The Research Bulletin of Jumonji Gakuen College, vol.27, 1996

The Physics of the Speech Sound/t/ Variants of the/t/Sound and their Application to Instruction*

Izumi Nonaka**

(Received 1996.5.9.)

1. Introduction

Listening comprehension ability develops when two related processes (bottom-up processing and 'top-down' processing) are involved in an interactional manner. In botto-m-up processing, listeners convert sounds into meanings, and in "top-down" processing, listeners try to grasp the core meaning by making use of their previous knowledge, their experience, or expectations. In my listening course, I encourage students to develop both types of processing and offer them two different tasks to perform. In helping students develop their bottom-up processing ability, I have been teaching phonetic structure, phonemes that we don't have in Japanese, sound changes (liaison, assimilation, reduction), and intonation. Based on several years of experience, I have identified the main obstacles to progress in students' bottom-up processing ability. It is the sound/t/. /t/is a tiny little phoneme, but it is by far the most kaleidoscopic. It often happens that owing to the sound/t/, many students perceive certain words as just an incomprehensible chunk for a long time and fail to make sense of entire passages. Top-down processing is also discouraged if there are too many of these incomprehensible chunks. The purpose of this paper is analyze the physics of variants in the sound/t/and to investigate how to apply the knowledge to instruction in listening and speaking.

2. Variants of/t/

2.1. [t̬] + [ɾ]

An unstressed medial/t/becomes a rather different sound, described as a "flap," when it occurs between voiced sounds, usually vowels. It is represented as/t̬/or/ɾ/in the phonetic alphabet, and sounds like/d/or/r/. In pronouncing a flap/t/in a word like "writer," the tongue rises rapidly and flicks past the upper tooth ridge. "Writer" and "rider" sound alike. It seems that many Americans make no distinction between this type

---

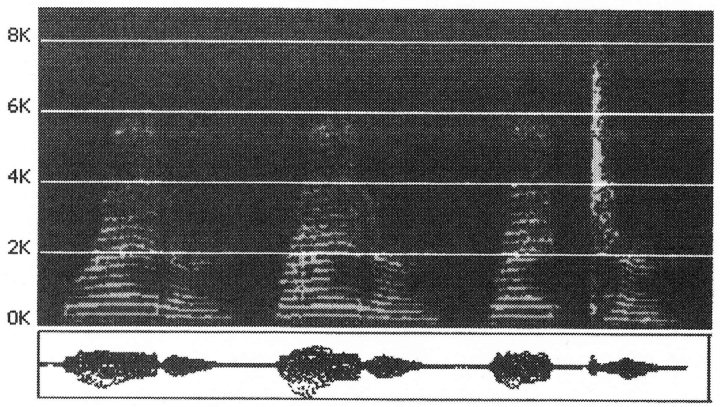
*発話における/t/音の変化に関する考察

**野中泉　文学科英語英文専攻（the Course of English Language & Literature）

of/t/and a/d/.

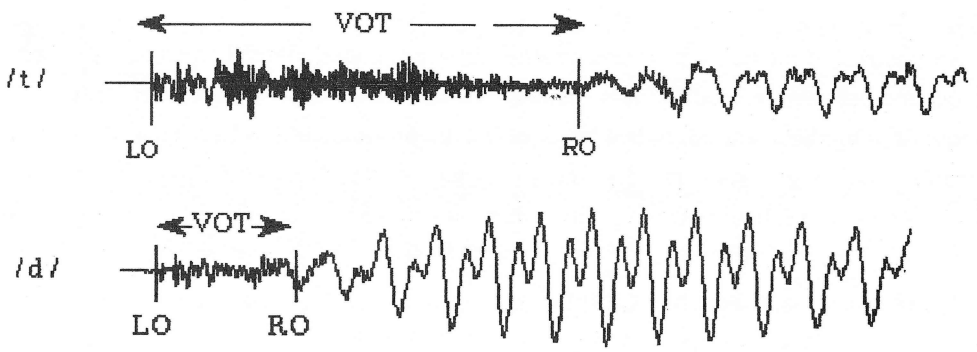
Figure 1 shows spectrograms of the words "writer", "rider", and "writer (British)." The vertical scale indicates the frequencies in Hz, and the horizontal scale shows time in milli-seconds. In the spectrogram, vertical lines that look like human ribs are visible during a vowel. The first two words are spoken by an American woman in her thirties, and the third word by a British woman in her thirties. There is a great deal of similarity between (1) "writer" and (2) "rider." As the/t/in (1) "writer" is flapped, we cannot see a difference between (1) writer/raidər/and (2) rider/raidər/. British people usually don't make the flapped/t/sound, so we can observe the original/t/sound in (3) writer. The acoustic structure of consonants is sucn that a sharp burst of noise appears in the upper frequencies. (3) "Writer" shows a distinctive high-frequency band that cannot be observed in (1) or (2).

Figure 1



What determines the difference between a voiced consonant and a voiceless consonant, in this case, /d/and/t/? P. Lieverman and S. E. Blumstein shows that the time spent on

Figure 2

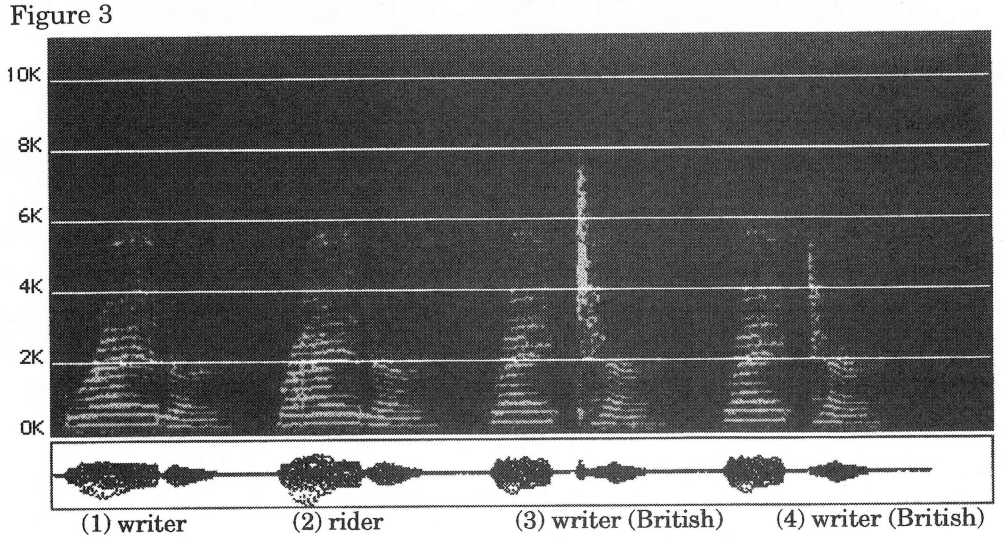


part of a consonant and on a vowel plays an important role. (1988) In pronouncing/t, d/, air pressure is built up while the front part of the tongue is held against the alveolus. When the closure is released, the air escapes forcefully. Figure 2 is an example of the voice onset-time (VOT) of stop consonant/t, d/. VOT is the amount of time between the appearance of a noise burst caused by vocal tract (alveolar) release and the appearance of a periodic waveform caused by vocal cord vibration.

Vocal tract (alveolar) closure is released at the point LO. The vocal cord begins to vibrate at the point RO. Figure 2 shows that the VOT of voiced consonant/d/is much shorter than that of voiceless consonant/t/. The difference in VOT is usually 20–50 msc. Liebermand and Blumstein point out that VOT determines whether a consonant is perceived as/t/or/d/. This study proves scientifically why flap/t/sounds like/d/, i.e., the VOT of rapidly pronounced unstressed/t/is attenuated.

Based on this theory, I conducted an experiment. The fourth spectrogram in Figure 3 is the sound "writer" as spoken by a British woman. It is made from (3) "writer," but the initial part of the onset time (worth 20 msc) was cut by a computer. Thirty-two out of 40 students in my class perceived (4) "writer" as "rider." This result supports the theory that the length of VOT determines whether a word is perceived as voiced or voiceless.

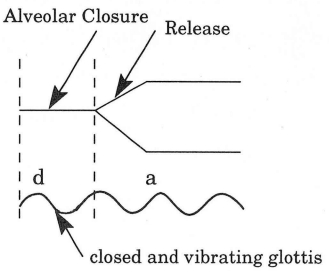
Figure 3



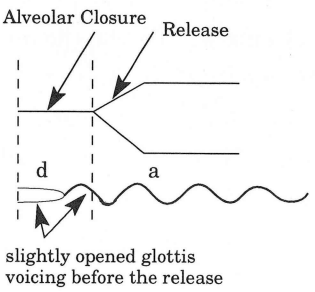
2.2. [t^h]

We saw earlier that the determinant of voiceless or voiced is VOT. The next question is what causes the difference in VOT? Figure 4 shows four different types of time lag between the release of a stop and the onset of voicing in pronouncing an initial stop

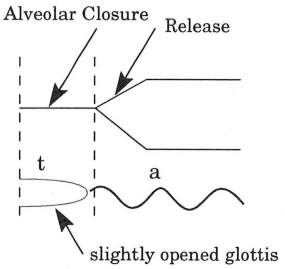
Figure 4 (Revised version of Catford's Chart)



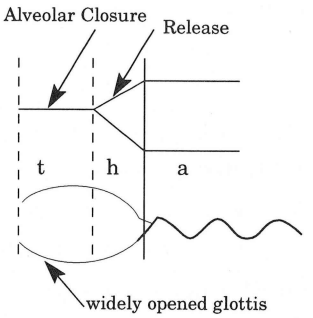
(1) Fully Voiced Stop /d/



(2) Slightly Voiced Stop /d/



(3) Voiceless Unaspirated Stop /t/



(4) Aspirated Stop /t/

consonant preceding a vowel.

In making a voiced sound, the vocal cords are held close together so that they will vibrate, i.e., the glottis (the space between the vocal cords) is closed. When we breathe or make a voiceless sound, the glottis is open, as (3) and (4) of Figure 4 shows. The difference between (3) and (4) is the width of the glottis. In (4), the glottis is wide open. It takes some time for the wide-open glottis to become narrow enough to vibrate the vocal cords, and hence the short period of voicelessness appears. Until the glottis closes and the vocal cords begin to vibrate, pressured air from the lungs will rush out through the glottis. This voiceless puff of breath is known as an "aspiration." Examples are as follows : pen [p^hen], team [t^hi : m], kiss [k^hɪs], appear [əp^hɪr], attend [at^hɛnd], akin [ək^hɪn].

Figure 5

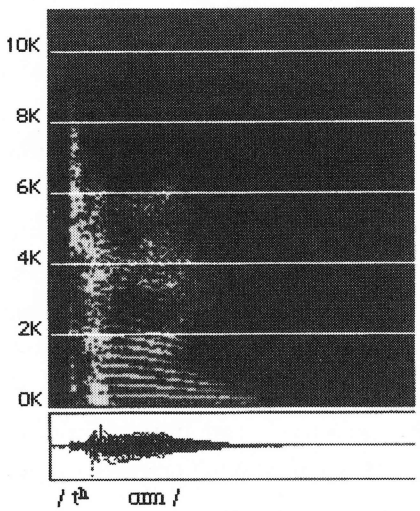


Figure 5 shows a sharp burst appearing in the upper frequencies after/t/in the word "time."

The fact is that Japanese ESL learners are not accustomaed to pronouncing consonants energetically with pulmonic pressure, as they don't have to do so in their native language. If they don't learn to pronounce stessed stop consonants with a strong puff of breath, their glottis will not open wide enough to generate aspiration. If initial voiceless stops/p. t. k/are pronounced without aspiration, they are likely to be perceived as voiced/b, d, g/.

2.3. Syllabic Consonants

Every syllable has a nucleus : Althongh it is most commonly a vowel, yet words like "cot・ton," "would・n't," and "cer・tain" have no vowel in their final syllable. These are known as syllabic consonants. They occur when a syllable ends in/t/, /d/, or/n/, and the following syllable is an unstressed/l/or/n/. In phonetic transcription, syllabic consonants are represented by a short vertical line below them, as in Clinton/klɪntn̩/, curtain/kərtn̩/, or settle/sɛtl̩/. This pronunciation of/t/is typical in American English and is regarded as formal or correct.

2.3.1/t̩/

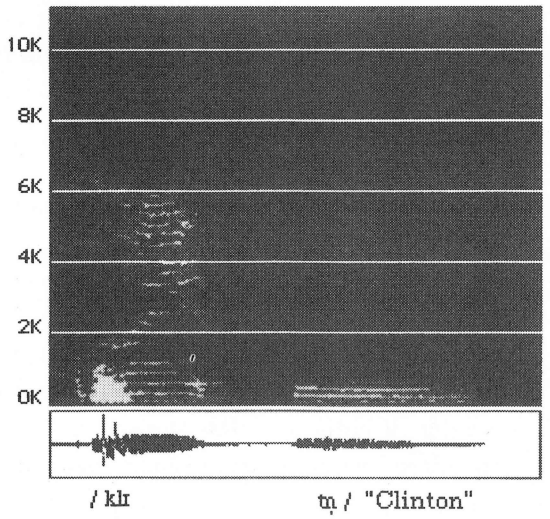
The/t/sound that is difficult for ESL learners to perceive is the voiceless alveolar stop that precedes a syllabic/n/. Both the/t/and/n/are alveolar (homorganic). In making a /t/, the tongue rises to make contact with the tooth ridge. The closure (stopping of the outflow of air from the lung) takes place simultaneoustly. Then the release of the air occurs. Normally, the tongue tip moves away from the tooth ridge when the air is released. However, before a syllabic consonant, the tongue tip remains where it is. As the oral passage remains closed, the air passes out through the nose since/n/is nasal.

Table 1

Clinton	clean	cling to	clean it	clip	cream	click	(others)	ø
8 %	32%	15%	12%	7 %	7 %	3 %	10%	6 %

Eighty students in my class listened to the word "Clinton," and then dictated it. The resuls are shown in Table 1. only 8 % of them answered correctly. Students who

Figure 6



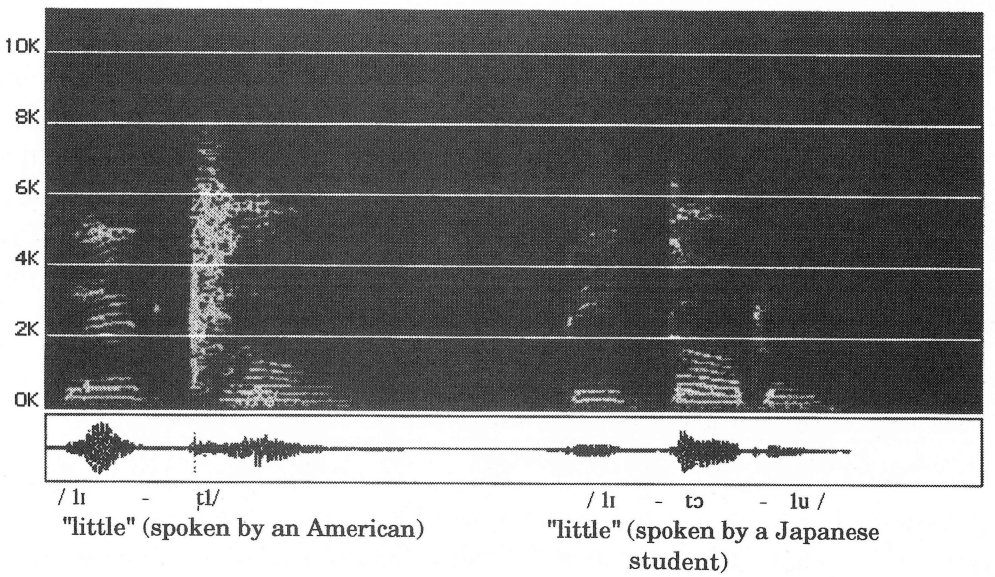
answered "cling to" and "clean it" seemed to perceive the syllabic consonant as two words. This may be because there is a rather long voiceless pause between two syllables in this case, as the spectrogram of "Clinton" (Figure 6) indicates.

Students who dictated "clean," "clip," "cream," or "click" failed to catch/tn̩/ as a syllable. Figure 6 indicates that the syllable of/tn̩/is produced at very low frequencies compared with the /klɪ/syllable. It can thus be inferred that students couldn't perceive the syllabic consonant partly because they don't have syllabic consonants in their mother tongue, and partly because/tn̩/is pronounced at too low a frequency for them.

2.3.2/tl̩/

Another special kind of/t/variant is a/t/followed by a syllabic/l/, which is transcribed as/l̩/. As soon as the tongue rises to contact the tooth ridge to produce the/t/, the release is made by a sudden lowering of the middle and sides of the tongue to make an/l/sound. The blocked airstream rushes out from both sides of the tongue. The tip of the tongue

Figure 7



remains at the tooth ridge. Basically, /tl/is a laterally released/t/.

As Figure 7 shows, an American pronounces/tl/with a rather large, puff-like aspiration. On the other hand, a Japanese ESL learner pronounces it without a puff, and/tl/is pronounced as/tɔlu/using two syllables.

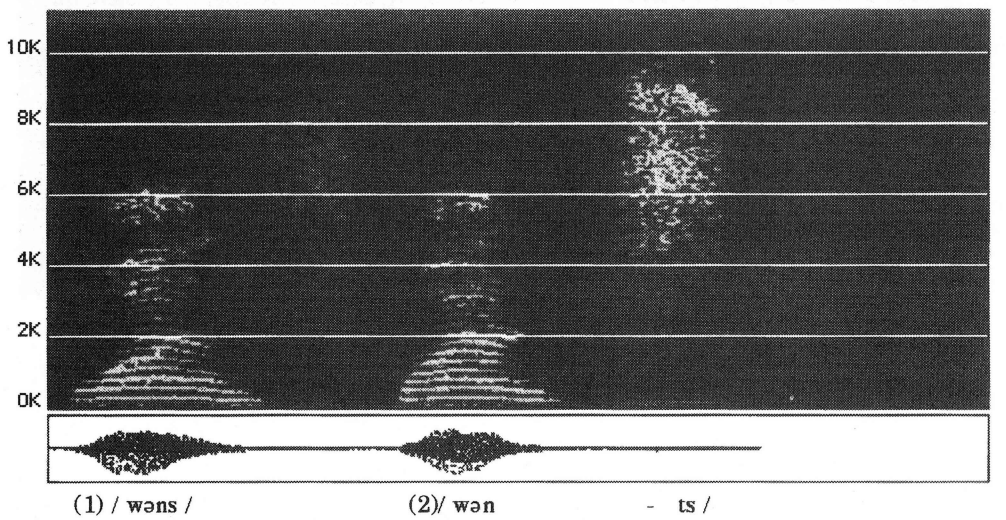
2.4. Intrusive/t/

The intrusive/t/is automatically inserted between two consonants to make one consonant sound transit smoothly into another. Voiced continuant consonants to a voiceless stop or fricative need transition help. Table 2 shows five types of transitions that need an intrusive/t/and examples of their use.

Table 2

Transition	Intrusive/t/	Example
/n/→/s/	/nts/	once/wənts/
/n/→/θ/	/ntθ/	tenth/tɛntθ/
/l/→/s/	/lts/	else/ɛlts/
/l/→/ʃ/	/ltʃ/	welsh/wɛltʃ/
/l/→/θ/	/ltθ/	wealth/wɛltθ/

Figure 8



In Figure 8(1) is the word "once" as spoken by an intermediate Japanese ESL learner, and (2) is the same word as pronounced by a native speaker from America. The/ts/sound

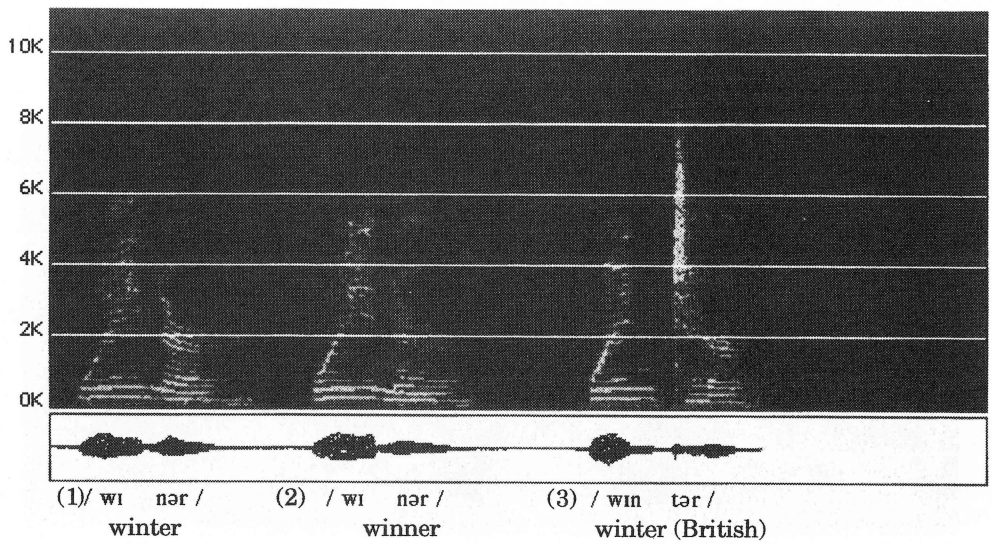
is obvious in (2), but not in (1). It is helpful to advise ESL students to learn to pronounce an intrusive/t/at an early stage of their study since this version of the/t/sound is so common both in British and American English.

25. Flapped/-nt-/

The linked/nt/after a stress is produced as a nasalized flap. The/nt/in/wɪntər/ ("winter") is a nasalized flap, and is pronounced like/wɪnər/. As a result, listeners have difficulty distinguishing "winner" from "winter." This special articulation characterizes American pronunciation. There are many examples that commonly occur : "international," "twenty," plenty," and "Atlanta."

The first two spectrograms of Figure 9 show little difference, but the first one is "winter" and the second one is "winner" as spoken by an American woman. The spectrogram of British woman pronouncing "winter" is shown in (3). We can see the presence of a plosive/t/in (3). The trouble is that this type of sound change also occurs in compound words with the prefix "inter-." ESL students may take "inter-college" for "inner-college," and "inter-culture" for "inner-culture."

Figure 9



2.6. /t̲/ ; Dental/t/

In the case of "eighth"/ɛɪtθ/, the/t/is produced at the same point of articulation of the/θ/that comes right after it. Note that this/t/doesn't explode. Since a/t/coming right

before a/θ/or/ð/is produced between the teeth, it is called a dental/t/, and is transcribed as/t̪/.

3. Implication of this study

As we have seen in this paper, it may be fair to say that the speech sound/t/is the most changeable phoneme depending on its acoustic surroundings. We can infer two major reasons for its changeability ; /t/is a stop consonant whose quality is changed by the length of VOT (amount of pulmonic air), and/t/is alveolar and has the most homorganic phonemes. The first reason would explain the phenomena discussed in 2.1 and 2.2, and the second reason would explain those discussed in 2.3, 2.3.1, 2.3.2, 2.4 and 2.5.

With regard to listening and speaking classes, I would like to propose that teachers place more emphasis on pronouncing consonants with an outflow of air from the lungs, and also make students more sensible to the point of articulation of each consonant. When a consonant cluster is made up of homorganic phonemes, Japanese students are apt to place a vowel after each consonant.

References

Bowen, J. Donald. 1975. Patterns of English Pronunciation. Massachusetts : Newbury House Publishers.

Catford, J. C. 1988. A Practical Introduction to Phonetics. Oxford : Clarendon Press.

Corder, S. Pit. 1981. Error Analysis and Interlanguage. Oxford : Oxford University Press.

Handel, Stephen. 1993. Listening. *An Introduction to the Perception of Auditory Events.* Massachusetts : The MIT Press.

Ladefoged, Peter. 1985. A Course in Phonetics (2 nd ed.). Florida : Harcourt Brace Jovanovich.

Ur, Penny. 1984. Teaching Listening Comprehension. Cambridge : Cambridge University Press.