Voice features of telephone operators predict auditory preferences of consumers

Vanessa André, Christine Petr, Nicolas André, Martine Hausberger and Alban Lemasson
Université de Rennes 1, France

What makes a human voice agreeable is a matter of scientific discussion. Whereas prosody was shown to play a role regarding “male-female” attraction, the impact of frequency modulations in “non-sexual”, notably commercial, contexts has attracted little attention. Another point unaddressed in the literature is auditory sensitivity to short-term frequency modulations as current studies focus more on sentence. Thirty French female operators were recorded over the phone. All “bonjour” greeting words were classified in terms of frequency modulation linearity and orientation at the syllable and word levels. Then, the different voices were played back to students and seniors who had to rate each voice according to their degree of agreeableness. Listeners preferred non-monotonous voices. Differences between age-classes were greater than between sex-classes. Results suggest that short-term frequency changes are important for auditory evaluation of voice agreeableness. This study opens new research perspectives concerning the importance of prosody during consumer-seller interactions.

Keywords: acoustic structure, human, phone interaction, prosody, voice perception

1. Introduction

Regardless of the syntactic and the semantic content of speech, the phonetic organization of voices is commonly used by conversing interlocutors to assess each other’s identity, personality, arousal state and motivation (Bruckert, Lienard, Lacroix, Kreutzer and Leboucher, 2006; Collins, 2000; Feinberg, Jones, Little, Burt and Perrett, 2005; Jones, Feinberg, DeBruine, Little and Vukovic, 2010; Scherer, 1972; Scherer, 1978; Smith, Brown, Strong and Rencher, 1975). Thus, the interpretation of a word by a receiver is based first on facial expressions (55%), second
on voice features (38%), and only third on lexical content (7%) (Mehrabian and Ferris, 1967), suggesting that in the absence of any visual clues, prosody plays a role in auditory perception. However, most studies have focused on women – men auditory sexual attraction and evaluation (Bruckert et al., 2006; Collins, 2000; Jones, Feinberg, DeBruine, Little and Vukovic, 2008; Jones et al., 2010; Klofstad, Anderson and Peters, 2012; Re, O’Connor, Bennett and Feinberg, 2012; Simmons, Peters and Rhodes, 2011). Interestingly, a few studies also investigated vocal attractiveness in other social contexts (e.g. Bruckert et al, 2010). Nevertheless, little is still known about voice agreeableness and its impact on social interactions and notably in contexts where voices play a crucial role as in commercial interactions over the phone.

A key acoustic feature playing a general role in auditory perception and evaluation is the pitch of a voice. Auditory recognition of gender and age is frequency-dependent; the voices of men and seniors are lower-pitched than those, respectively, of women and juniors (Bruckert et al., 2006; Latinus and Belin, 2011). Regardless of the age and the gender of both receivers and emitters, people with low-pitched voices are perceived as more dominant than people with high-pitched voices (Jones et al., 2010; Klofstad et al., 2012). Conversely, men judge women with high-pitched voices to be more feminine than women with low-pitched voices (Collins and Missing, 2003; Jones et al., 2008). Similarly, men with low-pitched voices are considered to be more masculine, more corpulent (larger and taller) and more attractive to women than men with high-pitched voices (Collins, 2000; Feinberg et al., 2005; Jones et al., 2010). Higher pitches are associated with anger, joy and anxiety, whereas lower pitches are associated with sadness and indifference (Zetterholm, 1998).

Voice variations in both temporal and frequency domains appear even more crucial for evaluation than pitch (Besson, Magne and Schon, 2002; Latinus and Belin, 2011). Some people with clinical disorders, such as schizophrenia or depression, are unable to detect these so-called prosodic variations in their interlocutors’ voices and thus have difficulties holding a proper conversation (Alpert, Pouget and Silva, 2001; Bach, Buxtorf, Grandjean and Strik, 2009; Péron et al., 2011). Among the different prosodic features, voice rhythm is of primal importance. For example, reasonably fast-speaking people are considered more competent and more persuasive than slow-speaking people in occidental cultures (Peng, Zebrowitz and Lee, 1993). However, speakers with very low or very high rhythms are considered less benevolent than speakers with intermediate rhythms (Brown, Strong and Rencher, 1973; Brown, Strong and Rencher, 1974; Brown, William and Alvin, 1975; Smith et al., 1975). Frequency variations can also play a crucial role. In general, less monotonous voices (i.e. with frequent frequency changes for instance while reading a speech) are associated with more positive personalities (Zukerman and Miyake,
1993). However, most of the past studies focus on rhythm variations or frequency changes at the sentence or the word level (Brown et al., 1973; Brown et al., 1974; Brown et al., 1975; Peng et al., 1993; McAleer et al., 2014; Smith et al., 1975), and comparatively little is known concerning the impact of subtle frequency variations such as sudden within-word changes (at the syllable level). Subtle frequency variations concern both frequency linearity (linear vs non-linear frequencies within the syllable) and orientation (upward vs flat vs downward frequencies across the syllable), but again, to our knowledge, nothing is known concerning the relative importance of frequency modulations for evaluation by listeners.

Unfamiliar interlocutors are frequently engaged in oral non-visual conversation, notably over the phone. A lot of companies use phone platforms to approach potential clients, so there is no visual signaling during conversations. Direct oral exchanges are rated more positively than electronically-written messages (Dillma et al., 2009), probably because they allow a better evaluation of the interlocutor. Reports suggest that prosody plays a role during phone interviews. Questionnaire survey results indicate that people say that they would agree more easily to answer questions over the phone when the caller’s voice was not monotonous in terms of sentence intonation (Oksenberg, Coleman and Cannell, 1986; Benkí, Broome, Conrad, Groves and Kreuter, 2011).

The current study evaluated the impact of subtle frequency changes (i.e. variation of frequency linearity and orientation) in phone operators’ voices (at the syllable and word levels) on their agreeableness rating by potential consumers. First, we phoned different grocery stores in order to record the voices of various female telephone operators and then classified the recorded voices in terms of frequency modulation linearity and orientation. Second, the different voices were played back to a panel of junior (students) and senior men and women for evaluation. Third, we crossed acoustic and agreeableness data. The impacts of the sex and the age of the listener were assessed. The sex of the listener is supposed to play a role in evaluation ratings (Collins, 2000; Feinberg et al., 2005; Jones et al., 2010), but it is not known whether this is also true in a commercial context. Although reports show that age impacts differently adult and children listeners’ agreeableness evaluations (Saxton, DeBruine, Jones, Little and Roberts, 2009), we do not know whether age impacts adults’ evaluations.
2. Methods

2.1 Participants

The group of participants (listeners) included 30 biology and psychology students at Rennes universities (France), between 18 and 26 years old (13 men, 17 women), and 30 retired persons from various socio-professional categories, between 60 and 75 years old (12 men, 18 women). There was no age difference between males and females in each of the two age groups (Mann Whitney tests; students: $U = 92.500$, $Z = -0.756$, $P = 0.457$ / retired persons: $U = 88.000$, $Z = 0.829$, $P = 0.415$). All participants were French native speakers, living in Rennes city, and were naive to playback experiments.

2.2 Protocol

2.2.1 Voice recording of telephone operators

The voices of 30 women operators, working in grocery stores in Rennes (France) were recorded during a phone conversation in February 2011. One experimenter (N. A.) phoned the reception of the store and inquired about their opening schedule. The conversation was recorded directly on a PC computer (Dell ® Latitude D600) using Audacity® (sampling rate 11kHz, resolution 16bit, .wav format). From these conversations, only the first greeting word pronounced by the operator was saved. As the study focused on the pronunciation of a single word with post-recording anonymized files, speaker identification was not possible. Hence, no approval was necessary. Hence, we collected a data set of 30 “Bonjour” ([bɔʒuʁ], meaning ‘Good morning’ in French) that was used for subsequent acoustic and playback analyses. The 30 recordings were homogenized in intensity using ANA® software (Richard, 1991) so that the acoustic stimuli were comparable.

2.2.2 Playback to participants and evaluation of the agreeableness of the voices

To avoid any bias in our interpretation, we first made sure that all the seniors passed a cognitive test (MMSE “mini mental state examination” - Fostein, Folstein and McHugh, 1975) and a test for geriatric depression (Yesavage et al., 1983). All the seniors successfully passed the screening tests. The Mini Mental State Examination (MMSE) is an 11-question measure. The maximum score is 30. A score of 23 or lower is indicative of cognitive impairment (Kurlowicz and Wallace, 1999). The students were exposed to playbacks in a quiet room of our laboratory and the seniors were tested at home. All raters declared to have no hearing problem. Each participant was informed in oral and written forms that this study aimed to determine the agreeableness of a person according to its vocal features.
Participants were also informed: that they were going to listen to several recordings of telephone operator voices pronouncing a particular word; that each voice will be broadcast only once; that they had to rate each voice by responding to the question “Did you find this voice agreeable?” and by using the following rating scale: 1 (not agreeable at all), 2 (No), 3 (neither-yes-nor-no), 4 (Yes), 5 (Yes very agreeable); and that there was no time restriction to answer. During the test, all the participants were in front of the researcher in the isolated experimental room. All participants completed all trials. We choose a five choice scale to make it comparable with what was used in comparative studies about evaluation of human feelings (Johnson, 1996; Kokkinos, 2007; Nagy, 2002; Trout, Magnusson and Hedges, 2000). The 30 voice stimuli were proposed to each participant in a different (random) order. Sounds were played back using a Hewlett Packard® house dv9000 computer connected to same Sennheiser® HD 25–1 II noise-cancelling headphones. In order to homogenize the amplitude of the different recorded voices, we used the function “Normalizer” of Audacity software to homogenize the amplitude of the different recorded voices. Participants informed the experimenter when they were ready to listen to the first voice, and had to rate the voice just heard before to ask for listening to the next one. The experimenter controlled the playbacks.

2.2.3 Acoustic analyses
The complexity of the frequency modulation (FM) of each voice stimulus was qualified using an acoustic classification, based on visual and audio analysis of spectrograms (Sampling frequency 11kHz, FFT-length 1024), commonly used in animal bioacoustics (Datta and Sturtivant, 2002; Hausberger, 1997; Lemasson and Hausberger, 2011; McCowan and Reiss, 1997): (1) FM linearity (linear vs non-linear), (2) FM orientation (upwards vs flat vs downwards) (Fig. 1). Hence, both syllables (‘bon’ vs ‘jour’) were blindly (i.e. the experimenter did not know the identity of the speaker or the score given to his/her voice) and independently scored in

![Figure 1. Frequency modulation patterns used to classify syllables](image-url)
terms of linearity and orientation by a second experimenter (V. A.) (See examples in Fig. 2). In addition to this latter experimenter (A), two other persons (B and C, both were naive with the experiment and the second one was also naive with bioacoustics) were asked to rate all sonograms according to the above-defined criteria of classification. Then, Cohen’s kappa was calculated to measure the agreement between raters (A-B: 0.97 / A-C: 0.88). This confirms the reproducibility of our rating as Cohen’s kappa greater than 0.85 is typically considered to be high (Cicchetti and Heavens, 1981).

a. non-linear downwards

b. non-linear upwards
2.2.4 Data crossing and statistical analyses

From the 1800 rates collected (i.e. 60 participants rated 30 voices each), we collected the following total of scores: 66 times “1”, 252 times “2”, 776 times “4” and 154 times “5”.

A first Binomial (agreeable vs disagreeable: when subjects gave a score equal to respectively ‘4 or 5’ vs ‘1 or 2’) Generalized Linear Model (GLM) compared the agreeableness scores of participants from the different age-sex classes in terms of FM linearity and orientation patterns for each syllable separately.

Another Binomial (agreeable vs disagreeable) GLM evaluated the relative importance of FM orientation and linearity on the agreeableness scores given by participants from the different age-sex classes at the word level. For instance, a word was considered flat when both syllables were flat. Words were classified as follows (with L = linear, NL = non-linear, F = flat, NF = non-flat for each syllable): non-linear and non-flat (NLLN FNF / LNL NFF / LNL FNF / NLNL FNF), linear and non-flat (LL NFF / LNL FFF), non-linear and flat (NLNL FF / LNL FF), linear and flat (LL FF). Analyses were run with R software with FDR (False Discovery Rate) correction for multiple comparisons.
3. Results

3.1 Analyses at the syllable level

The agreeableness scores given by subjects showed that 69.70% and 30.30% of the voices heard were considered agreeable and disagreeable respectively. However, appreciations varied according to the listeners’ characteristics as well as to the acoustic pattern of frequency modulations (FM). While the listener’s sex did not influence voice appreciation scores (Table 1), age appeared a major factor as the seniors found the voices in general more agreeable than did the students (Fig. 3).

Table 1. Impact of listener’s characteristics and syllable acoustic structures on agreeableness scores.

<table>
<thead>
<tr>
<th></th>
<th>Single factor effects</th>
<th>Interactions with Sex</th>
<th>Interactions with Age</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td>$F_{1/58} = 3.432,$</td>
<td>$F_{1/58} = 0.156,$</td>
<td>$F_{1/58} = 0.256,$</td>
</tr>
<tr>
<td></td>
<td>$P = 0.064$</td>
<td>$P = 0.0693$</td>
<td>$P = 0.613$</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>$F_{1/58} = 9.803,$</td>
<td>$F_{1/58} = 0.346,$</td>
<td>$F_{1/58} = 0.008,$</td>
</tr>
<tr>
<td></td>
<td>$P = 0.002$</td>
<td>$P = 0.556$</td>
<td>$P = 0.927$</td>
</tr>
<tr>
<td><strong>Linearity (Syllable 1)</strong></td>
<td>$F_{1/58} = 6.932,$</td>
<td>$F_{1/58} = 0.346,$</td>
<td>$F_{1/58} = 0.256,$</td>
</tr>
<tr>
<td></td>
<td>$P = 0.008$</td>
<td>$P = 0.556$</td>
<td>$P = 0.613$</td>
</tr>
<tr>
<td><strong>Linearity (Syllable 2)</strong></td>
<td>$F_{1/58} = 2.423,$</td>
<td>$F_{1/58} = 0.002,$</td>
<td>$F_{1/58} = 0.008,$</td>
</tr>
<tr>
<td></td>
<td>$P = 0.120$</td>
<td>$P = 0.965$</td>
<td>$P = 0.927$</td>
</tr>
<tr>
<td><strong>Orientation (Syllable 1)</strong></td>
<td>$F_{2/57} = 32.385,$</td>
<td>$F_{2/57} = 2.893,$</td>
<td>$F_{2/57} = 9.385,$</td>
</tr>
<tr>
<td></td>
<td>$P &lt; 0.001$</td>
<td>$P = 0.235$</td>
<td>$P = 0.009$</td>
</tr>
<tr>
<td><strong>Orientation (Syllable 2)</strong></td>
<td>$F_{2/57} = 15.794,$</td>
<td>$F_{2/57} = 9.293,$</td>
<td>$F_{2/57} = 10.447,$</td>
</tr>
<tr>
<td></td>
<td>$P &lt; 0.001$</td>
<td>$P = 0.010$</td>
<td>$P = 0.005$</td>
</tr>
</tbody>
</table>

1. A General Binomial (GLM) analysed single factor effects (Sex, Age, FM linearity and orientation of each syllable) and interactions between factors ($f_2 = 0.388; Pwr = 0.899$). Significant differences are shown in bold type.

Analyses of voice acoustic structures showed that both linearity and orientation of frequency modulations in the two syllables “bon” (1) and “jour” (2) were determinant (Table 1). Non-linear FM was preferred to linear FM, but only for syllable 1 (Fig. 4). Agreeableness scores regarding FM orientation patterns differed between the two syllables. Syllables 1 with an upward orientation were allocated the highest scores, whereas syllables 2 with a downward orientation were preferred (Fig. 5). Among non-flat voices, 81/90% was going upward and 19/10% was going downward on the first/second syllable. Significant interactions among factors were limited to sex and age with FM orientations (Table 1). Women found upward FM in syllable 2 more agreeable than men did (Fig. 6). The students found downward
syllable 1 and upward syllable 2 less agreeable than did the seniors. Flat syllables 1 and 2 were also found less agreeable by younger listeners (Fig. 7).

Figure 3. Voice agreeableness scores given by the students and the seniors (1: agreeable, 0: disagreeable). Binomial Glm test ** $P < 0.01$.

Figure 4. Agreeableness scores in relation to frequency modulation linearity of syllable 1. Binomial (1: agreeable, 0: disagreeable) GLM ** : $P < 0.01$. 

All rights reserved
Figure 5. Agreeableness scores in relation to frequency modulation orientation pattern of syllables 1 (a) and 2 (b). Binomial GLM NS : $P > 0.05$ ; * : $P < 0.05$ ; *** : $P < 0.001$.

Figure 6. Agreeableness scores given by the women vs the men according to FM orientation. Binomial GLM NS : $P > 0.05$ * $P < 0.05$. 
Voice features of telephone operators predict auditory preferences of consumers

Agreeableness score (mean +/- s.e.)

Students
Seniors

Downward Flat
NS
Upward

** **

a.

Figure 7. Agreeableness scores given by the students vs the seniors according to FM orientation in syllables 1 (a) and 2 (b). Binomial GLM NS : $P > 0.05$ ; * : $P < 0.05$ ; *** : $P < 0.001$.

3.2 Analyses at the word level

None of the voices heard by subjects were linear and flat, 10% were flat and non-linear, 13.33% were linear and non-flat and 76.67% were non-linear and non-flat. Focusing the analysis on the word level showed that variations in evaluation appeared mostly based on orientation of frequency modulations: words with non-linear and non-flat frequency modulations were preferred to non-linear and flat words, whereas agreeableness scores of linear and non-flat words did not differ from the two other classes (Table 2, Fig. 8).
Table 2. Impact of listener’s characteristics and word acoustic structures on agreeableness scores\(^1\).

<table>
<thead>
<tr>
<th>Factor effects</th>
<th>Factor effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>( F_{1/58} = 28.193, P &lt; 0.001 )</td>
</tr>
<tr>
<td>Word x Sex</td>
<td>( F_{1/58} = 0.051, P = 0.821 )</td>
</tr>
<tr>
<td>Word x Age</td>
<td>( F_{1/58} = 8.759, P = 0.003 )</td>
</tr>
</tbody>
</table>

1. A General Binomial (GLM) analysed single factor effects (Word acoustic structures) and interactions between the word acoustic structure and the listener characteristics (sex and age) (\( f^2 = 0.374; P_{\omega} = 0.980 \)). Significant differences are shown in bold type.

**Figure 8.** Relative influence of word orientation and linearity on agreeableness scores given listeners [linear (L) or not (NL), flat (F) or not (NF)]. Binomial Glm test NS: \( P > 0.05 \); ***: \( P < 0.001 \).

Figure 9 shows spectrograms of typical agreeable (syllable 1: upwards and linear, syllable 2: downwards and non-linear) and disagreeable (syllable 1: flat and non-linear, syllable 2: flat and linear) voices.
Figure 9. Spectrograms of typical agreeable (a) and disagreeable (b) voices pronouncing the French greeting word “Bonjour”.

Moreover, whereas sex did not impact agreeableness scores, it varied with age (Table 2). The seniors found non-linear words (both flat and non-flat) more agreeable than the students (Fig. 10).
4. Discussion

This study confirms that (1) subtle and sudden (at both syllable and word levels) frequency modulations in unfamiliar human voices impact auditory evaluation; (2) the voices of phone operators are not evaluated similarly by all potential consumers; (3) the age more than the sex of the listener impacts his/her agreeableness appreciation; and (4) variations in frequency orientations more than in frequency linearity determine the agreeableness of a voice, non-monotonous voices being allocated the highest rates of agreeableness. However, we first acknowledge that our conclusions are drawn from analyses conducted on fundamental frequencies only. Hence, further investigations on resonant frequencies would be interesting. Also, we used a relatively small sample size of raters and speakers. Hence, replicating this study with more varied speakers (e.g. different age-sex classes), more raters and rating contexts are now necessary.

Our study is, to our knowledge, the first to evidence the differential impact of different kinds of subtle and sudden prosodic frequency modulations of human voices on auditory evaluation, particularly at the beginning of a word (first syllable). Hence, short-term prosodic changes seem as crucial as long-term changes. In line with this, some authors have shown that humans are able to categorize very quickly (in less than 200ms) a voice as neutral or emotional (Chen and Yang, 2012; Paulmann, Schmidt, Pell and Kotz, 2008). Thus, people attribute particular emotional states to particular voice prosody (Bach et al., 2009; Latinus and Belin, 2011; Mehrabian and Ferris, 1967). For example, some authors showed that voices with
large frequency variations are associated with a friendly personality, and speakers with high pitched voices are considered to be helpful and sincere (Brown et al., 1975; Chen and Yang, 2012; Zetterholm, 1998; Weirich, 2008). However, authors underline a cultural impact on the prosody evaluation. One study stresses that Koreans and Americans do not judge a given speaker in similar ways. More precisely, Americans, contrary to Koreans, associate a fast voice with a powerful and qualified person. However, when Korean participants live in the United States, their judgments converge with those of Americans. Indeed, they associate a fast rhythm with competence (Peng et al., 1993).

Variations in frequency modulation orientation impact more than frequency modulation linearity, for each syllable as well as for an entire word. Non-flat voices are particularly positively evaluated. A first hypothesis predicts that changes in orientation are preferred to changes in linearity. A second hypothesis predicts that changes in orientation are more easily detectable than changes in linearity by a human ear. Indeed, some authors underline that some sounds (like tones) are more easily perceived by the human ear than others (like clicks) (Szymaszek, Szelag and Sliwowska, 2006). Even though linearity was not the most determining acoustic criterion, it did impact the evaluation of the first syllable of the word “Bonjour.” Thus we suppose that linearity variations are detected to some extent, and that evaluation is based more on a higher agreeableness score for orientation changes. Moreover, we showed that the type of orientation preferred in the first (upward) and in the second (downward) syllables differed. This should warn researchers of the difficulty to draw firm and general conclusions based on prosodic acoustic analyses conducted at a broad (e.g. sentence) level, as often found in the literature (Brown et al., 1973; Brown et al., 1974; Brown et al., 1975; Peng et al., 1993; Smith et al., 1975). We must acknowledge that our conclusions at the word level are based on a single greeting item and cannot be generalized to other words. Also, our data concern only French-speaking female phone operators and French student listeners. This limit is mentioned by other authors concerning the risk of “overgeneralization” (Montepare and Zebrowitz- McArthur, 1987), and the important variations due to the cultural background of the subjects (Peng et al., 1993) and to the experimental context (Jones et al., 2008). Finally, future studies may want to go beyond the simple analysis of the fundamental frequency pattern and notably investigate variations in formants’ distribution.

Here the listener sex effect on auditory evaluation was very limited, although most studies currently underline its importance, probably because they focus on the characteristics allowing people to attribute masculine and feminine features to speakers in a “sexual” attraction context (Collins, 2000; Feinberg et al., 2005; Feinberg et al., 2006; Jones et al., 2008; Jones et al., 2010; Little et al., 2010; Puts, 2005; Vukovic et al., 2008). However, these authors found differences in the
women’s preferences for masculine men’s voices according to their estrogen cycle, to the listening context (short-term versus long-term mating contexts), and to their self-rated attractiveness (Feinberg et al., 2006; Puts, 2005; Vukovic et al., 2008). As our study in a “commercial” context could not evidence strong differences between men and women, the impact of prosody possibly varies according to the context of the conversation. It seems that, independently of context, men and women do not detect emotions in voices similarly. Men’s treatment of emotional prosody is slower than women’s (Besson et al., 2002). Moreover, the brain organizations for processing prosody differ between sexes (Imaizumi, Homma, Ozawa, Maruishi and Muranaka, 2004; Rymarczyk and Grabowska, 2007). These differences in perception and treatment of human voices can explain the fact that we found subtle differences in the agreeableness scores of the participants according to their gender. Particularly, we underlined that women allocated higher scores to upward frequency modulations than men did.

Conversely to sex, age of listeners seems to be an important factor for voice appreciation. Indeed, older people allocated higher scores to some voices than did younger people. Two hypotheses could explain this. First, because of an age-dependent auditory sensitivity seniors could be less sensitive than juniors to acoustic changes or to the associated emotions. Several authors have reported that detection and perception of prosody varies with age. Older people are less efficient and accurate in the detection of emotionally associated prosodic changes in a voice (Mill, Alink, Realo and Valk, 2009; Mitchell, 2007). The second hypothesis predicts an age-dependent acoustic preference. Some authors have shown that the positive or the negative valence assigned to voices varies with age (Fecteau, Armony, Joanette and Belin 2005). This is confirmed by our data showing that seniors did not systematically rate higher all the voices, neither all the modulated voices, but only some voices.

5. Conclusion

To summarize, this study is the first to our knowledge to evidence the impact of human voice prosody on a listener in a commercial context when it is particularly important to use an agreeable voice, and it raises several fundamental and applied perspectives. Moreover, whereas studies generally focus on long-term frequency changes, our results underline the importance to focalize on subtle variations that have been neglected until now.
Acknowledgments

The idea of this study emerged from scientific discussions among members of the interdisciplinary research network GIS “Cerveau, Comportement, Société”. This study received financial support from the French Ministry of Research, the C.N.R.S. We thank the receptionists and all the students and seniors for their participation in this study. We are grateful to Ann Cloarec for correcting the English.

References


**Authors’ addresses**

Vanessa André  
C.N.R.S., Ethologie animale et humaine  
U.M.R. 6552 - Université de Rennes 1  
Campus de Beaulieu 263 Avenue du General Leclerc  
Bâtiment 25, 35042 Rennes  
France  
andre_vanessa@hotmail.fr

Christine Petr  
Laboratoire IREA-LEGO  
Université de Bretagne Sud  
Campus de Tohannic 56000 Vannes France  
christine.petr@univ-ubs.fr

Nicolas André  
C.N.R.S., Ethologie animale et humaine  
U.M.R. 6552 - Université de Rennes 1  
Campus de Beaulieu 263 Avenue du General Leclerc  
Bâtiment 25, 35042 Rennes  
France  
nicolas.andre@univ-rennes1.fr

Martine Hausberger  
C.N.R.S., Ethologie animale et humaine  
U.M.R. 6552 - Université de Rennes 1  
Campus de Beaulieu 263 Avenue du General Leclerc  
Bâtiment 25, 35042 Rennes  
France  
martine.hausberger@univ-rennes1.fr

Alban Lemasson  
Université de Rennes 1  
Ethologie animale et humaine  
UMR6552 - C.N.R.S., Station Biologique 35380, Paimpont  
France  
alban.lemasson@univ-rennes1.fr
About the authors

Vanessa André is a PhD student from University of Rennes 1 studying human ethology (sensory perception, communication and welfare).

Christine Petr is a Professor in management sciences studying the impact of new technologies on consumer behaviors.

Nicolas André holds a Master degree in ethology from University of Rennes 1.

Martine Hausberger is research director at CNRS in the Animal and Human Ethology (EthoS) lab. She studies social behavior, cognition and communication in songbirds and human-animal relationships.

Alban Lemasson is the head of EthoS lab and is Professor at University of Rennes 1. He studies vocal communication and social behaviors in mammals.