

Tone of voice or what was said? The impression non-native speakers of English make on Australian English native listeners

Chiharu Tsurutan
Griffith University

Abstract

This study investigates the relative impact of verbal expression and tone of voice when native speakers of English form an impression of non-native speech. Four expressions of inquiry uttered in two tones by non-native speakers were judged by native listeners and analyzed using an ordinal Probit model. Plain expressions received lower scores than polite expressions in both tones, suggesting that appropriate expression is more important than tone of voice. It was found that while a friendly tone can enhance listeners' impression on speakers, particularly when the expression is plain, the impact of tone of voice is less evident when appropriate expressions are used. The study revealed no statistically significant gender differences. On the basis of the study, we offer a pedagogical suggestion that beginners need some instruction to guide their choice of textbook expressions for an effective real-world interaction.

Keywords: tone of voice, impression of speech, non-native speech, communication, language attitudes

Introduction

In a multicultural society, people have plenty of opportunities to communicate with speakers from different language backgrounds. In the society of the target language, non-native speakers are often the subject of native speakers' negative observation (Eisenclas & Tsurutani, 2011; Gluszek & Doridio; 2010; Munro, 2003), ranging from comments "blunt or rude" to statements such as, "I cannot stand their tone of voice".

However, it is questionable whether listeners are really referring to the tone of voice or to the expressions the non-native speaker uses. The first language (L1) transfer could certainly affect the prosodic feature of their speech, and second language (L2) speakers can very often sound blunt because they do not have the capacity to pay attention to intonation while working out what to say. Yet non-native speakers' unconventional expressions may unintentionally convey an impolite impression. People at times may refer to the expression and to the sentence structure chosen for the utterance when they say the speaker's tone of voice is inappropriate. Non-native speakers may inadvertently use expressions that are inappropriate for a given situation

because they use a direct translation from their L1 or phrases from a textbook, while ignoring the context in which the expression should be used. It is therefore of interest to language educators to investigate the relative impact of tone and expression that L2 learners use in their utterances. This will also flag the importance of screening the expressions used in English phrase books published overseas.

Although prosodic features range from stress and rhythm to accentuation and intonation, intonation is the main prosodic feature explored in this study, since here our concern is the overall impression of speech based on speaker's tone of voice. The impacts of prosodic features of speech, which are difficult to describe and explain, are often marginalized in language teaching. Some people even undermine the role of prosody in second language acquisition and think their learned language will be acceptable as long as they use the right expression. This study investigated the relative impact of non-verbal and verbal behavior in the form of tone of voice and expression of inquiry, produced by non-native speakers of English. The result should also inform us of useful information about the role of prosody in communication for L2 learners. To this end, two contrasting tones of voice and expression were used as parameters of stimuli. Using these stimuli, the perception of L2 spoken English by Australian English native listeners was examined to see whether listeners reacted to unfriendly tones more strongly than to inappropriate expressions, or vice versa. Findings of the study provide useful information for language teachers and learners, and may inspire them to reconsider the impact of prosody in their teaching/learning.

Verbal content vs. non-verbal cues

People communicate successfully by using appropriate verbal content and non-verbal cues. Verbal content refers to the actual linguistic content delivered by the speaker, basically, what s/he has said. Non-verbal cues are usually given not only by the speaker's posture, physical movement, eye contact, facial expression and hand/body gestures but also by their tone of voice, namely any signals that convey the speaker's intention, apart from the linguistic content. The importance of non-verbal behavior in communication has been well acknowledged since the era Mehrabian (1971) presented a formula informed by his study on the impact of verbal and non-verbal behaviors. Mehrabian's now well-known formula is 7% verbal, 38% vocal (tone of voice) and 55% facial expression, when these three factors are inconsistent in expressing the speaker's feelings. His research was initially based on the feeling, "like-dislike", with suggestion of its possible application to feelings and attitudes in general (Mehrabian, 1971). However, this formula was challenged by various researchers and their findings imply that the different scenario prepared for particular communication setting and the different method of measurement used in their studies could bring out a different ratio. For example, Krauss, Apple Morency, Wenzel and Winton

(1981) reported verbal content to be the best predictor when judging speakers' emotions. However, Tusing and Dillard (2000) pointed to the significance of vocal marker in perceived dominance; non-native speakers' choice of linguistic expressions and foreign tone of voice could potentially add another dimension to their communication with native speakers.

The non-verbal cue this study aims to investigate is the tone of voice used by non-native speakers of English. Tone of voice was chosen to identify whether non-native speakers' tone of voice is an issue in communication rather than, or as well as, the expressions they use. The effect of tone of voice in communication has been investigated in various social settings where subtle nuance in tone can change the meaning of the speaker's message. The typical settings were conveying emotions and attitudes (Brown, Winter, Idemaru & Grawunder, 2014; Culpeper, 2011; Menezes, Erikson, & Franks, 2010; Nadeu & Prieto, 2011; Scherer, 2000; Shochi, Rilliard, & Erikson, 2009), telephone communication (Hecht & LaFrance, 1995), the speech of professionals working in medical or psychological health (Ambady, LaPlante, Nguyen, Rosenthal, Chaumeton, & Levinson, 2002) and computation (Pentland, 2005). On the other hand, studies that investigated the impact of tone of voice in relation to other verbal and non-verbal factors have been relatively few (Bryant & Fox Tree, 2005; Laplante & Ambady, 2002, 2003; Zuckerman, Amidon, Bishop & Pomerantz, 1982). Further, the scope of research in each of these studies is restricted to one parameter; i.e. prosodic features of ironic speech (Bryant & Fox Tree, 2005), which studies only irony among various emotions; and tone of voice vs. facial expression (Zuckerman et al., 1982), which discusses the relation between tone of voice and facial expression. This limited scope of studies is largely due to the difficulty of singling out the effect of tone of voice in natural communication and of creating an appropriate empirical setting to identify the role of tone of voice for expressing various emotions at once. No published study has reported on the impact of tone of voice in the communication of non-native speakers with native speaker interlocutors.

Laplante and Ambady's study (2002) compares the effect of tone of voice with the effect of verbal content, using native speakers of English. Their study involved the use of positive and negative comments, delivered in both positive and negative tones, in relation to students' academic results. They found that non-verbal cues played a limited role in changing the impression of the message, and tone of voice was less effective in delivering negative content. It was anticipated that participating students would naturally focus on the content of the message, as academic results are an extremely important part of students' lives and the tone of voice used in delivery did not make any difference to the impact of the result for them. The research design, using two tones and different expressions, can be replicated to test non-native speech in different message settings and content. It would be helpful for revealing the role/s that the tone of voice plays in native listeners' perceptions.

In this study, two tones of voice, friendly and blunt, were used to

inquire about directions in different degrees of polite expression uttered by a non-native speaker. The setting of inquiry is a very common and likely scenario that non-native speakers come across in their new language environment. The two tones, friendly and blunt, were chosen as friendly tones are believed to help communication in every social setting, whereas sounding blunt and arrogant is the last thing L2 learners wish for when they first enter into a new target language community. Thus, these factors were used as a parameter of stimuli. English does not have a variety of expressions to convey different degrees of formality or politeness as compared with Japanese or Korean in societies that are seen to be more vertically structured. Nevertheless, polite expressions play a major role when asking a favor even in English (Maynard, 1997). In the real world, the impact of tone of voice on messages is less easy to measure than some other non-verbal cues (e.g., facial expressions and gestures). As well, the context in which the message is delivered can interfere with the result. In this study, a controlled setting and a prepared scenario were used to extract information about utterances of the same sentence in different tones of voice. Although there are many ways to ask for direction, a few typical expressions were chosen after consulting with native English teachers. Increasing the number of stimuli lengthens the time of task unnecessarily and will make the semantic and pragmatic differences between sentences too subtle to make a judgement for listeners. Four sentences were sufficient to represent polite vs. plain versions and adequate to present the semantic and pragmatic differences between sentences. Four different inquiry sentences that differ in their level of politeness were prepared, ranging from a polite request to a very direct wh-question. Two expressions are considered to be polite expressions commonly used by native speakers, while the other two expressions are found under the section “asking directions” in a travel guide book published in Japan. One of latter two, “Where is XX?”, is the direct translation of a Japanese sentence of inquiry. The utterances captured in laboratory recordings are not exactly the same as natural utterances, however the stimulus sentences were recorded by four experienced language teachers to make sure the difference between the two tones was maintained. All four teachers’ first language was Japanese and all had similar academic backgrounds (postgraduate degrees).

Methodology of listening task

The focus of this study is the relative impact of non-native speakers’ tones of voice over the verbal content of their speech, as perceived by native listeners; i.e., how tone of voice influences individuals’ perceptions of different levels of polite expressions of inquiry. Two different tones of voice and four different expressions of inquiry recorded by non-native speakers were mixed to create 32 stimuli. Native speakers of Australian English listened to the speech stimuli and judged their impression of the speaker using a Likert scale, for example 3 was good, 2 was neutral, and 1 was bad. Analysis of their scores

should be able to determine the relative importance of tone of voice over verbal content in their judgement.

Materials and material construction

The following sentences of inquiry that are commonly addressed to strangers were used as materials. The word “central” [sɛntɹəl] has the alveolar approximants [ɹ] and [l], which are difficult sounds for Japanese native speakers (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999) and can easily enhance the trace of foreign accent in their production.

- 1) Excuse me. Could you please tell me the way to the central station?
- 2) Excuse me. Can you tell me the way to the central station?
- 3) Excuse me. Where is the central station?
- 4) Excuse me. I want to go to the central station.

Sentences 1) and 2) are considered to be polite ways to ask direction by native speakers. Sentences 3) and 4) are expressions that appear in the travel guidebooks available from bookshops in Japan. In particular, 3) is a direct translation of a Japanese expression of inquiry and is often used by beginners of English. It is expected that the first two sentences, 1) and 2), will give better impression of the speaker than the second two sentences, 3) and 4).

Four Japanese native speakers (two males and two females) who have each resided in Australia for more than 20 years recorded the four sentences in two different tones. These speakers are fluent in English, but had clear traces of a Japanese accent. All of the speakers were Japanese language instructors, and were good at acting to produce different tones of voice. Their ages ranged from mid-forties to mid-fifties. The speakers were given the following instruction: Please say the phrases 1) – 4) nicely (A). Then, say them again arrogantly (B). Four friendly versions were recorded first, then four arrogant versions followed. Speakers produced each version twice and the first trial was used unless there was an acoustic flaw in performance.

In their performances, the friendly tones had a higher pitch, wider pitch range and slower speech rate, while utterances with blunt tones were delivered in a low pitch and faster speech rate. (Only male speaker 1 used a consistent speech rate for all four sentences in opposite ways.) This corresponds with prosodic characteristics of friendly and arrogant speech reported in previous studies (Menezes et al., 2010; Nadeu & Prieto, 2011; Tsurutani, Shi, & Minematsu, 2016) The following table presents the acoustic measurements of their performance. The longest sentence, Sentence 1, was used for the purpose of presenting a clear difference between two versions of tone in measurement.

Table 1

Acoustic measurements of 4 speakers (measured in Sentence 1)

Speakers	Speech rate (syllable/sec)*		Duration of utterance (msec)		Average Pitch (Hz)		Pitch range (Hz)	
	F	UF	F	UF	F	UF	F	UF
Male 1	5.99	5.40	2.67	2.96	153	> 147	174	> 107
Male 2	4.13	4.98	3.87	3.21	156	> 147	167	> 131
Female 1	4.18	5.59	3.83	2.86	239	> 210	223	> 161
Female 2	4.66	4.68	3.43	3.42	260	> 231	220	> 167

*Faster speech has a higher figure.

F=friendly, UF=unfriendly

Their level of performance was checked by two other native speakers who have knowledge of linguistics. It was confirmed that the two tones clearly presented intended tones.

Participants in the listening task

Ten male and 12 female Australian English speakers participated in the listening task, receiving a small payment for doing so. Their ages ranged from 39 to 69 (average age 57 years). The researcher contacted people who do not have regular contact with non-native speakers, to avoid possible bias in judgement by having a foreign friend who has a similar accent. In a multicultural society like Australia, people cannot avoid having contact with non-native speakers while carrying out their everyday activities, such as shopping or dining in ethnic restaurants. However, these contacts do not occur on a regular basis and were considered to be a minor part of the participants' everyday lives.

Method of listening task

The stimuli were given to the participants either as a CD or sound file. In the sound file, following on from three practice sentences, the stimulus sentence was played twice, each with a 1 second interval and a 3 second inter stimulus interval. The listeners were asked to judge whether the utterance gave a good impression, a neutral, or a bad impression on a 3-point Likert scale. A 3-point scale was sufficient for a quick impressionistic judgement and was suitable for calculating the result by ordinal probit modeling. Two different versions of randomized order of stimuli presentation were used and distributed randomly among participants. The entire task took approximately 15 to 20 minutes for participants to complete, including the information sheet for their background. Whether the perceived impression was ranked according to the goodness of tone; A -> B (1) A -> 4) A, 1) B -> 4) B), or ranked according to the appropriateness of expressions; 1) A, 1) B -> 4) A, 4) B would determine the strength of the two factors.

Results

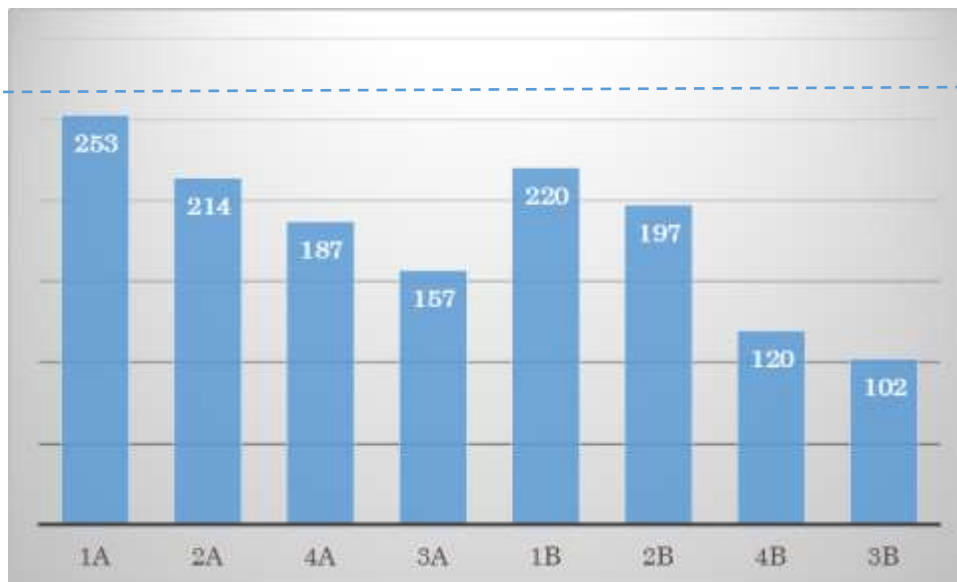
The data were analyzed using an ordinal Probit model (Agresti, 2010). The gender of both listeners and speakers was found to be a non-significant variable at the 95% level. From the parameter estimates we observed, Sentence 1 scored the highest response, followed by Sentence 2, 4 then 3. Tone B (blunt) generally received a lower score than Tone A (friendly), with this effect being lower in sentences with more polite expressions.

Table 2

Parameter estimates and credible intervals for model terms.
(Significance is indicated with an asterisk)

Term	Parameter Estimate
Intercept	2.797 (2.412, 3.204)*
Sentence 2	-1.284 (-1.737, -0.849)*
Sentence 3	-2.396 (-2.851, -1.958)*
Sentence 4	-1.820 (-2.272, -1.386)*
Tone B	-1.134 (-1.594, -0.690)*
Sentence 2: Tone B	0.766 (0.205, 1.337)*
Sentence 3: Tone B	-0.217 (-0.806, 0.382)
Sentence 4: Tone B	-0.251 (-0.821, 0.326)
γ_2	1.435 (1.286, 1.589)

The following graph shows the total scores the stimulus sentences received. As expected, Sentence 1 had the highest score, followed by Sentences 2, 4 and 3.



264

Figure 1. Total scores obtained from 22 listeners (highest possible score=264)

The performance of 1A is close to the highest possible score, which means that native listeners did not mark down non-native speech due to the speakers' accent. Between the two versions of the same sentence, Tone A always had a higher score than Tone B, which suggests that tone helps to give a better impression particularly in the blunt expressions, as uttered in Sentences 4 and 3. However, the effect of tone is weaker in Sentences 1 and 2. This suggests that a friendly tone can be particularly helpful when using a blunt expression. If speakers use a friendly tone even when uttering a plain expression such as that of 4A, "I want to go to the central station", it can sound almost as good as the more polite expressions spoken in an arrogant tone, in Sentence 2B. It is very likely that the wrong choice of expression by a non-native speaker annoys the local people who the non-native speaker asks for help. The posterior probability of each sentence and tone being scored 1–3 is given in Table 3.

Table 3
Posterior probability (95% confidence interval) of each sentence and tone being perceived as scores 1–3

	Score 1	Score 2	Score 3
Sentence 1 Tone A	0.003 (0.001, 0.008)	0.088 (0.040, 0.151)	0.910 (0.842, 0.959)*
Sentence 1 Tone B	0.050 (0.026, 0.083)	0.361 (0.284, 0.436)	0.589 (0.489, 0.686)*
Sentence 2 Tone A	0.067 (0.037, 0.107)	0.402 (0.329, 0.472)	0.531 (0.433, 0.628)*
Sentence 2 Tone B	0.128 (0.080, 0.187)	0.485 (0.426, 0.541)*	0.387 (0.298, 0.480)
Sentence 3 Tone A	0.346 (0.260, 0.437)	0.502 (0.447, 0.554)*	0.152 (0.098, 0.216)
Sentence 3 Tone B	0.826 (0.742, 0.896)*	0.164 (0.100, 0.239)	0.009 (0.003, 0.020)
Sentence 4 Tone A	0.166 (0.110, 0.234)	0.509 (0.456, 0.560)*	0.324 (0.242, 0.414)
Sentence 4 Tone B	0.657 (0.560, 0.749)*	0.309 (0.232, 0.386)	0.034 (0.016, 0.060)

*The highest percentage in each column

This table provides the distribution of each score, which was not revealed in Figure 1, and supports the same result. Sentence 1A provides a good impression 91% of the time, while that possibility is reduced to 58.9% by the use of Tone B. This reduction is as wide as the increase in the bad impression made by the use of Tone B in Sentences 3 and 4, at 48% (82.6-34.6) and 49.5% (65.7-16.6) respectively. On the other hand, the decrease is not so obvious in Sentence 2, as the score goes down only from 53.1% to 38.7%.

The findings in this section are summarized in the following two points:

- 1) A friendly tone helps to improve the impression of the speaker, particularly when the expression they use is blunt.

- 2) When polite expressions are used, the role of tone is not as significant as in blunt expressions.

Discussion and Implications

In this study, utterances with a polite expression had higher scores than the ones with a plain expression regardless of tones used for the utterance. It means that listeners primarily form their impression of a speaker based on the content of their utterance. In everyday communication, to some extent speakers are able to predict what their interlocutor will say next in the course of conversation. Non-native speech does not always follow this norm, and an expression that is unusual and unexpected for the native speaker-interlocutor can hinder smooth communication. Dismayed native speakers in this situation could end up commenting on the non-native speech as an unfriendly tone of voice.

In this study, the tone of voice did impact on the speakers' impression, but not as strongly as verbal content, which supports the finding of the previous study on native speakers' performance (Laplante & Ambady, 2003). That is, it is more likely that native listeners are referring to the expression the non-native speaker uses when they say "tone of voice". No one intends to offend someone they have not met in their brief first encounter. Non-native learners would not intentionally use an arrogant tone of voice, however, due to their lack of pragmatic knowledge, they could sometimes use an inappropriate expression as observed in the selection of inquiry expressions in the Japanese travel guide. This would be the case particularly for beginners who rely heavily on a textbook or a guidebook for their choice of expression to use. The expressions in the phrase books should be checked carefully by educators to avoid unnecessarily unpleasant experiences for both listeners and speakers. At the same time, language instructors need to make learners aware that expressions in textbooks for beginners use a simple grammar and are not necessarily appropriate for some social settings. Learners are also to be reminded to check the context when they use the expression they have newly learned. The expressions and context used in this study limited the scope of the investigation to inquiry of directions in non-native speech. A different context and setting could be explored in a future study.

In order to improve the quality of communication in a multicultural society, native listeners need to be considerate of the difficulty non-native speakers experience, while non-native speakers need to pay more attention to the appropriateness of their expression. This study provides evidence that tone does play a role in the way the listener judges the speaker, however this is a secondary factor. The primary factor is the verbal content in which the sentence is expressed. In the study, two contrasting tones were used to examine the role of tone of voice in oral communication. However, in real life it is highly unlikely that a deliberately arrogant tone of voice would be used by non-native speakers when asking directions. When listeners do not form a

good impression of non-native speakers in their first, brief encounter, the problem could lie in the speaker's lack of pragmatic knowledge in choosing the right expression. Both native listeners and non-native speakers need to be aware of this point and work towards better communication.

Acknowledgements

I thank the Griffith Social and Behavioral Research College for their statistical support.

This study was funded by the School of Languages and Linguistics, Griffith University.

References

- Agresti, A. (2010). *Analysis of Ordinal Categorical Data* (2nd ed.). USA: Wiley.
- Ambady, N., LaPlante, D., Nguyen, T., Rosenthal, R., Chaumeton, N., & Levinson, W. (2002). Surgeons' tone of voice: A clue to malpractice history, *Surgery*, 132(1), 5-9.
- Bradlow, A. R., Akahane-Yamada, R., Risoni, D., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production, *Perception and Psychophysics*, Vol. 61 Issue 5, 977-985.
- Brown, L., Winter, B., Idemaru, K., & Grawunder, S. (2014). Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, *Journal of Pragmatics*, 66, 45-60.
- Bryant, G. A., & Fox Tree, J. E. (2005). Is there an ironic tone of voice? *Language and Speech*, 48(3), 257-277.
- Culpeper, J. (2011). It's not what you said, it's how you said it! Prosody and impoliteness. In Linguistic Politeness Research group (Ed., pp. 57 to 83), *Discursive approach to politeness*, Berlin: De Gruyter Mouton.
- Eisenchlas, S., & Tsurutani, C. (2011). You sound attractive: Perceptions of foreign accent in a multicultural context. *Australian Review of Applied Linguistics* 34 (2), 218-238.
- Gluszek, A., & Dovidio, J. F. (2010). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review*, 14(2) 214-237.
- Hecht, M. A., & LaFrance, M. (1995). How (fast) can I help you? Tone of voice and telephone operator efficiency in interaction. *Journal of Applied Social Psychology*, 25(23), 2086-2098.
- Krauss, R. M., Apple, W., Morency, N., Wenzel, C., & Winton, W. (1981). Verbal, vocal and visible factors in judgements of another's affect. *Journal of Personality and Social Psychology*, 40(2), 312-320.
- Laplante, D., & Ambady, N. (2003). On how things are said: Voice tone,

- voice intensity, verbal content and perceptions of politeness. *Journal of Language and Social Psychology* 22 (4),434-441.
- Laplante, D., & Ambady, N. (2002). Saying it like it isn't: Mixed messages from men and women in the workplace. *Journal of Applied Social Psychology*, 32(12), 2435-2456.
- Maynard, S. K. (1997). *Japanese communication*. University of Hawaii Press.
- Mehrabian, A. (1971). *Silent messages*. California, USA: Ward worth publishing.
- Menezes, C., Erickson, D., & Franks, C.. (2010). Comparison between linguistic and affective perception of sad and happy – A cross-linguistic study. *Proceedings of Speech Prosody*, Chicago, Illinois. Retrieved October 7, 2013, from <http://isle.illinois.edu/sprosig/sp2010/program.html>
- Munro, M. (2003). A primer on accent discrimination in the Canadian context, *TESL Canada Journal*, 20(2), 38-51.
- Nadeu, M., & Prieto, P. (2011). Pitch range, gestural information, and perceived politeness in Catalan. *Journal of Pragmatics*, 43(3), 841-854.
- Pentland, A. (2005). Socially aware computation and communication. *Computer*, 38(3), 33-40.
- Scherer, K. (2000). A cross-cultural investigation of emotion inferences from voice and speech: Implications for speech technology. *Proceedings of Interspeech 2000*. Beijing. Vol 2, 379-382.
- Shochi, T., Rilliard, A., & Erickson, D. (2009). Intercultural perception of English, Japanese and French social affective prosody. In S. Hancil (Ed.), *The role of prosody in affective speech: Linguistic insights* (pp. 31-59). Bern: Peter Lang.
- Tsurutani, C., Shi, S., & Minematsu, N. (2016). Prosodic characteristics of Japanese polite speech spoken by native and non-native speakers, *Proceedings of International Australian Speech Science and Technology Conference*, Sydney. <http://www.assta.org/sst/2016/programme.html>
- Tusing, K. J., & Sillard, J. P. (2000). The sounds of dominance: Vocal precursors of perceived dominance during interpersonal influence. *Human Communication Research*, 26(1), 148-171.
- Zuckerman, M., Amidon, M. D., Bishop, S. E., & Pomerantz, S. D. (1982). Face and tone of voice in the communication of deception. *Journal of Personality and Social Psychology*, 43(2), 347-35.

Note on Contributor

Chiharu TSURUTANI is an adjunct senior lecturer in Japanese language and linguistics at Griffith University. Her research area is second language acquisition, particularly in the area of phonetics and phonology. She has been conducting research in various issues around foreign accented speech. In recent years the focus of her research is on the sociolinguistic and pragmatic aspects of foreign accent. Email: c.tsurutani@griffith.edu.au