〔Article〕

Teaching Connected Speech and High School English Education in Japan

KODERA, Masahiro

Introduction

Connected speech is the area on which little research has been done and is not taught in any systematic manner in EFL programs, including high school English education in Japan. No teaching material has been published to show a comprehensive view of connected speech. Yet teaching connected speech is essential for listening comprehension as well as pronunciation. The purpose of this article is to present an overview of connected speech that will be helpful for teaching English pronunciation. First, we will see the important of teaching connected speech in English education, especially for listening comprehension. Next, we will look at the present situation of pronunciation teaching at high school in Japan, where no systematic teaching of pronunciation, let alone that of connected speech, is given. Then an overview of connected speech is presented with special emphasis on the articulatory explanations of linking, elision, and assimilation. Unfamiliarity with these combinatory articulatory phenomena is the main cause that hinders the improvement of L2 learners' listening comprehension of fluent native speaker English.

1. Importance of teaching connected speech

1.1. Connected speech and problems for listening comprehension

Connected speech, also known as sandhi-variation, reduced forms, etc., is the continuous chains in normal conversation, which includes such phenomena as reduction, elision, assimilation, intrusion, juncture, linking, and contraction (Brown 2006:284). It is a combinatory articulatory phenomenon in which words are not pronounced in isolation but run together (Celce-Murcia 2010:163). In connected speech, L2 learners of English often find some words missing where they expect to hear and they try their best, often in vain, to figure out where the word boundaries are in a stream of sounds.

What is most frustrating for L2 learners, especially for those with a good knowledge of English grammar and a sufficient amount of vocabulary, is their inability to decipher native speech. In an EFL environment, language instructors speak clearly and teach English with listening materials that are full of clearly pronounced and articulated speech (Rogerson 2006:85). After having been exposed to the adapted English for a few years, L2 learners are "going to have a very rude awakening when [they try] to understand native speech in natural communicative situations" (Ur 1987:10).

It is important to teach how to pronounce each vowel and consonant, especially those phonemes that the Japanese language does not have, and practice pronouncing each individual word clearly. In natural communicative situations, however, a native speaker does not pronounce in the way L2 learners are taught in the classroom. Their inability to decipher native speech comes from the fact that they develop their listening and speaking skills based on the adapted English speaking style they experience in an EFL class. "Then they arrive at the host country and are shocked and dismayed to find that native speakers don't

actually speak in the way the language is written or in the ways their teachers and listening materials represent the language. The language outside of the classroom seems unfamiliar and fast, and the students are unable to decipher word boundaries or recognize words or phrases"（Rogerson 2006:85-86）.

L2 learners need to be aware of the differences between citation forms（i.e. clear pronunciation of a word when it is stressed or pronounced in isolation, out of context）and modifications in connected speech（i.e. contracted, elided, reduced, assimilated, or linked pronunciation that occur between syllables or words）so that they know what to expect when listening to fluent, native speaker English（Rogerson-Revel 2011:177）. Without knowing the differences, they will have a frustrating experience and may have a serious misunderstanding in communication.

Among various connected speech phenomena, elision of consonants and linking, especially consonant-to-vowel and consonant-to-consonant linking, present the biggest problem for Japanese learners of English since these phenomena occur only in certain circumstances in Japanese. The difference of syllable structure and rhythm causes this difficulty. Elision of consonants occurs in English when consonant clusters（i.e. sequences of two or more consonants within the same syllable）are simplified. Consonant-to-vowel and consonant-to-consonant linking occur at word boundaries when a word ends with a consonant and the following word begins with a vowel or a consonant. These phenomena do not occur in Japanese since the syllable structure of Japanese is basically a sequence of VC（a consonant followed by a vowel）and words do not end with a consonant except for 'yo-on', i.e. contracted sounds like [kja], [kju], [kjo].

'I'll ask her' many sound similar to 'Alaska' [əlæskə] to Japanese learners of English. 'I'll' is a contracted form of 'I will', and both 'I' and 'will' are function words and unstressed. The word-final [l] of 'I'll' links to the initial [æ] of 'ask', resulting in [læ]. The pronoun 'her' is a function word and is unstressed. The initial consonant [h] is dropped and 'her' becomes [əɻ], which links to the final consonant [k] of 'ask", becoming just [kəɻ]. As a result, 'I'll ask her' is pronounced [aɪlæskəɻ], which sounds almost identical to 'Alaska' [əlæskə] to the Japanese learners.

Consonant-to-vowel linking causes another problem of resyllabification. When a word ends in a consonant cluster and is followed by a word that begins with a vowel, the word-final consonant is often pronounced as part of the following syllable（Celce-Murcia 2010:167）. For example, 'find out' may sound like 'fine doubt', and learners try in vain to figure out what it means. Recognition of the position of juncture（or word boundary）is very important in listening comprehension since it can cause a difference in meaning and may lead to misunderstanding.

Consonants in complex consonant clusters, particularly [t] and [d] in medial position, are often elided, which makes it very difficult for L2 learners to recognize the phonetic cue to past tense（Rogerson-Revel 2011:176, Cruttenden 2008:304）. For example, 'I walked back', where [t] of [wɔːkt bæk] is dropped, sounds identical to the present tense 'I walk back'. In the same way, L2 learners find no difference between 'They seemed glad' and 'They seem glad', when [d] of [siːmd] is dropped.

Assimilation also influences consonant clusters. A word-final consonant assimilates to the initial consonant of the following word and changes to a homorganic consonant. For example, the word-final [t] of 'that boy' becomes [p], being influenced by the following bilabial consonant [b]. This change of [t] to [p] gives the impression that [t] is dropped altogether.

1.2. Stress-timed English and syllable-timed Japanese

English is a stress-timed language that shows a tendency for unstressed syllables to become weak, whereas Japanese is a syllable-timed language and unstressed vowels tend to keep the quality and quantity found in their stressed counterparts（Roach 2009b: isochrony）. In unstressed syllables, vowels are often

reduced to schwa [ə], and both vowels and consonants may be elided. Reduction and elision occur most typically in function words in English, which have both weak and strong forms of pronunciation.

Function words in English are normally pronounced in their weak forms, and reduced forms of some function words may sound virtually identical. For example, the indefinite article 'a', which is pronounced [ei] in its strong form, the preposition 'of' [ʌv], and the auxiliary 'have' [hæv] may all be reduced to the schwa sound of [ə]. In the same way, pronouns 'him' and 'them' are often reduced to [əm], which makes the two sentences 'Did you see him?' and 'Did you see them?' sound identical, i.e. [didʒəsiːʲəm], and there is no way to distinguish them without context.

Another example of function words is the modal 'can'. The vowel in 'can' is usually reduced to schwa [ə], while it is pronounced with the full vowel [æ] in the negative form 'can't'. The only difference between the strong form of 'can' and the negative form is the [t] phoneme. A voiceless stop [t] is usually not released at the end of a word or it may be replaced by a glottal stop when followed by another consonant. As a result, the following two sentences sound almost identical when pronounced in citation form:

'You can park here.'	vs.	'You can't park here.'
[juː kæn pɑːrk hɪr] (strong)	vs.	[juː kænt˺ pɑːrk hɪr] or [kæn? pɑːrk hɪr]
[juː kən pɑːrk hɪr] (weak)		

The modal 'can' like other function words is pronounced in its weak form [kən] except in certain limited circumstances. The distinction between these two sentences can be made by the difference between the two vowels [æ] and [ə] or by the glottal stop (Cartier 1967:52, Celce-Murcia 2010:375-377). However, if the L2 learner is taught only the strong forms, which is often the case with high school English education in Japan, they cannot tell whether the speaker is encouraging or prohibiting the parking. In natural communicative situations, the pronoun 'you' is reduced to [jə] and 'can' to [kən], which makes the affirmative sentence sound like [jəkənpɑːrkhɪr], while 'can' in the negative sentence is pronounced in its strong form [kænt], but with [t] unreleased ([kænt˺]) or with [t] glottalized ([kæn?]). Without the knowledge of these reduced forms, L2 learners find native speech hard to decipher. Unfortunately, reduced forms in connected speech do not receive serious attention in high school education in Japan.

2. Connected speech in high school English education in Japan

There are two major problems with teaching connected speech: research and teacher training. First, little research has been done on connected speech and language teachers do not find any systematic treatment of connected speech. Brown and Kondo-Brown (2006:6) argues:

> One of the problems with connected speech as a discipline is that it isn't a discipline or even a sub-discipline. True, connected speech has interested some teachers and researchers over the past 30 years, but only a few people have worked systematically on connected speech and those few have done so only sporadically. In the meantime, teachers continue to teach the phonemes of their language of focus and later wonder why their students' pronunciation is still inadequate, that is, why their students cannot put the phonemes together in anything approaching a native-like manner.

Language teachers need a single body of information on connected speech that they can consult when teaching pronunciation. Unfortunately, there is no teacher resource that can claim to be comprehensive on the topic of connected speech. After reviewing nine books that are focused on helping students learn

reduced forms in English, Brown (2006) has found that few teacher-training books cover the topic in depth and that the ESL/EFL student textbooks' coverage of the topic are scattered in and among other information about pronunciation and listening comprehension.

Another issue that makes teaching connected speech difficult is the language teachers' lack of knowledge. Rogerson (2006:89) has found that less than half of the 45 teachers surveyed in Hawaii had taken courses on teaching listening comprehension or had received instruction on the role of reduced forms in teaching ESL listening comprehension. The similar situation can be found in Japan as well. Shibata, et al. (2006) reports that, among 77 junior and 147 senior high school English teachers surveyed, 69.8 percent had taken a phonetics/phonology class and 47.9 percent had received any training to improve their English pronunciation.

What makes matters worse is the fact that English teachers in Japan do not recognize the importance of teaching pronunciation, which makes a sharp contrast with the ESL instructors surveyed in Hawaii, where 42 of the 45 teachers considered reduced forms instruction "somewhat" to "very important" and 43 considered it "very helpful" to "somewhat helpful" to teach reduced forms (Rogerson 2006:90). On the other hand, teaching pronunciation is not considered an important part of English education at high school in Japan. Arimoto (2005) reports that about 30 percent of the English teachers who participated in his English pronunciation training session do not teach pronunciation at all, and many of the high school English teachers who use the English textbook he co-authored found the pronunciation part of the textbook not necessary. He also says that English teachers do not give any systematic description of phonetic alphabets and that they just play a CD to show English native speaker's pronunciation and have their students repeat it. With this way of teaching, Arimoto argues, students never learn to correct their pronunciation.

Still another issue that makes teaching pronunciation and connected speech difficult is the English textbooks used in high schools. Ueda and Otsuka (2010) has checked six English textbooks used in junior high schools to see how pronunciation teaching is treated in the textbooks and has found that they all provide some coverage of phonetically related items, such as phonetic alphabets, stress, intonation, connected speech, etc., but none gives a systematic description of connected speech.

3. An overview of connected speech phenomena

Japanese learners of English learn words in citation form but not in reduced form. In actual conversation, however, they most likely hear words in reduced form and have hard time figuring out what a native speaker is speaking to them. They often find some sounds missing in native speech, which results from various connected speech phenomena, such as reduction, elision, linking, and assimilation. These phenomena can be explained from two aspects of English pronunciation: rhythm and syllable structure.

3.1. Stress-timed rhythm, reduction, and elision

English has an unusually high number of weak forms and contracted forms (Rogerson-Revel 2011:178). When a syllable in English is unstressed, it is often pronounced differently from when stressed. Vowels are centralized and become more schwa-like. This phenomenon of vowel reduction is said to come from the difference between "stress-timed" and "syllable-timed" languages (Roach 2009b: reduction). According to the stress-timed rhythm hypothesis, English is a stress-timed language, which has a rhythm that allows us to divide its speech up into more or less equal intervals of time called feet, each of which begins with a stressed syllable (Roach 2009b: rhythm).

Stressed timed languages tend to give each stress group approximately the same weight, while syllable

timed languages (e.g. Japanese, Spanish) tend to give each syllable approximately the same prominence and the length of each syllable remains more or less the same as that of its neighboring syllables whether or not it is stressed. In English, the more syllables a stressed segment has, the less time each syllable spares for its pronunciation, resulting in reduced forms. In the following examples, each sentence is divided up to three stress groups, each of which is given about the equal time to pronounce (Table 1).

Table 1

•		•	•
CATS		CHASE	MICE
THE CATS	have	CHASEd	MICE
THE CATS	will	CHASE	the MICE
THE CATS	have been	CHASing	the MICE
THE CATS	could have been	CHASing	the MICE

(cited from Celce-Murcia 2010:210)

Each sentence has three stressed elements, but each stressed element has a different number of syllables. For instance, the first stressed element of the first sentence has just one syllable while that of the last sentence has as many as five syllables. Generally, content words are stressed and function words are unstressed, which makes 'could have been' in the last sentence pronounced very quickly, resulting in the reduced form of [kədəbɪn] rather than [kʊd hæv bɪn].

Reduction is the processes in which phonemes (vowels or consonants) are changed, minimized, or eliminated in order to facilitate pronunciation. Reduction occurs when a syllable is unstressed and the vowel in the syllable is reduced to schwa [ə]. It most typically occurs to function words, such as articles and auxiliary verbs (Roach 2009: reduction).

The distinction between stressed-timed and syllable-timed languages is not universally accepted (Celce-Murcia 2010: 208). Roach (2009b: rhythm) says: "Most evidence from the study of real speech suggests that such rhythms only exist in very careful, controlled speaking, but it appears from psychological research that listeners' brains tend to hear timing regularities even where there is little or no physical regularity." It is generally agreed among pronunciation researchers that stress-timing represents a strong tendency in English and it should be included in the pronunciation curriculum (Celce-Murcia 2010:208). Unaware of this distinction, Japanese students, whose language is syllable-timed, tend to stress syllables in English equally, without reducing unstressed syllables (Celce-Murcia 2010:210).

3.1.1. Weak forms of unstressed function words

Function (or grammatical) words, as opposed to lexical or content words (e.g. main verbs, nouns, adjectives, adverbs, and demonstratives), are those words whose role is largely or wholly grammatical, such as auxiliary verbs, conjunctions, determiners, prepositions, pronouns, and quantifiers. They have both a strong form (i.e. stressed citation form) and one or more weak forms (i.e. unstressed reduced form). In connected speech, function words are normally unstressed and pronounced in their weak forms. Strong forms are used only in exceptional circumstances: when the word is being quoted (e.g. 'How is the word "for" spelt?'); when it is at the end of a sentence or phrase (e.g. 'What are you looking at?', 'I know I can.'); when it is being contrasted (e.g. 'There are arguments for and against.').

Some function words sound almost identical when reduced in connected speech, which causes problems for L2 learners in listening comprehension. For example, 'Did you see him/them?' are both pronounced [didʒəsiːʲəm]. Table 2 shows some function words that sound alike in reduced speech.

Table 2　Function words that sound alike in reduced speech

him	[əm]	as	[əs]	or	[ər]	a	[ə]	in	[ən][n]
						of		on	
them		is		are		have		an	
								and	

(Cited from Celce-Murcia 2010:377)

Reduction often results in the loss of phonological segments. For instance, 'and' [ænd] is reduced to a weak form [ənd], or [ən] with [d] elided, or a syllabic [n̩] with both [ə] and [d] elided (Giegerich 1992:287). The preposition 'of' [ʌv], when followed by a word that begins with a consonant, may be pronounced just [ə] with the vowel [ʌ] reduced to [ə] and the consonant [v] dropped. As a result, 'cup of tea' is pronounced [kʌpətiː] rather than [kʌp ʌv tiː] (Crystal 2006:158-159). Table 3 shows a list of strong and weak forms of some function words with some segments elided (e.g. 'her' [ər] with [h] dropped, 'them' [əm] with [ð] elided).

Table 3　Strong forms and weak forms of function words

Function Words	Strong Form	Weak Form	
Determiners			
a	[eɪ]	[ə]	
an	[æn]	[ən]	
the	[ðiː]	[ðə]	
some	[sʌm]	[səm]	
Pronouns (Loss of initial [h] and [ð] in pronominal forms)			
he	[hiː]	[hi] [iː] [i]	'Where did he go?' ⇒ [dɪdiː], [diː]
his	[hɪz]	[ɪz]	'They found his car.' [hɪz] ⇒ [ɪz]
him	[hɪm]	[ɪm]	'Tell him.' [hɪm] ⇒ [ɪm]
her	[hər]	[ər]	'I'll ask her.' [hər] ⇒ [ər]
them	[ðem]	[ðəm] [əm]	'I'm waiting for them.' [ðəm] ⇒ [əm]
you	[juː]	[jə] [jʊ]	'What do/are you' ⇒ [wədəjə]
your	[jɔːr]	[jə]	
us	[ʌs]	[əs]	
Copula 'be'			
are	[ɑːr]	[ər]	
am	[æm]	[əm] [m]	
is	[ɪz]	[əz] [z] [s]	
Prepositions			
of	[ʌv]	[əv] [ə]	'cup of tea' ⇒ [kʌpətiː]
to	[tuː]	[tə]	
for	[fɔːr]	[fər]	
from	[frʌm]	[frəm]	
in	[ɪn]	[ən]	
on	[ɑːn]	[ən] [n]	
at	[æt]	[ət]	
as	[æz]	[əz]	
Conjunctions			
and	[ænd]	[ənd] [ən] [n]	'boys and girls' ⇒ boys 'n girls
or	[ɔːr]	[ər]	'soup or salad' ⇒ [suːpər] salad

that	[ðæt]	[ðət] [ət]	
Auxiliary Verbs			
have	[hæv]	[həv] [əv] [v] [ə]	'could have been' ⇒ [kədəbɪn]
has	[hæz]	[həz] [əz] [z]	'Has it' ⇒ [(ə)zɪt]
can	[kæn]	[kən] [kn]	'I can go.' [kæn] ⇒ I [kən]/[kn] go.
could	[kʊd]	[kəd] [kd]	
does	[dʌz]	[dəz] [z]	'Does it work?' ⇒ [zɪt] work?
should	[ʃʊd]	[ʃəd]	

3.1.2. Elision

Elision is the process of elimination or dropping of phonemes (both vowels and consonants) or whole syllables that would be present in the citation form of a word or phrase (Brown 2006:284). It is most often found in unstressed function words (see 3.1.1), unstressed syllables in polysyllabic words (e.g. 'philosophy' [fəlɑsəfi] ⇒ [fəlɑsfi]), and complex consonant clusters (e.g. 'twelfths' [twelfθs] ⇒ [twelθs] or [twelfs]). Elision may occur word-initially (aphesis: e.g. 'again' [ɑgen] ⇒ [gen]), word-medially (syncope: e.g. 'interesting' [ɪntərestɪŋ] ⇒ [ɪntrɪstɪŋ]) or word-finally (apocope: e.g. 'of' [ʌv] ⇒ [ə]).

3.1.3. Elision of vowels

Vowels in unstressed syllables are often reduced to schwa [ə] or incorporated into a syllabic consonant. Elision of vowels usually happens when a short, unstressed vowel occurs between voiceless consonants (e.g. potato [pəteɪtoʊ]⇒ [pteɪtoʊ]) (Roach 2009b: elision). It also happens when a vowel occurs between an obstruent consonant and a sonorant consonant (e.g. nasal, lateral), resulting in syllabic consonants such as syllabic nasals (e.g. 'sudden' [sʌdən] ⇒ [sʌdn]) and syllabic [l] (e.g. 'bicycle' [baɪsɪkəl] ⇒ [baɪsɪkl]).

A vowel between consonants in a weak syllable is not pronounced when it follows a strongly stressed syllable in multisyllabic words (e.g. 'camera', 'philosophy'), or when the preceding consonant is a voiceless stop [p], [t] or [k] (e.g. 'police', 'tomato', 'correct'), or when the following consonant is [n], [l] or [r], resulting in syllabic consonants (e.g. 'tonight', 'police', 'history') (Celce-Murcia 2010:172, Roach 2009:113-114, Giegerich 1992:287, Rogerson-Revell 2011:166) (Table 4). Whole syllables may be elided, when the syllables are unstressed just after or before a stressed syllable when the elided syllable contains a consonant that is repeated in the following syllable, especially [r] (e.g. 'library') (Crystal 2006:158, Skandera 2011:96) (Table 5). The initial unstressed vowel or syllable can also be dropped (e.g. 'about', 'because') (Hagen1988:5, Celce-Murcia 2010:173, Crystal 2006:28) (Table 6).

Table 4 Loss of an unstressed vowel between consonants

1st syllable	'police' [pəliːs] ⇒ [pliːs] 'garage' [gərɑːʒ] ⇒ [grɑːʒ]	'potato' [pəteɪtoʊ] ⇒ [pteɪtoʊ] 'correct' [kərekt] ⇒ [krekt]
2nd syllable	'camera' [kæmərə] ⇒ [kæmrə] 'probably' [prɑbəbli] ⇒ [prɑbli]	'bicycle' [baɪsəkl] ⇒ [baɪskl] 'history' [hɪstəri] ⇒ [hɪstri]
3rd syllable	'philosophy' [fəlɑsəfi] ⇒ [fəlɑsfi]	

Table 5 Loss of a whole syllable between a repeated consonant (haplology)

'February' [februerɪ] ⇒ [febrɪ]	'library' [laibreri] ⇒ [laibriː]
'probably' [prɑːbəbli] ⇒ [prɑːbli]	'particularly' [pətikjuləli] ⇒ [pərtikjuli]

Table 6　Loss of an initial unstressed vowel or syllable (aphesis)

'about' [əbaʊt] ⇒ 'bout	'another' [ənʌðəɹ] ⇒ 'nother
'around' [əraʊnd] ⇒ 'round	'because' [bɪkɔ:z] ⇒ 'cause
'exactly' [ɪgzæktli] ⇒ 'xactly	'remember' [rɪmembəɹ] ⇒ 'member

Elision of vowels can be found in syllabic consonants (Roach 2009b: syllabic consonant). A syllable in English contains a central vowel, but a vowel may be elided in natural rapid speech and is replaced by a syllabic consonant. For example, 'student' is often heard as [stju:dnt] rather than [stju:dənt] with the weak vowel [ə] being replaced by a syllabic consonant [n̩]. Syllabic consonants regularly occur when a syllable final alveolar [t], [d] or [n] is followed by an unstressed homorganic consonant [n] or [l] (e.g. 'mountain' [maʊntn], 'total' [toʊtl]). In addition to [n] and [l], [m], [ŋ], and [r] can also occur as syllabic consonants (e.g. 'open' [oʊpm], 'broken' [broʊkŋ], 'history' [hɪstri]).

Syllabic consonants occur under the following conditions: 1) when an alveolar [t] or [d] is followed by a homorganic consonant [n]; 2) when a bilabial [p] or [b] is followed by a homorganic consonant [m]; 3) when a velar [k] or [g] is followed by a homorganic [ŋ]; 4) when a fricative [f] or [v] is followed by [n]; 5) when a consonant is followed by a lateral [l]; 6) when [t] is followed by [r]. The syllable that ends with a nasal ([m, n, ŋ]) has a nasal release and that with a lateral [l] has a lateral release (Celce-Murcia 2010:94, Rogerson-Revell 2011:124, Cruttenden 2008:167) (Table 7). A final syllabic nasal is less frequent and normally have [ə] in the last syllable when [s] precedes (e.g. 'Boston'), and it is unusual when it follows a consonant sequence of a nasal and a stop precedes (e.g. 'London', 'abandon') (Roach 2009:70).

Some syllabic consonants are practically obligatory; 'bottle' is pronounced [bɑtl] not [bɑtəl], and 'button' [bʌtn] not [bʌtən] (Roach 2009b: syllabic consonant). Many can be pronounced either with or without a schwa. For example, 'open' can be pronounced [oʊpən] or [oʊpm]. There are some regional differences about inserting a vowel before [l]; 'bicycle' may be pronounced [baɪsɪkəl] or [baɪsɪkl] and 'total' may be [toʊtəl] or [toʊtl] (Roach 2009:68-69). It would be advisable to teach L2 learners to pronounce them without a vowel.

Table 7　Syllabic consonants and elision of vowels

Obstruent Consonant + Unstressed Vowel + Sonorant Consonant				
Nasal Release: Vowel + Homorganic Stop [t][d][p][b][k][g] + [ə] +Nasal [m][n][ŋ]				
Alveolar				
[t]	[ə]	[n]	'written' [rɪtnən] ⇒ [rɪtn]	'mountain' [maʊntən] ⇒ [maʊntn]
			'gotten' [gɑtən] ⇒ [gɑtn]	'curtain' [kətən] ⇒ [kətn]　'certainly'
[d]			'sudden' [sʌdən] ⇒ [sʌdn]	'garden' [gɑ:rdən] ⇒ [gɑ:rdn]
			'wooden' [wʊdən] ⇒ [wʊdn]	'burden' [bəɹdən] ⇒ [bəɹdn]
Bilabia				
[p]	[ə]	[m]	'open' [oʊpən] ⇒ [oʊpn] ⇒ (assimilation) ⇒ [oʊpm]	
			'happen' [hæpən] ⇒ [hæpn] ⇒ (assimilation) ⇒ [hæpm]	
[b]			'ribbon' [rɪbən] ⇒ [rɪbn] ⇒ (assimilation) ⇒ [rɪbm]	
Velar				
[k]	[ə]	[ŋ]	'broken' [broʊkən] ⇒ [broʊkŋ]	'bacon' [beɪkən] ⇒ [beɪkŋ]
[g]			'slogan' [sloʊgən] ⇒ [sloʊgŋ]	

Fricative				
[f]	[ə]	[n]	'often' [ɔfən] ⇒ [ɔfn]	'deafen' [defən] ⇒ [defn]
[v]			'seven' , [sevən] ⇒ [sevn]	'heaven' [hevən] ⇒ [hevn]
Lateral Release: Consonant + Unstressed Vowel + [l]				
[t]	[ə]	[l]	'total' [toʊtəl] ⇒ [toətl] 'lentil' [lentəl] ⇒ [lentl]	'bottle' ⇒ [bɑtˡl], battle, little' (obligatory)
[d]			'orderly' [ɔrdəˌli] ⇒ [ɔrdli] 'pedal' [pedəl] ⇒ [pedl]	'badly' [bædlli] (obligatory)
[p]			'chapel' [tʃæpəl] ⇒ [tʃæpl]	'couple' [kʌpəl] ⇒ [kʌpl]
[b]			'label' [leɪbəl] ⇒ [leɪbl]	'trouble' [trʌbəl] ⇒ [trʌbl]
[k]			'technical' [teknɪkəl] ⇒ [teknɪkl]	'nickel' [nɪkəl] ⇒ [nɪkl]
[g]			'triangle' [traɪæŋgəl] ⇒ [traɪæŋgl]	'struggle' [strʌgəl] ⇒ [strʌgl]
[n]			'kernel' [kəˌnəl] ⇒ [kəˌnl]	'tunnel' [tʌnəl] ⇒ [tʌnl]
[s]			'parcel' [pɑːrsəl] ⇒ [pɑːrsl]	'cancel' [kænsəl] ⇒ [kænsl]
[z]			'chizel' [tʃɪzəl] ⇒ [tʃɪzl]	'diesel' [diːzəl] ⇒ [diːzl]
[f]			'beautiful' [bjuːtəfəl] ⇒ [bjuːtəfl]	'awful' [ɔːfəl] ⇒ [ɔːfl]
Others:				
[t]		[r]	'history' [hɪstəri] ⇒ [hɪstri]	

3.2. Linking

3.2.1. Syllable structure: closed syllable and open syllable

Every language has syllables (Roach 2009b: syllable), and a consonant+vowel (CV) sequence is the simplest and most universal syllable structure, which seems to be found in all languages (Crystal 2006:448). The syllable with the CV sequence pattern is not closed by another consonant and is called an open syllable type, while a syllable that ends with a consonant is a closed syllable type. Japanese is open-syllable and its syllable structure is basically the CV sequence, except for the structure that ends with [N]. The consonant-to-vowel and the consonant-to-consonant linking patterns across word boundaries never occur in Japanese since the word-final [N] does not link to the initial vowel of the following word. Japanese has only the VV and VC sequence patterns across word boundaries.

Linking is the connecting of the final sound of a word or syllable to the initial sound of the next (Celce-Murcia 2010:166). It occurs between vowels and vowels (e.g. blue ink), consonants and vowels (e.g. skip it), and consonants and consonants (e.g. top person, hot cake) (Brown 2006:285). Japanese is open-syllable, and the vowel-to-vowel linking should not cause a serious problem for Japanese learners of English. On the other hand, they find the consonant-to-vowel and the consonant-to-consonant linking very difficult in listening comprehension and pronunciation, since no word in Japanese ends with a consonant except for a uvular nasal [N]. For example, most Japanese learners of English are unaware of the distinction between the uvular nasal [N] and the alveolar nasal [n] and pronounce 'an apple' as [əN æpl] rather than [ənæpl] without linking the word-final consonant [n] and the initial vowel [æ] of the next word.

3.2.2. Consonant clusters, cluster reduction, and resyllabification

English has a wide configuration of consonant clusters, and a word or syllable can end with either a vowel or a consonant. English syllables can take various shapes: one vowel, V ('eye' [aɪ]); a vowel with up to three initial consonants, CV ('pie' [paɪ]), CCV ('spy' [spaɪ]), CCCV ('spray' [spreɪ]); a vowel with

up to three final consonants, VC (‘at’ [æt]), VCC (‘ask’ [æsk]), VCCC (‘asked’ [æskt]); a vowel with one or more initial consonants and up to four final consonants CVC (‘ten’ [ten]), CVCC (‘tent’ [tent]), CVCCC (‘tempt’ [tempt]), CVCCCC (‘tempts’ [tempts]); a vowel with almost the full range of possible initial and final clusters, CCVCC (‘trust’ [trʌst]), CCCVCCC (‘splints’ [splɪnts]) (Celce-Muricia 2010: 103).

Consonant clustering of English presents a great difficulty to Japanese learners of English (Celce-Murcia 2010:98-100), since consonant clusters do not occur in Japanese with only two exceptions: 1) the double consonant cluster (or geminate consonant) (soku-on 促音 in Japanese) with two plosives like [tt], [kk], [pp] as in ‘gakko’ (meaning ‘school’) or that with two fricatives like [ss], [ʃʃ] as in ‘assari’ (‘simply’, ‘easily’) and ‘isshou-kenmei’ (‘very hard’, ‘desperately’), which makes two morae in Japanese; 2) the contracted sound (yo-on 拗音) such as [kja], [kju], [kjo] as in ‘kyu’ (‘nine’ pronounced [kjuː] similar to ‘queue’ in English). The double consonant cluster does not occur syllable-finally.

Native English speakers try to make consonant clusters easier to pronounce, and one of their strategies is cluster reduction, with which one of the consonants, usually a middle consonant in the final cluster of three or four consonants, is dropped. For example, ‘asked’ [æskt] is simplified to [æst] or ‘facts’ [fækts] to [fæks] (Celce-Muricia 2010:100). Consonant clusters also occur across word boundaries, and assimilation is one strategy to make the pronunciation easier. Geminate (or two identical) consonants are usually pronounced as one long consonant (e.g. stop playing).

In the sequence of two homorganic consonants, the first consonant has only a closing stage and the second consonant has only a release stage (e.g. ‘hot dog’). The first consonant is not released or it may be replaced by a glottal stop (e.g. ‘great smile’ [ɡreɪʔ smaɪl]). In other cases of assimilation, the first consonant takes the characteristics of the following consonant (e.g. ‘that boy’ [ðæp bɔɪ]) or the two consonants fuse into a new segment (e.g. ‘need you’ [niːdʒʊ]).

Still another strategy is resyllabification, i.e. simplification of final consonant clusters at word boundaries. When a word or syllable terminating in a consonant cluster is followed by a word or syllable commencing with a vowel, the final consonant of the cluster is often pronounced as part of the following syllable. For instance, ‘moved it’ is pronounced [muːv dɪt] rather than [muːvd ɪt]. This applies not only to the CC+V sequence (e.g. ‘find out’ [faɪn daʊt]), but also to the VC+V sequence (e.g. ‘keep out’ [kiː paʊt], ‘dream on’ [driː mɔːn]) (Rogerson-Revel 2011:169, Celce-Murcia 2010:166-167, Avery and Ehrlich 1992:85).

Resyllabification makes it difficult for L2 learners to find the word boundary. When they hear the sound [kiːpstɪkɪŋ], for example, it would be almost impossible to figure out whether it means ‘keep sticking’ or ‘keeps ticking’ without context. The only phonetic cue to tell the difference is whether [t] is aspirated or not.

3.2.3. Vowel-to-vowel linking

In vowel-to-vowel linking, a junctural [ʲ] glide or [ʷ] glide is often inserted when a word or syllable ends in a tense vowel or diphthong and the next word or syllable begins with a vowel (Celce-Murcia 2010:165, Cruttenden 2008: 227, 230, Cook 2000: 63-64). A junctural [ʲ] glide is inserted when a vowel follows [iː], [eɪ], [aɪ] or [ɔɪ] (e.g. seeing [siːʲɪŋ]), and a junctural [ʷ] glide is inserted when a vowel follows [uː], [oʊ] or [aʊ] (e.g. ‘blue ink’ [bluːʷɪŋk]) (Table 8).

The junctual [ʲ] glide is different from the phoneme [j] in that the finishing point of the diphthong is not sufficiently prominent and the glide is not long enough as can be seen in the opposition between ‘my ear’ [maɪ ʲɪə] and ‘my year’ [maɪ jɪə] (Cruttenden 2008:227). In the same way, the difference between the junctural [ʷ] glide and the phoneme [w] can be seen in the opposition between ‘two-eyed’ [tuːʷaɪd] and ‘too wide’ [tuː waɪd] (Cruttenden 2008:230).

Table 8 Word-final vowel + junctural [ʷ] [ʲ] glide + word-initial vowel

Junctural [ʲ] glide following [iː] [eɪ] [aɪ] [ɔɪ]			
[iː] + V	[eɪ] + V	[aɪ] + V	[ɔɪ] + V
seeing [siːʲɪŋ]	saying [seɪːʲɪŋ]	crying [kraɪʲɪŋ]	toying [tɔɪʲɪŋ]
he isn't	say all	try out	toy airplane

Junctural [ʷ] glide following [uː] [oʊ] [aʊ]		
[uː]+ V	[oʊ]+ V	[aʊ]+ V
doing [duːʷɪŋ]	going [goʊʷɪŋ]	however [haʊʷevəɹ]
do it	go away	how is it

3.2.4. Consonant-to-vowel Linking

Japanese learners of English experience great difficulty finding the position of a word boundary without the knowledge of consonant-to-vowel linking and resyllabification. Two or more words linked together in connected speech may sound like one word. For example, they may hear 'Alaska' when they are expected to hear 'I'll ask her'. They might hear 'supersalad' and wonder what kind of salad it would be when they are actually asked if they prefer 'soup or salad.' Resyllabification causes a serious problem as well. When they hear [maɪtəɹn] for instance, they need to figure out whether it means 'my turn' or 'might earn.' Table 9 shows various combinations of a consonant and a vowel at word boundaries.

Table 9 Consonant-to-vowel linking

Labial + V		Dental + V		Alveolar + V		Palato-alveolar + V		Velar + V	
[p]	stop it	[θ]	beneath it	[t]	washed it	[ʃ]	cash out	[k]	back out
[b]	grab it	[ð]	with it	[d]	depend on	[ʒ]	camouflage it	[g]	drag out
[m]	come in		breathe it	[s]	pass out	[tʃ]	march in	[ŋ]	sing it
[f]	laugh about			[z]	Does it?	[dʒ]	damage it		
[v]	leave early			[n]	run around				
				[l]	fool around				
				[r]	fair enough				

(Adapted from Avery, P. and S. Ehrlich 1992:85)

3.2.5. Consonant-to-consonant linking

3.2.5.1. Geminate consonants linking

Two identical (geminate) consonants across word boundaries are usually pronounced as one long consonant (Table 10). For example, [p] in 'stop playing' is not pronounced twice; rather two [p]s are pronounced [pː] (Celce-Murcia 2010:167, Cruttenden2008:166).

Table 10 Geminate consonants linking

[p][p]	⇒ [pː]	stop playing	[b][b]	⇒ [bː]	Bob brushed his teeth.
[t][t]	⇒ [tː]	hot tea	[d][d]	⇒ [dː]	red dress
[k][k]	⇒ [kː]	take care	[g][g]	⇒ [gː]	big glass
[m][m]	⇒ [mː]	from many countries	[n][n]	⇒ [nː]	Rain never falls.
[f][f]	⇒ [fː]	knife for cutting	[v][v]	⇒ [vː]	five vacant schools
[s][s]	⇒ [sː]	nice school	[z][z]	⇒ [zː]	loves zebras
[θ][θ]	⇒ [θː]	both things	[ð][ð]	⇒ [ðː]	with them
[l][l]	⇒ [lː]	She will lend you money.	[ʃ][ʃ]	⇒ [ʃː]	I wish she would come.

3.2.5.2. Elision of homorganic consonants

In the sequence of two homorganic consonants at word boundaries, the first consonant is assimilated to the following consonant (Table 11). When the first consonant is a stop followed by a homorganic stop, the first stop is not released (e.g. 'term paper', 'hot dog', 'big cat'). It has only a closing stage with no release stage while the second stop has only a release stage with no closing stage. When the first consonant is a stop [t] followed by a consonant, especially a homorganic consonant (e.g. [t, d, s, z, ʃ, tʃ, dʒ, n, l, r]), it may be replaced by a glottal stop (e.g. 'great smile' [ɡreɪʔ smaɪl]) (Cruttenden 2008:180). When a voiced fricative or stop is followed by a voiceless consonant, the first consonant becomes voiceless (e.g. 'his students' hi[s s]tudent, 'bad time' ba[t t]ime) (Imai 2007:118).

Table 11　Elision of homorganic consonants

Bilabia + Bilabial: [p] [b]] [m]					
[m]	[p]	term paper	[p]	[m]	put up my hand
[b]	[p]	web page	[p]	[b]	stop by
[m]	[b]	come back	[b]	[m]	rub my hands
Alveolar + Alveolar: [t] [d] [s] [z] [tʃ] [dʒ] [n] [l]					
[t]	[d]	'Sit down' ⇒ [si (ʔ)] down　'get down'　'hot dog' 'What do you want?' ⇒ [wadə juː want]			
	[s]	'great smile' ⇒ grea[ʔ s]mile			
	[z]	'great zeal' ⇒ [grea[ʔ z]eal			
	[ʃ]	'right shape' ⇒ righ[ʔ ʃ]ape			
	[tʃ]	'that chair' ⇒ tha[ʔ tʃ]air			
	[dʒ]	'great joke' ⇒ grea[ʔ dʒ]oke			
	[n]	'I don't know.' ⇒ I [douʔ] know.　'at night' ⇒ 'a[ʔ n]ight'　'not now' ⇒ no[ʔ n]ow			
	[l]	'at least' ⇒ a[ʔ l]least　'not like'			
	[r]	'that ring' ⇒ tha[ʔ r]ing			
[d]	[t]	'bad time' [bæd taɪm] ⇒ [bæt taɪm]　'used to', 'supposed to'			
	[n]	'good news' [ɡʊd nuːz] ⇒ [ɡʊn nuːz]			
	[l]	'good luck' ⇒ goo[dˡ l]uck			
[z]	[s]	'his students' ⇒ hi[s s]tudent　'She was sick.' ⇒ [wəs sik]			
Velar + Velar: [g] [k] [ŋ]					
[g]	[k]	'big cat' ⇒ 'bi[ʔ k]at			
[k]	[g]	'back garden'			
Dental + Dental [θ], [ð] (Imai 2007:118)					
[ð]	[θ]	'with thanks' ⇒ wi[θ θ]anks			
[θ]	[d]	'both these areas' ⇒ bo[ð ð]ese			
Dental Fricatives + Alveolar Fricative [θ] [ð] + [s] [z] ⇒ [s] [z]					
[ð]	[z]	'clothes' [kloʊðz] ⇒ [kloʊz]			
[θ]	[s]	'months' ⇒ 'mo[ns]'　'fourth square', 'North Star'			
[ð]	[s]	'breathe slowly' ⇒ brea[θ s]lowly			
Labiodental + Labiodental					
[v]	[f]	'We've found it.' ⇒ We[f f]ound it.			
[f]	[v]	'brief victory' ⇒ brie[v v]ictory			
Labiodental fricative + Bilabial ⇒ Bilabial nasal					
[v]	[m]	'give me' [ɡɪv miː] ⇒ [ɡɪm miː] ⇒ [ɡɪmɪ] (Cruttenden 2008:300, 304)			

3.2.5.3. Unreleased stop consonants

When a word that ends with a stop consonant is followed by a word that begins with a consonant, especially a stop or an affricate, the first stop is usually not released (e.g. 'pet cat' [pet˺kæt], 'big church' [bɪg˺tʃəɹtʃ] (Celce-Murcia 2010:166-167, Avery and Ehrlich 1992:85) (Table 12).

Table 12 Unreleased stop + consonant

	[p]	[b]	[t]	[d]	[k]	[g]
Stop						
[p]	deep pond	web page	white post	good person	black pepper	big party
[b]	stop by	web browser	that boy	good boy	black book	big boy
[t]	keep track	lab test	white tie	good time	black tiger	big time
[d]	soap dish	rub down	white dog	good deal	black dog	big deal
[k]	pop culture	job creation	white cake	red cap	rock climbing	big cat
[g]	keep going	Bob got it	that girl	bad girl	black god	big game
Nasal Stop						
[m]	stop me keep moving cheap meat	rub my hands	that man not/just me let/get me	Good morning old man bad man	black magic	big man
[n]	cheap nuts	rub now	not now	red nose	black net	big nose
Affricate						
[dʒ]	grape juice	Bob jumped	great joke	mad judge	black judge	big judgment
[tʃ]	top choice	superb choice	that chair	old church	pork chop	big church
Glide						
[w]	top winners	Bob won	last week hot water	And what?	black woman	big wing
[r]	top row	lab records	that ring Get rid of it	vivid reminder	bank robber	big root
[j]	cheap yacht	superb year	not yet	and yet	back yet	big yacht
Fricative						
[s]	top secret	job seeking	great smile	said so	Black Sea	big smile
[z]	top zone	superb zoo	great zoo	bad zoo	black zone	big zoo
[ʃ]	cheap shoes	superb shoes	great shoe	bad shoes	deck shoes	big shoes
[f]	pop fly	verb forms	not for me	bad foot	black feet	big feet
[h]	top half	job hunting	at home not here	good humor	black hair	big house
[v]	sharp vision	the job varies	not very	loud voice	black velvet	big valley
[θ]	keep thinking	Bob thought	what thing	old theater	black theater	big theater
[ð]	up the hill	rob this bank	out there	and the car	crack the code	dig this hole
Lateral						
[l]	up late	rub lightly	at last	good luck	look like	big land

3.2.5.4 Consonant cluster simplification and elision

It is sometimes difficult even for native speakers to pronounce two or more consonants together, and they omit one of the consonants to make the cluster easier to pronounce (Celce-Muicia 2010:100). Elision of consonants in English occurs most commonly when a speaker simplifies a complex consonant cluster, for example 'acts' becoming [æks] with [t] dropped (Roach 2009b: elision). The word-final [t] or [d] in

clusters of two consonants at a word boundary is often deleted when the following word begins with a consonant (e.g. 'last chance' [læs tʃæns]). In the same way, the [t] of the negative [-nt] is often elided before a following consonant (e.g. 'You mustn't lose it [yʊ mʌsn luːz ɪt]) (Table 13). Elision of word-final [t] or [d] eliminates the phonetic cue to paste tense when the following word has an initial consonant (Cruttenden 2008:304). For example, 'I walked back' and 'I walk back' sounds identical in connected speech. (Rogerson-Revell 2011:167).

Table 13　Consonant cluster C+[t]/[d]+C and consonant elision

[s]	[t]	C	'last chance' [læst tʃæns] ⇒ [læs tʃæns]		'just one'
[f]			'left wheel' [left wiːl] ⇒ [lef wiːl]		'drift by'
[ʃ]			'mashed potatoes' [mæʃt pəteitoʊz] ⇒ [mæʃ pəteitoʊz] 'finished now', 'pushed them'		
[p]			'kept quiet' [kept kwaɪət] ⇒ [kep kwaɪət] 'helped me', 'jumped well', 'stopped speaking'		
[k]			'exactly' [ɪgzæktli] ⇒ [ɪgzækli]		
			PAST TENSE	'walked back' [wɔːk bæk]	'thanked me', 'looked like'
[tʃ]			'reached Paris' [riːtʃt pærɪs] ⇒ [riːtʃ pærɪs]		'fetched me'
[n]			'went down' [went daʊn] ⇒ [wen? daʊn]		
			NEGATIVE [-nt]	'You mustn't lose it.' ⇒ [mʌsn luːz ɪt] 'Doesn't she know?' ⇒ [dʌzn ʃɪ nəʊ]	
[n]	[d]	C	'bend back' [bend bæk] ⇒ [ben bæk]		'kindness' [kaɪndnəs] ⇒ [kaɪnnəs]
[l]			'hold tight', 'old man', 'cold lunch'		
[z]			'refused both'		
[ð]			'unclothed body'		
[v]			'moved back'		
[b]			'robbed both', 'grabbed them'		
[g]			'lagged behind', 'dragged down', 'jogged by'		
[dʒ]			'changed color', 'urged them', 'arranged roses', 'judged fairly'		
[m]			PAST TENSE	'They seemed glad.' [siːm glæd]	

3.3. Assimilation

Assimilation is the process whereby one phoneme is changed into a different phoneme or a different allophone because of the influence of a nearby phoneme, causing a change in the place of articulation or in voicing (Brown 2006:283, Roach 2009b: assimilation). A given sound takes on the characteristics of a neighboring sound both within words and between words. (Celce-Murcia 2010:167). There are three types of assimilation in English: progressive (or perseverative), regressive (or anticipatory), and coalescent. Assimilation is referred to as co-articulation in experimental phonetics with an emphasis on articulatory explanations for why the assimilation occurs (Roach 2009b: assimilation).

3.3.1. Progressive assimilation

Progressive assimilation is most often found in the behavior of the plural suffix 's', the third person singular suffix 's', the possessive marker 's', and the past tense suffix '-d'. They are voiced when the preceding consonant is voiced (e.g. 'dogs' [dɔgz]) and devoiced when preceded by a voiceless consonant (e.g. 'docks' [dɑks]) (Table 14). Except for the above cases, progressive assimilation is relatively uncommon. It may occur when a plosive is followed by a syllabic nasal. For example, 'happen' is

pronounced [hæpm] rather than [hæpən] with the weak vowel [ə] being elided and the alveolar [n] changing to a bilabial [m], assimilating to the preceding bilabial [p] (Cruttenden 2008:303) (Table 15).

Table 14 Progressive assimilation and suffixes

	voiced		voiceless	
Plural Morpheme	dogs	[dɔgz]	docks	[dɑks]
Third Person Singular	he's	[hiːz]	it's	[ɪts]
Possessive Marker	John's	[ðʒɑnz]	Jack's	[ðʒæks]
Past Tense	moved	[muːvd]	worked	[wəːk]

Table 15 Progressive assimilation

[n] ⇒ [m]	'happen' [hæpən] ⇒ [hæpm] (See: Table 7 Syllabic Consonants)
[n] ⇒ [ŋ]	'bacon' [beɪkn] ⇒ [beɪkŋ]
[s] ⇒ [ʃ]	'bridge score' [brɪdʒ skɔːr] ⇒ [brɪdʒ ʃkɔːr] 'Church Street' : [tʃəːtʃ striːt] ⇒ [tʃəːtʃ ʃtriːt]

The alveolar stop [t] in a [nt] sequence between two vowels or before syllabic [l] is often dropped in connected speech, assimilating to the preceding alveolar nasal [n]. For example, 'twenty' [twenti] may be pronounced [tweni] (Celce-Muricai 2010:172). This phenomenon can occur across word boundaries as well. For instance, 'went away' may be pronounced [wenəweɪ] (Cruttenden 2008:304). In the same way, 'want to' [wɑːnt tə] becomes [wɑnnə] and 'going to' [goʊɪŋ tə] becomes [gənə] (Table 16).

Table 16 [nt] reduction

'twenty' [twenti] ⇒ [tweni]	'identify' [aɪdentɪfaɪ] ⇒ [aɪdenɪfaɪ]
'Internet' [ɪntəːnet] ⇒ [ɪnəːnet]	'winter' [wɪntəː] ⇒ [wɪnəː]
'He went away.' [went əweɪ] ⇒ [wen əweɪ]	
'want to' ⇒ [wɑːnt tuː] ⇒ [wɑːnt tə] ⇒ [wɑnttə] ⇒ [wɑntə] ⇒ [wɑnnə]	
'going to' ⇒ [gouiŋ tuː] ⇒ [gouin tə] ⇒ [gouin nə] ⇒ [gounə] ⇒ [gənə]	
'mantle' [mæntl] ⇒ [mænl]	

3.3.2. Regressive assimilation

Regressive assimilation is more pervasive than is progressive assimilation (Celce-Murcia 2010:168). Word-final [t, d, n, s, z] readily assimilate to the following word-initial consonant while keeping the original voicing, and the first consonant becomes more like the second one (Avery1992:87). Alveolars [t, d, n] are replaced by bilabials before bilabial consonants [p, b, m] (labialization), and by velars before velar consonants [k, g] (velarization); alveolar fricatives [s, z] are replaced by palato-alveolars [ʃ, tʃ, ʒ, dʒ] before consonants containing a palatal feature (palatalization); alveolars [t, d, n, l] get closure at the teeth when they are followed by a dental fricative [θ] or [ð] (dentalization). Voiced obstruents often become voiceless when they are followed by a voiceless consonant (devoicing). Assimilation may be total or partial. It is total in 'ten mice' [tem maɪs], where the assimilating sound [n] is identical with the conditioning sound [m]. It is partial in 'ten bikes', where the [n] falls under the influence of the following [b] to become [m], adopting the bilabiality of [b] but not the plosiveness (Crystal 2006:38).

3.3.2.1. Labialization of alveolars

Alveolars [t, d, n] assimilate to a following bilabial [p], [b] or [m] while keeping the original voicing, with the [t] replaced by [p], [d] by [b], [n] by [m]. For example, [t] in 'that boy' [ðæt bɔɪ] changes to [p] in anticipation of the following bilabial [b], resulting in [ðæp bɔɪ] (Table 17).

Table 17　Labialization of alveolars

Alveolar [t][d][n] + Bilabial [p][b][m] ⇒ [p][b][m]		
[t]	[p]	'that person' [ðæt pəːsən] ⇒ [ðæp pəːsən]
	[b]	'football' [fʊt bɔːl] ⇒ [fʊp bɔːl] ⇒ [fʊ? bɔːl] (Glottal Stop highly likely) 'that boy' [ðæt bɔɪ] ⇒ [ðæp bɔɪ] 'Don't be late.' [doʊnt bɪ] ⇒ [doʊnp bɪ] ⇒ [doʊmp bɪ]
	[m]	'that man' ⇒ [ðæt mæn] ⇒ [ðæp mæn] 'just me' [dʒʌst mi] ⇒ [dʒʌsp mi] ⇒ [dʒʌs mi] 'let me' [let mi] ⇒ [lep mi] ⇒ [lemɪ] or [le? mi] (Glottal Stop)
[d]	[p]	'good pen' ⇒ goo[b p]en
	[b]	'bad boy' ⇒ ba[b b]oy
	[m]	'sad man' ⇒ sa[b m]an
[n]	[p]	'ten ponies' ⇒ te[m p]onies
	[b]	'seven boys' ⇒ seve[m b]oys
	[m]	'thirteen men' ⇒ thirtee[m m]en

3.3.2.2. Velarization of alveolars

Alveolars [t, d, n] assimilate to a following velar [k] or [g] while keeping the original voicing, with the [t] being replaced by [k], [d] by [g], [n] by [ŋ] (e.g. 'fruit cake' [frʊːk keɪk]) (Davenport 2010:25-26) (Table 18).

Table 18　Velarization of alveolars

Alveolar [t][d][n] + Velar [k][g] ⇒ [k][g][ŋ]		
[t]	[k]	'that case' ⇒ [ðæk keɪs]　　'It can wait.' ⇒ [? k(ə)n weɪt] 'He won't come.' [woʊnt kʌm][t]+[k] ⇒ [woʊnk kʌm] [n]+[k] ⇒ [woʊŋk kʌm] [ŋ]+[k]
[d]		'sad king' [sad kɪŋ] ⇒ [sæg kɪŋ]
[n]		'ten cakes' [ten keɪks] ⇒ [teŋ keɪks]
[t]	[g]	'quite good' [kwaɪt gʊd] ⇒ [kwaɪk gʊd]
[d]		'bad girl' [bæd gəːl] ⇒ [bæg gəːl]
[n]		'ten girls' [ten gəːrlz] ⇒ [teŋ gəːrlz]　　'I can go.' [aɪ kæn gou] ⇒ [aɪ kæŋ gou]

3.3.2.3. Palatalization of alveolar fricatives

Alveolar fricatives [s] and [z] assimilate to a following palatal glide [j], palato-alveolar fricative [ʃ] or palato-alveolar affricates [tʃ] [dʒ] while keeping the original voicing, with [s] and [z] replaced by a palato-alveolar fricative [ʃ] and [ʒ] respectively (Davenport 2010:29). For example, 'this shop' [ðɪs ʃɑp] becomes [ðɪʃ ʃɑp] and 'Does she?' [dʌz ʃɪː] becomes [dʌʒ ʃɪː] (Table 19).

Table 19 Palatalization of alveolar fricatives

[s]	[j]	[ʃ]	'this year' [ðɪs jɪəɹ] ⇒ [ðɪʃ jɪəɹ]
	[ʃ]		'this shop' [ðɪs ʃɑp] ⇒ [ðɪʃ ʃɑp]
	[tʃ]		'this chapter' [ðɪs tʃæptəɹ] ⇒ [ðɪʃ tʃæptəɹ]
	[dʒ]		'this judge' [ðɪs dʒʌdʒ] ⇒ [ðɪʃ dʒʌdʒ]
[z]	[j]	[ʒ]	'those young men' [ðoʊz yʌŋ men] ⇒ [ðoʊʒ yʌŋ men]
	[ʃ]		'Does she?' [dʌz ʃiː] ⇒ [dʌʒ ʃiː]
	[tʃ]		'those churches' [ðoʊz tʃətʃɪz] ⇒ [ðoʊʒ tʃətʃɪz]
	[dʒ]		'those judges' [ðoʊz dʒʌdʒɪz] ⇒ [ðoʊʒ dʒʌdʒɪz]

3.3.2.4. Dentalization of alveolars

Alveolars [t, d, n, l] assimilate to a following dental fricative [θ] or [ð] and get closure at the teeth, with [t] replaced by [t̪], [d] by [d̪], [n] by [n̪], and [l] by [l̪] (Table 20).

Table 20 Dentalization of alveolars: dental [t][d][n][l]

Alveolars [t][d][n][l] + Dental fricatives [θ][ð] ⇒ [t̪][d̪][n̪][l̪]		
[t]	[θ]	'hot thing' ⇒ ho[t̪θ]ing 'at three' ⇒ a[t̪θ]ree
[d]		'good thing' ⇒ [gʊd̪θɪŋ]
[n]		'born thieves' ⇒ bor[n̪θ]ieves
[l]		'health' ⇒ [hel̪θ]
[t]	[d]	'at that' ⇒ a[t̪ð]at
[d]		'bad though' ⇒ ba[d̪ð]ough
[n]		'in that' ⇒ i[n̪ð]at 'in the' ⇒ [ɪn̪ðə]
[l]		'sell them' ⇒ [sel̪ð]em

3.3.2.5. Final obstruent devoicing

Voiced fricatives [v, z, ð] and a voiced palato-alveolar affricate [dʒ] often become voiceless word-finally when followed by a voiceless obstruent (Davenport 2010:29). The word-final labiodental fricative [v] becomes voiceless, being replaced by [f] when preceding a voiceless obstruent [t, k, s, f] (e.g. 'have to' [hæftə]). The sound may be lost altogether in unstressed function words. For example, the auxiliary 'have' in 'could have been' and the preposition 'of' in 'a piece of cake' are both reduced to [ə] and become identical in connected speech.

In the same way, the word-final alveolar fricative [z] becomes voiceless when followed by a voiceless obstruent [t, s, h] (e.g. 'has to' [hæstə]). The dental fricative [ð] becomes a voiceless [θ] when it assimilates to a following voiceless alveolar fricative [s] or voiceless dental fricative [θ]. The palato-alveolar affricate [dʒ] becomes a voiceless [tʃ], when assimilating to a following voiceless alveolar fricative [s] (Table 21).

Table 21　Devoicing: final obstruent devoicing

Voiced obstruents + Voiceless consonant		
[v] ⇒ [f]		
[v]	[t]	'have to' [hæv tə] ⇒ [hæftə]
	[k]	'of course' [əv kɔːrs] ⇒ [əf kɔːrs]
	[s]	'move slowly' [muːv sloʊli] ⇒ [muːf sloʊli]
	[f]	'We've found it.' ⇒ We[f f]ound it.
[z] ⇒ [s]		
[z]	[t]	'has to' ⇒ [hæz tə] ⇒ [hæstə] 'plays tennis' play[z t]ennis ⇒ play[s t]ennis 'yours truly' ⇒ your[s t]ruly
	[s]	'his students' ⇒ hi[s s]tudent (Table 11 Elision of Homorganic Consonants)
	[h]	'men's hat' ⇒ men[s h]at
[ð] ⇒ [θ]		
[ð]	[s]	'with sugar' wi[ð ʃ]ugar' ⇒ wi[θ ʃ]ugar　'with sympathy'
	[θ]	'with thanks' wɪ[ð θ]æŋks ⇒ wɪ[θ θ]æŋks
[dʒ] ⇒ [tʃ] (Table 15 Progressive Assimilation)		
[dʒ]	[s]	'bridge score' [brɪdʒ skɔr] ⇒ [brɪtʃ skɔr]

3.3.3. Coalescent assimilation

Coalescent assimilation (or reciprocal assimilation) is the process where the first and second sounds in a sequence come together and mutually condition the creation of a third sound with features from both original sounds (Celce-Murcia 2010:170). Final alveolar [s] and [z] combine with initial palatal glide [j], typically of a pronoun 'you' or 'your', to form palatalized fricatives [ʃ] and [ʒ] respectively. In the same way, final [t] and consonant sequence [ts] combine with [j] to form affricate [tʃ], while final [d] and consonant sequence [dz] form afficate [dʒ] with [j]. For example, 'miss you' [mɪs juː] changes to [mɪʃʊ] and 'need you' [niːd juː] becomes [niːdʒʊ] (Table 22). [t] may be replaced by a glottal stop before palatal glide [j]; for example 'not yet' may be pronounced [nɑʔ jet] rather than [nɑtʃet] (Cruttenden 2008:180, Roach 2009:111).

Table 22　Coalescent assimilation

Palatalization: Final alveolar consonants [s, z, t, d, ts, dz] + initial palatal [j]				
[s]	[j]	⇒ [ʃ]	'miss you' [mɪs juː] ⇒ [mɪʃʊ]	'kiss you', 'in case you'
[z]		⇒ [ʒ]	'He knows your name.' ⇒ [noʊʒəˌ]	'It was yellow.'
[t]		⇒ [tʃ]	'I hate you.' ⇒ [heɪtʃə]	'I want you to go there.'
[ts]			'He hates you.' ⇒ [heɪtʃə]	'She lets you stay up late.'
[d]		⇒ [dʒ]	'I need you.' ⇒ [niːdʒʊ]	'would/could/should/did you'
[dz]			'She needs you.' ⇒ [niːdʒʊ]	'He never heeds your advice.'

4. Concluding Remarks

"The learners' unfamiliarity with reduced forms in English is one major source of difficulties in their listening comprehension as well as in their pronunciation skills," says Ito (2006:18). In connected speech, various phenomena, such as reduction, elision, assimilation, intrusion, juncture, linking, and contraction,

occur. Learners should be taught what is happening in natural speech. However, it would not be practical or useful to teach all of these modifications that occur in native speech to the extent that the learners can produce them. Rogerson-Revell (2011:175) says: "Recognition is more important than production for most L2 learners." She also says (2011:172): "Regarding the modifications that occur in connected speech, many phonologists and pronunciation teachers agree that it seems advisable to help learners *recognize* native speakers' production of such features in fluent connected speech but not encourage learners to *produce* all of them themselves. Priority should be given to practice with accentuation and rhythm, while linking and elisions should be given greater emphasis than assimilations."

First, we need to provide an overview of connected speech in English systematically so that the students will be able to recognize what hinders their listening comprehension. An awareness of some most common types of connected speech is useful for listener comprehension (Rogerson-Revel 2011:177). The areas that should be highlighted are: 1) linking, especially consonant-to-vowel and consonant-to-consonant linking, which often gives problems for Japanese learners of English when they try to find word boundaries; 2) elision of consonants, which deletes the sounds that the learners expect to hear; 3) consonant assimilation, which changes the sounds that they expect to hear to different sounds or deletes some sounds as a result of assimilation.

Once the learners understand how the sounds they expect to hear change or disappear in connected speech, they will not have "a very rude awakening when [they try] to understand native speech in natural communicative situations" (Ur 1987:10). Then, their listening comprehension as well as pronunciation will improve with less frustration compared to the case where only citation forms are taught.

REFERENCES

有本純 (2005)「発音指導における教師の役割：怪しい発音指導の正体」,『英語教育』12月号, 27-29, 大修館, 東京.

Avery, Peter and Susan Ehrlich (1992) *Teaching American English Pronunciation*, Oxford University Press, Oxford.

Brinton, Laurel J. (2000) *The structure of modern English: a linguistic introduction*, John Benjamins, Amsterdam.

Brosnahan, L.F. and Bertil Malmberg (1976) *Introduction to Phonetics*, Cambridge University Press, Cambridge.

Brown, Gillian (1990) *Listening to Spoken English*, 2nd ed., Longman, Essex.

Brown, James Dean (2006) "Teacher Resources for Teaching Connected Speech," *Perspectives on Teaching Connected Speech to Second Language Speakers*, ed. by Brown, James Dean and Kimi Kondo-Brown, 27-47, University of Hawaii Press, Honolulu.

Brown, J. D. and K. Kondo-Brown (Eds.) (2006) *Perspectives on Teaching Connected Speech to Second Language Speakers*, University of Hawaii Press, Honolulu.

Brown, J. D. and K. Kondo-Brown (2006) "Introducing Connected Speech," *Perspectives on Teaching Connected Speech to Second Language Speakers*, ed. by Brown, James Dean and Kimi Kondo-Brown, 1-15, University of Hawaii Press, Honolulu.

Cartier, Francis A. (1967) 'Some Important Oddities of English Phonetics', *TESOL Quarterly* Vol.1, No. 3, 52-55.

Celce-Murcia, Marianne, Donna M. Brinton, and Janet M. Goodwin (2010) *Teaching Pronunciation: A Course Book and Reference Guide*, 2nd ed., Cambridge University Press, Cambridge.

Cook, Ann (2000) *American Accent Training*, 2nd ed., Barron's, New York.

Cruttenden, Alan (2008) *Gimson's Pronunciation of English*, 7th ed., Hodder Education, London.

Crystal, David (1995) *The Cambridge Encyclopedia of the English Language,* Cambridge University Press, Cambridge.

Crystal, David (2006) *A Dictionary of Linguistics and Phonetics*, 5th ed., Blackwell, Oxford.

Davenport, Mike and S.J. Hannahs (2010) *Introducing Phonetics and Phonology*, 3rd ed., Hodder, London.

Giegerich, Heinz J. (1992) *English Phonology*, Cambridge University Press, Cambridge.

Hagen, Stacy A. (1988) *Sound Advice*, Prentice Hall Regents, Englewood Cliffs.

今井邦彦 (2007)『ファンダメンタル音声学』ひつじ書房, 東京.

今井由美子, 他 (2010)『英語音声学への扉』英宝社, 東京.

Ito, Yasuko（2006）“The Significance of Reduced Forms in L2 Pedagogy,” *Perspectives on Teaching Connected Speech to Second Language Speakers*, ed. by Brown, James Dean and Kimi Kondo-Brown, 17-25, University of Hawaii Press, Honolulu.

川越いつえ（1999）『英語の音声を科学する』大修館書店，東京.

Ladeforges, Peter（2006）*A Course in Phonetics,* 5th ed., Thomson, Boston.

Lujan, Beverly A.（2006）*The American Accent Guide*, 2nd ed., Lingual Arts, Salt Lake City.

牧野武彦（2005）『日本人のための英語音声学レッスン』大修館書店，東京.

Matthews, Peter H.（1997）*Oxford Concise Dictionary of Linguistics*, Oxford University Press, New York.

Mayer, Charles（2009）*Introducing English Linguistics*, Cambridge University Press, Cambridge.

松坂ヒロシ（1986）『英語音声学入門』研究社，東京.

Roach, Peter（2009a）*English Phonetics and Phonology: A Practical Course*, Cambridge University Press, Cambridge.

Roach, Peter（2009b）*English Phonetics and Phonology Glossary: A Little Encyclopaedia of Phonetics*. <http://www.cambridge.org/servlet/file/EPP_PED_Glossary.pdf?ITEM_ENT_ID=2491706&ITEM_VERSION=1&COLLSPEC_ENT_ID=7>.

Rogerson, Moana（2006）“Don’cha Know? A Survey of ESL Teachers’ Perspectives on Reduced Forms Instruction,” *Perspectives on Teaching Connected Speech to Second Language Speakers*, ed. by Brown, J. D. and K. Kondo-Brown, 85-97, University of Hawaii Press, Honolulu.

Rogerson-Revell, Pamela（2011）*English Phonology and Pronunciation Teaching*, Continuum Press, London.

佐久間淳一，加藤重広，町田健（2004）『言語学入門』研究社，東京.

柴田雄介，横山志保，多良静也（2006）「音声指導に関する教員の実態調査」日本英語音声学会第 6 回九州沖縄四国支部研究大会.

Shockey, Linda（2003）*Sound Patterns of Spoken English*, Blackwell, Oxford.

Skandera, Paul and Peter Burleigh（2011）*A Manual of English Phonetics and Phonology*, Gunter Narr Verlag, Tubingen.

上田洋子，大塚朝美（2011）「発音と音声のしくみに焦点をあてた中学校英語教科書分析：インプットの基礎を考察する」『大阪女学院大学紀要』 7 号 , pp.15-32.

Ur, Penny（1984）*Teaching Listening Comprehension*, Cambridge University Press, Cambridge.

Weinstein, Nina（2001）*Whaddaya Say?* 2nd ed., Longman, New York.

Weisser, Martin（2005）（last edited: 16-Oct-2007 15:27）*Phonetics and Phonology*.
<http://ell.phil.tu-chemnitz.de/phon/phon.html.>

（2011年11月25日掲載決定）