Perception of American and Australian English "can" and "can't" by Japanese Listeners: How to Teach "can" and "can't"

Kinuko Takahashi[1], Tomohiko Ooigawa[2]

[1,2]Sophia University, Japan,
[1]*kinuko@dc4.so-net.ne.jp*, [2]*ooigawaferchichi@gmail.com*

Abstract

In order to follow up on a past study ([1]), the present research examined Japanese listeners' perception of American and Australian English utterance-final "can" and "can't" ([kæ̃n]-[kæ̃ʔ] and [kɐ̃n]-[kɐ̃ʔ], respectively).[1] The results showed that Japanese listeners discriminated the Australian contrast better than the American contrast because of their ability to detect the differences in the vowel quality between [kɐ̃n] and [kɐ̃ʔ]. The results of the present study are eventually going to be applied to teach Japanese EFL students how to discriminate the contrast, as they tend to misunderstand messages due to poor discrimination between the two.

Keywords: Perception, English, can, can't, Japanese.

1. Introduction

1.1. Background

The present research was initiated by a real-life episode. "Next week I can/can't." This was uttered by a person who appeared to be an English teacher and native speaker of English. He happened to be sitting in the coffee shop where we (i.e., the authors of the present study) were. The moment we overheard this, we looked at each other and both of us whispered, "can, or cannot?" We wondered whether the first author's colleagues, interpreters and other learners have the same problem with identifying the two. In order to obtain an answer to this question, we carried out a survey using a simple questionnaire, asking interpreters whether they had had any trouble identifying "can" and "can't" on their interpreting assignments.

The results of the survey in [2] showed that ten out of eleven professional interpreters had had trouble with identifying the two on their interpreting assignments. According to the information obtained from the interpreters who participated in the study, they had witnessed their clients, native speakers, confirming and clarifying the message with either of *can/can't* with each other during important business meetings.

We may use the non-contracted form "cannot" in order to avoid confusion. However, when the first author used the non-contracted form at a university in London, she was told by a native speaker of English "No native speaker says that." Also, a reference ([3]) states: "If you**'re** a non-native learner of English, remember that weak and contracted forms are necessary for anyone with the goal of approaching fluent native-speaker English." (p.18). In summary, "can't" is often used by native speakers of English in the real world. Therefore, even non-native speakers are expected to be accurate in discriminating between the two. However, this is not an easy task, as the final "t" of the contracted form "can't," which Japanese learners are instructed to use as a clue to discriminate between the two, is often deleted. So, there is a question as to how native speakers discriminate between these two. What is their clue to identify positive and negative. This is where our "*can-can't* Project" was launched.

1.2. Production and perception by native speakers

In English, the negative form of the auxiliary verb "can" is "cannot," and its contracted form is "can't." Some English accents such as Received Pronunciation (RP) and Australian English have a difference in vowel quality between "can" and "can't," while American English does not have any phonemic-level vowel distinction in contrast. According to [4] and [5], the sequences of phonemes of "can't" are /kænt/, /kɑnt/ and /kɐnt/ for American English, RP and Australian English, respectively, while that of "can" is /kæn/ in common.[2]

Some function words, including auxiliary verbs have two kinds of realizations known as "weak (i.e., unstressed) forms" and "strong (i.e., stressed) forms" [3, 6, 7]. The realization of "can" is [kɔ̃n] for the weak form and [kæ̃n] for the strong form (see, the acoustic data shown in [2, 8, 9, 10]).[3] On the other hand, the contracted forms (e.g., "can't") are usually realized as strong forms ([3]). In addition, "can" is pronounced as a strong form at the end of a sentence, so that "can" is realized as [kæ̃n] in utterance-final

position [3, 6]. As to the phonetic realizations of "can't," it is usually realized without the [n] sound [11, 12, 13, 14, 15] in RP and American English. While, to our knowledge, no study has indicated the phonetic details of "can't" in Australian English, according to the sounds recorded for our studies, Australian "can't" tokens were also realized without [n] (see, Figure 1). Moreover, a syllable-final /t/ can be realized as [ʔ] before a consonant or a silence in English [12, 14, 15, 16].[4]

In summary, both in American and Australian English, utterance-medial "can" can be [kɔ̃n] or [kæ̃n], while utterance-final "can" is [kæ̃n]. Both in medial and final positions, American English "can't" can be [kæ̃t] or [kæ̃ʔ], while Australian English "can't" can be [kɐ̃t]/[kɐ̃ʔ] (see, Table 1).

Table 1: Realizations of "can" and "can't."

American	Weak form	Strong form
can	[kɔ̃n]	[kæ̃n]
can't	n/a	[kæ̃t] / [kæ̃ʔ]
Australian		
can	[kɔ̃n]	[kæ̃n]
can't	n/a	[kɐ̃t] / [kɐ̃ʔ]

American	Medial	Final
can	[kɔ̃n] / [kæ̃n]	[kæ̃n]
can't	[kæ̃t] / [kæ̃ʔ]	[kæ̃t] / [kæ̃ʔ]
Australian		
can	[kɔ̃n] / [kæ̃n]	[kæ̃n]
can't	[kɐ̃t] / [kɐ̃ʔ]	[kɐ̃t] / [kɐ̃ʔ]

Nakayama [17] has pointed out the perception difficulty on "can" and "can't," indicating as follows; "can" originally has the sound [n] at the end, so that [n] does not function as an acoustic cue to discriminate between "can" and "can't." On the other hand, the sound [n] can be an acoustic cue to discriminate the positive-negative contrasts of the other auxiliary verbs (e.g., *should-should**n**'t*). Nakano [18] has indicated the perception difficulty in identifying "can" and "can't," if the two words are stressed and embedded in a sentence such as: "I can/can't teach him," mentioning that the cue is the duration of "can" (shorter) and "can't" (longer).

So far, three past studies ([9, 19, 20]) have examined the perception of the American English "can" and "can't" by native speakers of English. These studies re-synthesized an original "can [kæ̃n]" sound token, and generated a continuum ranging from actual "can" to virtual "can't." In [9], the closure duration between the end [n] of "can" and the next plosive [g] in a sentence token ("I can go") was made longer. In [19] and [20], both the durations of [æ̃n] and the closure of the next plosive were shortened and lengthened in tokens: "I can teach," "I can get it." According to the results in the three studies, the shorter the duration of [æ̃n] and the

longer the closure duration were, the more frequently native speakers identify the re-synthesized "can" sounds as "can't." The studies ([19]) also examined the perception in the utterance-final context (i.e., "I can"), where the closure duration never existed. The results revealed that the shorter the duration of [æ̃n] was, the more frequently the native speakers identify the re-synthesized "can" sound as "can't." The results can be interpreted as showing that "can" becomes "can't," if the [n] sound is very short or absent. It is very interesting to note that two non-American English native speakers participated in the test of the perception of American English "can" and "can't" in [19, 20]. The listeners were eight Americans, a British from Essex and an Australian from Melbourne, and the studies did not report the differences in the results among the listeners. Therefore, the production and perception studies shown in this section indicate that the most important acoustic cue to discriminate/identify "can" and "can't" must be the existence of a clear and long [n] at least across some English accents.[5]

It is very interesting to note that [n] in the case of *can-can't* functions in the opposite direction from the other auxiliary verbs for the positive-negative contrast distinction. In other words, [n] functions as a positive-marker for *can-can't*, while it functions as a negative-marker for the other auxiliary verbs. It is also interesting that the stressed "can't" is shorter than the stressed "can" as indicated in [8, 19, 20], which is an opposite pattern to that mentioned in [18].

1.3. Perception by Japanese listeners

The study ([2]) showed that Japanese non-returnee (i.e., home-grown/monolingual) listeners did not perfectly identify the contrast of "So this system can/can't work." in a passage. The studies ([1, 10]) firstly dealt with Japanese listeners' perception of the contrast in utterance-final position. According to the results, the listeners (i.e., non-returnees) poorly discriminated the American [kæ̃n]-[kæ̃ʔ] contrast, while they discriminated the Australian [kæ̃n]-[kɐ̃ʔ] contrast better than the American contrast. In these studies ([1, 10]), a forced-choice identification test was also carried out, in which the choices were "カン /kaN/," "カントゥ /kaNtu/," "カ /ka/," "キャ /kja/," "キャントゥ /kjaNtu/," and "キャン /kjaN/." According to the results, the listeners tended to perceptually assimilate both American "can [kæ̃n]" and "can't [kæ̃ʔ]" to the identical choice /kjaN/, so that the American contrast was difficult to discriminate, which was understandable. However, the listeners also tended to assimilate both Australian "can [kæ̃n]" and "can't [kɐ̃ʔ]" to the

identical choice /kaN/, but the discrimination was easier than that of American English, which was not understandable. A research question was raised as to the reason why the Australian English "can [kæn]" and "can't [kɐ̃ʔ]" contrast is easier to discriminate for Japanese listeners than the American English "can [kæ̃n]" and "can't [kæ̃ʔ]" contrast.

1.4. Purpose of the study

In order to solve the research question, we conducted another forced-choice identification test using different choices from those used in the past study, replicating the discrimination test. The results of the present study will eventually be applied to teach Japanese EFL students how to discriminate the contrast, as they tend to misunderstand messages due to poor discrimination of the two.

As mentioned in 1.1, the present research originated in a real episode encountered by the present authors and was accelerated by experiences of professional interpreters on their interpreting assignment. The misperception of *can/can't* sometimes misdirects the conversation, which might result in serious consequences, as the meaning of the two is opposite. Therefore, it is quite important for educators to familiarize Japanese EFL students with the two.

We dealt with utterance-final "can" and "can't" (without audible "t" release) for the following reason; The contrast seems the most difficult to discriminate in utterance-final position. If one can discriminate the contrast in the context, it is also the case with other contexts. We dealt with the contrast in both American and Australian English for the following reason; Given that there are a variety of Englishes, we might choose to produce only one of the English accents as a speaker. However, as a listener, we have no control over the pronunciation. We need to be able to recognize all possible variations of "can" and "can't" phonetic realizations, because we may encounter English speakers who produce these variations. We have to listen to utterances in various accents and understand them in the real world.

2. Experiment

In order to follow up on the past study, we carried out another forced-choice identification test, replicating the discrimination test. None of the speakers and listeners had any perceptual and physical difficulties participating in the recording or the tests.

2.1. Identification test

2.1.1. Stimuli

The stimuli used in the present study were identical to the target stimuli in the two past studies ([1, 10], see, Table 2). The paired words "can" and "can't" were used as the target contrast. All the words used in the tests, including the distracters, were embedded in the same carrier sentence: "Next week I _." This sentence makes sense with all of the words.

Two native speakers of American English (male and female) and a native speaker of Australian English (female) individually participated in the recording. They were 18, 20 and 20 years old, and from the states of California, Connecticut and New South Wales (Sydney), respectively. The speakers had grown up in a monolingual family. The speakers produced these stimuli in a soundproof room. They read out the materials, including distracters, in random order at least ten times. The utterances were recorded onto a digital recorder (PCM-M10) through a microphone (ECM-MS957) and digitized at 48 kHz with 16 bits. Four tokens per type were selected as stimuli from the recorded materials.

Table 2: The target stimuli.

"Next week I **can / can't**."
    [kæ̃n] / [kæ̃ʔ] : by two American speakers
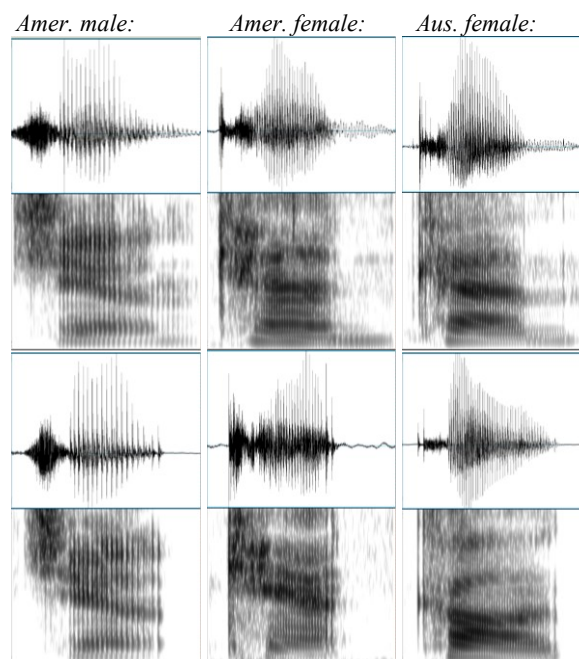   [kæ̃n] / [kɐ̃ʔ] : by an Australian speaker

Figure 1: The waveforms (upper) and spectrograms (lower) of a token of *can* (upper) and *can't* (lower) (duration: 380 ms, view range: 0.0-5000.0 Hz, window length: 0.005 s, dynamic range: 50.0 dB.).

Figure 1 shows the waveforms and spectrograms of one token of the target stimuli displayed by the software program Praat (version 5.4.22) [21]. In order to show the acoustic data, each token of the stimuli was extracted from the last zero-crossing of the end of the vowel (i.e., [aɪ]) preceding "can" or "can't." The durations were aligned to 380 ms. As they are shown in the waveforms and the spectrograms, the common acoustic difference between "can" and "can't" is the existence of the [n] sound; "can" ends with [n], while "can't" ends without [n] both in American and Australian English. In order to ensure the validity of the stimuli, two male native speakers of American English (20 and 31 years old) under identical conditions to those in the recording, who had not participated in the recording, were asked to identify the stimuli. They identified all the stimuli perfectly.

2.1.2. Listeners and procedure

The listeners were 25 students (19-23 years old) of a university in Japan, and they asserted that their native language was Japanese, and that they were not bilinguals according to the questionnaire. The test was conducted in a Computer-Assisted Language Learning (CALL) classroom in the university in Japan. Praat [21] was used as an interface. In front of the computer, the listeners completed a forced-choice identification task while wearing headphones (CZ530-A). The stimuli were presented in random order to each listener over headphones from the computer. In each trial, the participant listened to a stimulus (e.g., "Next week I can.") and judged which button on the screen indicated the most similar Japanese transcription for the word that they had heard after "Next week I." Ten choices written in Japanese *katakana* syllabaries were presented on the screen of the computer: "カン /kaN/," "キャン /kjaN/," "クン /kuN/," "ケン /keN/," "コン /koN/," "カーン /ka:N/," "キャーン /kja:N/," "クーン /ku:N/," "ケーン /ke:N/" and "コーン /ko:N/." The test included 72 trials (2 sentences × 4 tokens × 3 speakers × 3 repetitions). The listeners were allowed to make a correction to their response before clicking the *Next* button to proceed to the next trial. Once they clicked the *Next* button, they were not allowed to return to the previous trial. Prior to the test, the volume was adjusted to a comfortable listening level.

2.1.3. Results of the identification test

Figure 2 shows the perceptual assimilation proportion obtained by the identification test. In order to avoid complexity, the Japanese vowel length contrasts (e.g., /kaN/ vs. /ka:N/) were not considered in the graph. The assimilation patterns of American English "can" and "can't" are almost identical (/kja(:)N/ or /ke(:)N/), while those of the Australian contrast are different (almost perfectly assimilated to /ka(:)N/, or not). The listeners chose /kja(:)N/ (83% vs. 89%) and /ke(:)N/ (15% vs. 11%) for "can [kæn]" and "can't [kæʔ]" produced by the American male speaker. They chose /ke(:)N/ (50% vs. 39%) and /kja(:)N/ (39% vs. 55%) for "can [kæn]" and "can't [kæʔ]" produced by the American female speaker. Almost all the stimuli of Australian "can't [kɐ̃ʔ]" were assimilated to /ka(:)N/ (99%), while the listeners chose /ka(:)N/ (51%), /kja(:)N/ (27%) and /ke(:)N/ (21%) for Australian "can [kæn]."
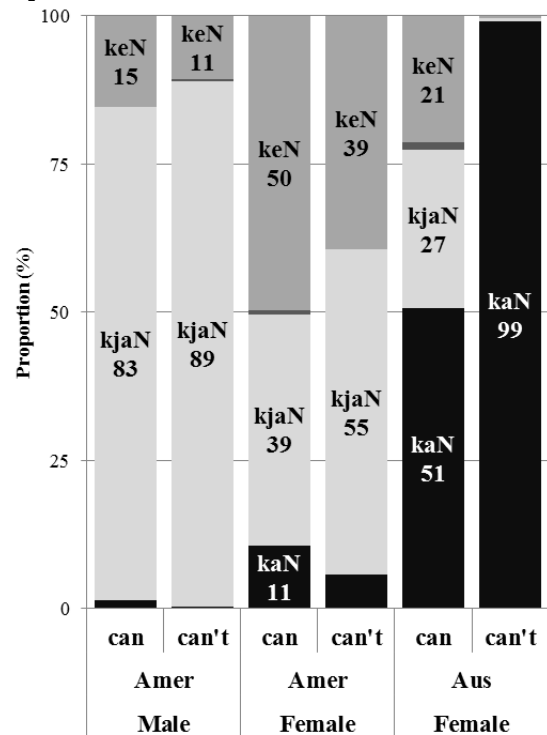

Figure 2: Identification proportion (%).

2.2. Discrimination test

Thirty listeners under identical conditions to the identification test participated in the discrimination test. In each trial of the task, the participant listened to two stimuli (i.e., AA or AB or BB or BA) that were always physically different (i.e., different tokens) but the first stimulus was either the same sentence as the second one, or a different sentence (e.g., *Next week I can-Next week I can* and *Next week I can-Next week I can't*). The listeners judged whether the two stimuli were the same sentences or not, and clicked the *Same* or *Different* button on the screen. When one play was not sufficient to make a choice, they were permitted to play the recording only one more time by clicking the *One more* button.

The test included 56 trials ((8 *can-can't* contrast combinations × 3 speakers) + 32 distracters). The distracters were 24 trials of *should-shouldn't* contrast produced by the three speakers, and 8 trials of *can('t)-should(n't)* contrasts produced by another native speaker of American English. The inter-stimulus intervals were 500 ms.

2.2.1. Results of the discrimination test

Figure 3 shows mean correct rates obtained by the discrimination test. The listeners discriminated the Australian "can [kæ̃n]" vs. "can't [kɛ̃ʔ]" contrast (77.5%) better than the American "can [kæ̃n]" vs. "can't [kæ̃ʔ]" contrasts (62.5% and 63.3%). A one-way repeated-measures ANOVA ($F(2, 58) = 9.887$, $p < .001$) and the multiple comparisons (Bonferroni) showed that there were significant differences in the scores between American male and Australian female speakers' contrasts ($p < .01$), and between American female and Australian female speakers' contrasts ($p < .01$), and that there was no significant difference in the scores between the American speakers' contrasts ($p = 1.000$).
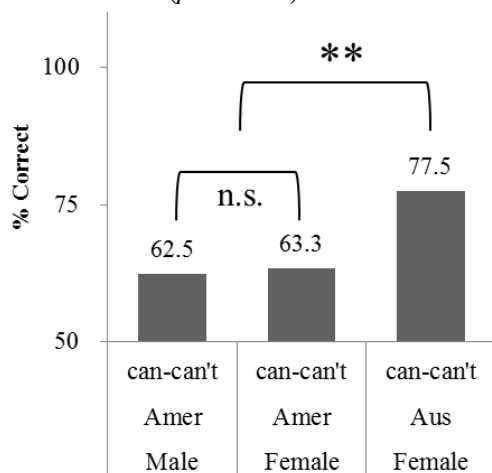

Figure 3: Mean correct discrimination rates (%).

3. Discussion

The results of the present discrimination test showed that the Japanese listeners discriminated Australian English "can [kæ̃n]" and "can't [kɛ̃ʔ]" significantly better than "can [kæ̃n]" and "can't [kæ̃ʔ]" in American English. According to the results of the present identification test, the perceptual assimilation patterns of American English "can [kæ̃n]" and "can't [kæ̃ʔ]" are almost identical: to /kja(:)N/ or /ke(:)N/. In contrast, the listeners mostly assimilated Australian "can't [kɛ̃ʔ]" to /ka(:)N/, while they assimilated Australian "can" to either /ka(:)N/ or /kja(:)N/ or /ke(:)N/. The results suggested that the listeners could detect some vowel quality differences between "can" and "can't" in

Australian English, but not for the American contrast. Therefore, it is assumed that the Australian contrast was easier to discriminate than the American contrast for the Japanese listeners. Table 3 shows the comparisons between the results of the past ([1, 10]) and present studies. The past and present results of the discrimination tests and the American English [kæ̃n] and [kæ̃ʔ] identification tests were consistent with each other. As to the assimilation patterns of Australian English stimuli, the present study showed clearer differences between [kæ̃n] and [kɛ̃:ʔ] than in the past study.

Looking more closely at the identification results, it is interesting to note that the assimilation patterns of [kæ̃] differed among the three speakers. The proportion of /ke(:)N/ was higher in the results of [kæ̃] produced by the American speaker from the East Coast than in the results of [kæ̃] produced by the speaker from the West Coast. The proportion of /ka(:)N/ was higher in the results of [kæ̃] produced by the Australian speaker than the results of [kæ̃] produced by the American speakers. The differences between American and Australian English were found in the past study ([1]), while the differences between the East and West Coasts were unknown due to the lack of the choice /ke(:)N/. The phenomena might be caused by detailed acoustic differences among the accents of the three speakers, which need to be examined in the future.

Table 3: Summary of the results in the past and present studies.

Difficult			***Past***	***Present***
American				
can	[kæ̃n]	→	/kjaN/	/kjaN/ or /keN/
can't	[kæ̃ʔ]	→	/kjaN/	/kjaN/ or /keN/
Easier				
Australian				
can	[kæ̃n]	→	/kaN/	/ka(:)N/, /kja(:)N/…
can't	[kɛ̃:ʔ]	→	/kaN/	/ka(:)N/ only

4. Conclusions

In order to follow up on a past study ([1]), the present research examined Japanese listeners' perception of American and Australian English utterance-final "can" and "can't" (without audible [t] release). According to the results, the discrimination of the Australian contrast (i.e., "can [kæ̃n]" and "can't [kɛ̃ʔ]") was easier than that of the American contrast (i.e., "can [kæ̃n]" and "can't [kæ̃ʔ]") for native speakers of Japanese because of their ability to detect differences in the vowel quality between [kæ̃n] and [kɛ̃ʔ], which is the answer to the research question (see, 1.4).

5. Educational implications

The studies revealed the easiness of the Australian contrast relative to the American one. However, as accurate discrimination of 77.5% was still not perfect (it was 80.8% in the past study ([1])), we *can't* simply say that the Australian contrast is easy for Japanese listeners. As was indicated in 1.2, at least, in RP, American and Australian English, the most important acoustic difference between "can" and "can't" is the existence of [n]. In order to be able to discriminate the contrast more accurately, Japanese EFL learners should get used to detecting the existence of [n] after vowels, instead of relying on the existence of [t] and the vowel quality/duration as they have been instructed. The findings of the present study can be applied to pronunciation training methods. However, such training should not be limited to mere pronunciation training that is based on practice, but should also include some phonetic theory by which students can become familiar with the acoustic and articulatory differences. The following instructions can be effective; "Touch the alveolar ridge with your tongue tip at the end of "can;" do not close your mouth and do not move your tongue at the end of "can't;" and after "can't," you should take a short pause before the next sound." We intend to develop an effective teaching/training method for both pronunciation and listening, focusing on [n] in *can/can't*, and to evaluate its effectiveness in the future.

6. Acknowledgements

7. References

[1] Ooigawa, T. 2013. A pilot study on perception of Australian English utterance-final "can" and "can't" by Japanese listeners. *Sophia University Working Papers in Phonetics*, 17-28.

[2] Takahashi, K., Ooigawa, T. 2012. Perception of American English "can" and "can't" by Japanese professional interpreters. *Interpreting and Translation Studies* 12, 249-262.

[3] Collins, B. S., Mees, I. M. 2008. *Practical Phonetics and Phonology: A Resource Book for Students* (2nd ed.). New York: Routledge.

[4] Wells, J. C., (Ed.). 2008. *Longman Pronunciation Dictionary* (3rd ed.). Edinburgh, UK: Pearson Education.

[5] Cox, F. 2012. *Australian English Pronunciation and Transcription*. New York: Cambridge University Press.

[6] Cruttenden, A. 2008. *Gimson's Pronunciation of English* (7th ed.). London, UK: Hodder Education.

[7] Ladefoged, P. 2006. *A Course in Phonetics* (5th ed.). Boston, USA: Thomson Wardsworth.

[8] Sasaki, M. 1993. Acoustic differences between 'can' and 'can't' in American English [written in Japanese]. *Sophia University Working Papers in Phonetics*, 3-10.

[9] Matsui, J.-k. 1998. The production and perception of the 'deleted' [t] in English. *Annual Review of English Language Education in Japan (ARELE)* 9, 97-106.

[10] Ooigawa, T. 2014. Perception of American English Utterance-Final "can" and "can't" by Japanese Listeners. *Sophia Linguistica* 62, 31-43.

[11] Malécot, A. 1960. Vowel Nasality as a Distinctive Feature in American English. *Language* 36(2), 222-229.

[12] Shockey, L. 2003. *Sound Patterns of Spoken English*. Oxford, UK: Blackwell.

[13] Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

[14] Takahashi, K., Ooigawa, T. 2011. A pilot study on perception of American English "can" and "can't" by Japanese professional interpreters. *Sophia University Working Papers in Phonetics*, 37-52.

[15] Stevens, K. N., Keyser, S. J. 2010. Quantal theory, enhancement and overlap. *J. of Phonetics* 38(1), 10-19.

[16] Tollfree, L. 2001. Variation and change in Australian English consonants. In: Blair, D., P. Collins (eds), *English in Australia*. Amsterdam/Philadelphia: John Benjamins, 45-67.

[17] Nakayama, K. 2000. Perceptual considerations on the contracted auxiliary verb can't in American English [written in Japanese]. *English Pronunciation and Transcription Association Biannual Papers (EPTA Bulletin)* 2, 51-59.

[18] Nakano, K. 1972. *The English Vowels: as contrasted with the Japanese Vowels [written in Japanese]*. Tokyo: Gakushobo.

[19] Sasaki, M. 1995. Perception of unreleased *t* in word-final position in American English. *Unpublished MA thesis, Sophia University*.

[20] Sasaki, M. 1995. Perception of unreleased *t* in word-final position in American English. *Proc. of SLS* Sophia Univ., 70-104.

[21] Boersma, P., Weenink, D. 2015. Praat: doing phonetics by computer [Computer program] (Version 5.4.22).

---

[1] In [1], the vowel in Australian English "can't" was transcribed into [ɑ], but in this paper, it was transcribed into [ɐ] according to [5].

[2] We did not use the symbol [ː], because we did not regard the vowel length distinction as a phonetic-level distinction.

[3] In order to avoid confusion, transcriptions of phonetic symbols were rendered as follows; As vowels were nasalized in syllables closed by a nasal consonant, all the vowels followed by a nasal consonant were transcribed with the diacritic [˜] such as in [kæ̃n] and [kɔ̃n], even if the diacritic was not used in the references. Also, the aspirations ([ʰ]) of [k] and [t] are ignored because they were not regarded as an important acoustic cue.

[4] The distinctions among [ʔt̚], [̚] and the other possible variants are ignored and these allophones were transcribed as [ʔ], because the acoustic differences between them were unknown by the authors.

[5] The phenomenon reminds us of the acoustic cue to discriminate English /l/ and /r/ (i.e., F3). The detailed acoustic/articulatory differences of English liquids differ across the accents, but the main acoustic cue is shared in many accents.