoliteness*”, “*declaration*”, “*doubt-incredulity*”, “*simple-politeness*”, “*exclamation of surprise*”, “*irritation*” and “*interrogation*”) have been recognized without any particular confusion.

“*Authority*” was confused with “*evidence*”. One possible explanation is that, concerning the self-confidence of speaker, these two attitudes are, in fact, similar: when imposing authority, the speaker is certainly sure of himself.

“*Evidence*” was confused with “*arrogance-impoliteness*”. “*Evidence*” shows that speaker is confident of himself, and this expression of certainty can sometimes be perceived like disrespect to interlocutor.

Curiously, two typical Japanese attitudes like “*sincerity-politeness*” and “*kyoshuku*” were confused with each other. These two attitudes express essentially the humility of a speaker facing a superior person in the social hierarchy. It is important to note that “*sincerity-politeness*” was also confused with “*simple-politeness*”, whereas this confusion with “*simple-politeness*” was absent for “*kyoshuku*”.

Concerning the attitude of “*admiration*”, we observed confusion with “*simple-politeness*”. These two attitudes are interconnected each other in the Japanese society. This evidence can be explained by the lexical polysemy of items like “*sonkee*” [admiration / politeness], and “*keifuku*” [admiration / politeness]

Table 2. Confusion matrix in percentages for 15 Japanese listeners: well recognized  ; recognized well with confusion  ; significant recognition, but confused with other attitudes  ; significant confusion (over 16.6%) 

		Percentage											
		Presented attitudes											
Japanese	Recognized attitudes	AD	AR	AU	DC	DO	EV	EX	IR	KYO	PO	QS	SIN
	AD	21.9%	0.0%	0.0%	0.0%	1.9%	0.0%	1.0%	0.0%	1.9%	4.8%	0.0%	3.8%
	AR	1.9%	72.4%	5.7%	10.5%	5.7%	21.0%	0.0%	2.9%	1.9%	0.0%	1.0%	0.0%
	AU	0.0%	9.5%	51.4%	2.9%	1.0%	9.5%	1.0%	11.4%	15.2%	0.0%	1.9%	1.0%
	DC	4.8%	5.7%	10.5%	65.7%	0.0%	12.4%	0.0%	0.0%	2.9%	8.6%	1.9%	9.5%
	DO	0.0%	0.0%	0.0%	0.0%	56.2%	0.0%	14.3%	0.0%	0.0%	1.0%	3.8%	0.0%
	EV	8.6%	5.7%	17.1%	7.6%	0.0%	45.7%	5.7%	0.0%	10.5%	1.9%	1.0%	3.8%
	EX	9.5%	0.0%	0.0%	0.0%	14.3%	4.8%	59.0%	0.0%	1.0%	1.9%	1.0%	1.9%
	IR	1.0%	5.7%	4.8%	1.0%	13.3%	2.9%	4.8%	85.7%	12.4%	0.0%	1.0%	1.0%
	KYO	11.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	24.8%	9.5%	1.0%	27.6%
	PO	26.7%	1.0%	4.8%	11.4%	0.0%	1.0%	0.0%	0.0%	2.9%	64.8%	8.6%	17.1%
	QS	0.0%	0.0%	0.0%	0.0%	7.6%	2.9%	14.3%	0.0%	0.0%	1.0%	77.1%	1.9%
	SIN	14.3%	0.0%	5.7%	1.0%	0.0%	0.0%	0.0%	0.0%	26.7%	6.7%	1.9%	32.4%

5.2. Behavior of the French Listeners - level 0

The distribution of all attitudes was above of chance. A significant effect of the length was identified between the one and two-moras sentences. It was not possible to identify any significant effect of the lexical accent for French subjects.

By using the same criterion that for the Japanese listeners, the following results were extracted. Figure 1. shows that "authority", "irritation" and "admiration" have been perceived with no significant confusion according to our criteria. But, attitude of "arrogance-impoliteness" showed week identification score by French listeners. This attitude was confused with "declaration" and "authority".

French listeners did not recognized two particular attitudes of politeness connected in Japanese society like "sincerity-politeness" and "kyoshuku". Moreover, "Sincerity-politeness" was confused with "simple-politeness" and "kyoshuku" which are others expressions of politeness. On the contrary, attitude of "kyoshuku" was recognized like "irritation", "arrogance-impoliteness" and "authority". This result was expected, since these attitudes do not exist at all in the French society, nor such voice quality does not match any politeness expression in French. French listeners confused also "interrogation" with "declaration". This result shows a possibility for French people to perceive Japanese Yes-No question like simple declaration. They show significant reciprocal confusions between "declaration" and "evidence", between "doubt-incredulity" and "exclamation of surprise", and between "simple-politeness" and "sincerity-politeness".

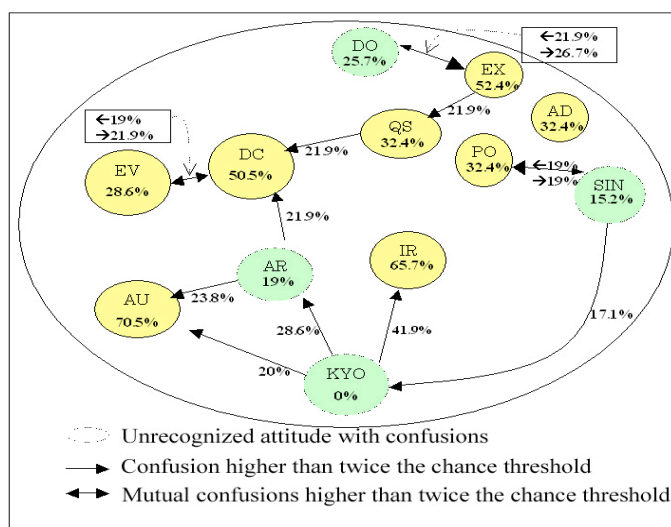


Figure 1. Confusion graph for French listeners level 0: percentages inside the frame indicate the confusion rate in percentages. Other

percentages written under the labels of each attitude represent the identification rates of attitude in percentages.

5.3. Behavior of the French Listeners - level 1

At this level of Japanese (beginners), the subjects have learned to identify the sincerity and the doubt. They have changed some confusions and misinterpret sincerity-serious with declaration and arrogance with evidence. They are some mutual confusions between doubt and surprise. They have learned to learned discriminate arrogance vs. authority, politeness vs. sincerity-serious, declaration -> evidence.

But however, are still confused Kyoshuku with (irritation, arrogance-impoliteness, authority), arrogance with declaration and mainly interrogation with declaration, that can be a strong communicative handicap for such subjects starting to interact in Japanese.

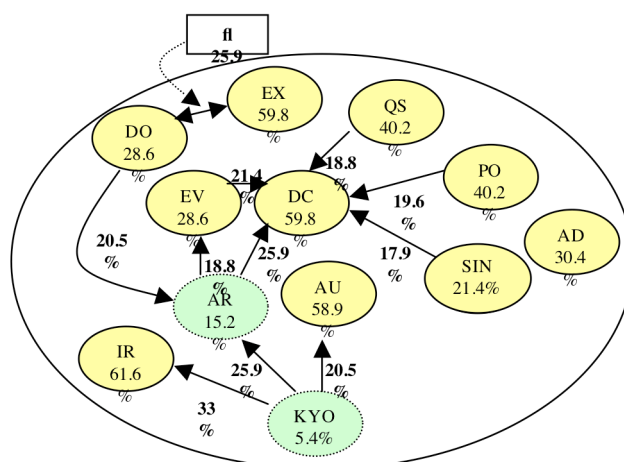


Figure 2. Confusion graph for French listeners level 1: percentages inside the frame indicate the confusion rate in percentages. Other percentages written under the labels of each attitude represent the identification rates of attitude in percentages. Only Arrogance and Kyoshuku are not significantly recognized.

5.4. Behavior of the French Listeners - level 2

At this level, where the French subjects start to be fluent in Japanese, they did not learn to identify any other attitudes. Arrogance and Kyoshuku are still not recognized. They however learned at least to discriminate question and declaration, but Kyoshuku is still confused with (irritation, arrogance-impoliteness, authority), and arrogance with (declaration,evidence). They have changed confusion of declaration with sincerity-serious. Doubt and surprise are still mutually confused (which is not the case for Japanese listeners, and not the case for French in French doubt and surprise [1])

