

# Orthographic Representation of Reductions and Rapid Speech in English

Timothy M. HARRIS

## Introduction

One of the greatest difficulties faced by second language learners is aural comprehension of the target language when spoken by native speakers at natural speed. Perhaps even more difficult is the development of a fluent, native-like accent that can be easily understood by native speakers. These difficulties are largely due to the difference between the language presented in textbooks and the language as actually spoken by native speakers on a daily basis. There are a number of factors that play a role in this, such as the use of slang and colloquialisms, but the focus here will be limited to phonetic and phonological issues.

Both English and Japanese exhibit the phonetic phenomenon of *reduction*, but the syllabic/moraic structure of Japanese, and the modern kana orthography, which involves an almost one-to-one correspondence between syllabic sound and symbol, results in far less complexity than is found in English. Some common examples of reduction in Japanese include the following:

(1) -kereba > -kerya > -kya

lenition of [b], fusion of [rea] to [rja], lenition of [r], fusion of [keja] to [kja]

Ex: 行かなければ > 行かなけりゃ > 行かなきゃ

(2) -te wa > -cha

---

平成17年3月31日 原稿受理  
大阪産業大学 教養部

lenition of [w], palatalization of [t] through anticipatory assimilation to [e],  
subsequent affrication: [tɕ]

Ex: 行かなくては > 行かなくちゃ

(3) de wa > ja (voiced counterpart of #2 above)

lenition of [w], palatalization of [d] through anticipatory assimilation to [e],  
subsequent affrication: [dʑ]

Ex: それでは > それじゃ

(4) -te shimau > -chimau > -chau

lenition of [e], fusion of stop [t] and fricative [ɕ] to affricate [tɕ], lenition of [m]

Ex: 食べてしまう > 食べちまう > 食べちゃう

(5) -ranai > -nnai

deletion of [a], anticipatory assimilation of [r] to [n]

Ex: わからない > わかんない

(6) -te iru > -teru deletion of [i]

Ex: 食べている > 食べてる

(7) no > n deletion of [o]

Ex: 僕のところ > 僕んところ、僕のうち > 僕んち

(8) -te oku > -toku deletion of [e]

Ex: 開けておく > 開けとく

(9) desu > ssu deletion of [dɕ]

Ex: わからない です > わからないッス

(10) ko-ka / ku-ka > kka

haplology involving [k] and deletion of intervening vowel

Ex: どこか > どっか、よく考えたら > よっ考えたら

(11) sumimasen > suimasen / summasen

lenition of [m] OR lenition of [i]

Ex: すみません > すいません / すんません

(12) sayônara > sayonara > sainara

shortening of long vowel [o:], deletion of short vowel [o]

Ex: さようなら > さよなら > さいなら

## Contractions

One issue which complicates the situation in English is the existence of a clear dichotomy between contractions like “can’t” or “you’re,” which fall within the realm of standard written English, and nonstandard spellings of reduced forms like “wanna” or “gonna,” which are generally found only in media such as comics or advertising. Almost all standard written contractions involve the modal verbs “would” and “will,” forms of the verb “be” (am, are, is), forms of the auxiliary verb “have” (have, has, had), or the negative “not.” One of the few exceptions is “let’s” (< “let us”). The negative “not” is shortened only with modal verbs, present and past forms of the verbs “be” and “have,” and the auxiliary verb “do” : *can’t, won’t, shouldn’t, wouldn’t, couldn’t, mustn’t, aren’t, isn’t, wasn’t, weren’t, haven’t, hasn’t, hadn’t, don’t, doesn’t, and didn’t*, as well as archaic forms like “shan’t” and the substandard form “ain’t.” Only one standard contraction is irregular: “won’t” instead of the expected “willn’t.”

The most common contractions of the verbs “be” and “have” occur after subject pronouns (e.g. I’m, you’re, she’s, we’ve) or nouns (e.g. “Bill’s going, but his father’s not.”) Other forms, while common in spoken English, are less common in written English. For example, “How’re you?” or “How tall’re you?” Examples like these, even when not actually written as contractions, are often pronounced with their reduced forms when read aloud, placing them perhaps more in the category of reductions than actual contractions. There are also two ambiguous contractions (’s and ’d) that can be distinguished only by context:

he’s < he is OR he has

he’d < he had OR he would

Another common form of contraction consists of modal verb plus the auxiliary verb “have” : *would’ve, could’ve, should’ve, must’ve, will’ve, might’ve*. While these forms are less common in written English, they are actually far more common in spoken English than their uncontracted counterparts. Finally, there are double contractions involving the modal verbs “will” or “would” in combination with the auxiliary verb “have.” While these are extremely common in spoken English, they tend to be avoided in writing, due to their

cumbersome appearance: *I'll've*, *I'd've*, etc.

The majority of native speakers of English use contractions in their everyday speech whenever possible, and one measure of native-like linguistic competence is the ability to use contractions accurately. For example, the stand-alone sentence “Yes, I will,” is never contracted to “Yes, I’ll.” The speech of ESL/EFL learners is often conspicuously lacking in contractions, resulting in stilted, unnatural-sounding English. Fixed expressions like “can’t” are usually less of a problem, providing that the student has had adequate exposure to such forms, but contractions which must be produced on the spot (ex: “My pen’s missing.”) are much less likely to be employed.

### Strong and Weak Forms

A second issue which complicates the issue of reductions in English is the coexistence of *strong* and *weak* forms of words. As mentioned earlier, modern kana orthography represents an almost one-to-one correspondence between the sounds and symbols of Japanese, albeit in syllabic units. Of course, this was not always the case. Historical kana orthography reveals the extent of phonetic change in the spoken language (e.g. てふ [teɸu] > ちょう [tɕo:]). Unlike Japanese, which underwent spelling reform in the mid 1940’s, current English spelling often reflects historical pronunciation, and in this regard can be considered analogous to historical kana orthography. One exception in modern standard Japanese is the fact that [dzi] and [zu] can each be written in two ways (じ・ぢ and ず・づ respectively), although some dialects reportedly make a phonetic distinction. “Silent” vowels also represent an exception to the one-to-one correspondence. For example, the [w] which is normally silent in わかりますか? [wakarimaska] is obligatorily pronounced in the question わかります? [wakarimasu].

The fact that English spelling does not observe a one-to-one correspondence between orthography and the spoken language leaves room for considerable variation in pronunciation. The developmental history of English and the complexities of its irregular spelling have contributed to the coexistence of strong and weak forms of words. The latter almost invariably involve syllabic consonants or a reduced vowel such as [i] or the schwa [ə]. This is simply not an issue in Japanese, since all five vowels are pronounced consistently clearly, with no occurrence of the schwa. Ladefoged (2001) provides the following chart of examples of common strong and weak forms in English, and mentions

that “more than five times as many could easily have been listed.” Syllabic consonants are of course represented with the diacritic [̩] underneath.

<u>Word</u>	<u>Strong Form</u>	<u>Weak Form</u>
a	eɪ	ə
and	ænd	ənd, ɹd, ən, ɹ
as	æz	əz
at	æt	ət
can	kæn	kən, kɹ
has	hæz	həz, əz, z, s
he	hi	i, hɪ, ɪ
must	mʌst	məst, məs, mʃ
she	ʃi	ʃɪ
that	ðæt	ðət
to	tu:	tʊ, tə
would	wʊd	wəd, əd, d

Of course, the existence of weak forms in English stems from the fact that English has stress accent (as opposed to Japanese, which has pitch accent). In English, vowels in unstressed syllables are often reduced, with a schwa or syllabic consonant representing the maximum potential reduction. Words that are almost identical in written form are often pronounced entirely differently, depending on the placement of stress, due to differences in the position of reduced vowels. The following two words, which differ only in the morphological ending “-y,” do not share a single phonetic vowel in common (in the same position):

photograph	['fɔ:təgræf] (American), ['fəʊ:təgrɒf] (British)
photography	[fə'tɑ:gɹəfi] (American), [fə'tɔ:gɹəfi] (British)

The prevalence of katakana-influenced pronunciation of English in Japan, exacerbated by textbooks which stubbornly insist on using katakana as a pronunciation guide, pulls many learners away from the use of weak forms. Katakana-influenced pronunciation, if

anything, is “stronger” than the strong forms listed above, and has the unfortunate effect of reinforcing them at the expense of the weak forms, which are far more common in the speech of native English speakers. Note how the two words above are rendered in katakana orthography, according to the *Kôjien* (1998). (Phonetic transcription is mine.)

フォトグラフ [ɸotoguraɸu]

フォトグラフィ [ɸotoguraɸi]

Like the English spelling, they are identical except for the final vowel. Even putting aside the phonetic distinctions between [ɸ] and [f], and [r] and [ɹ], there is simply no way to accurately render a schwa in katakana, or in this case, to even accurately render [æ]. The use of [a] in the first word, and [o] for the second vowel in the second word is a reasonable approximation of the British pronunciation, but this is less true for the American pronunciation. The closest one could come to accurately representing standard American pronunciation is perhaps the following:

フォータグラフ [ɸo:taguraɸu]

ファターグラフィー [ɸata:guraɸi:]

Katakana representations of English often tend to reflect spelling more than actual pronunciation, and British pronunciation rather than American, meaning that traditional katakana spellings are not at all a reliable indication of natural American pronunciation.

### Linking

Finally, a third issue which complicates the picture in English is the phenomenon of *linking* or *liaison*, in which word-final consonants are pronounced together with a following vowel, with no intervening pause in articulation. Linking leads to reductions that cross word boundaries, such as the infamous “gonna,” which consists of the two words “going” and “to.” In cases like this, it is sometimes difficult to determine exactly where one word ends and the next begins. This is simply not an issue in Japanese, due to its vowel-final syllabic structure. The sole exception is of course the nasal *n* (ん), which is never linked to a following vowel. Conversely, native English learners of Japanese often

incorrectly link the nasal *n*, pronouncing, for example, the name Ken'ichi (けんいち) as [kɛnitɕi] (ケニチ) instead of [kɛnitɕi] or [kɛitɕi].

Most Japanese learners of English receive little or no overt instruction in linking, and traditional katakana orthography encourages students to pronounce each word in isolation. Moreover, katakana forcibly squeezes English words into the Japanese phonological system, turning consonant-final words into vowel-final ones: [dzisw izw hizw ɛggw] ("This is his egg.") instead of [ðɪzɪzɪzɛg]. Katakana orthography therefore pulls learners away from linking in the same way that it pulls them away from weak forms.

Unfortunately, even textbooks dedicated to teaching natural English pronunciation often give short shrift to the topic of linking. Baker & Goldstein (1990), for example, include only 3 short sections on linking ("Joining Sounds") in a book consisting of 46 units, and these sections deal only with the linking of similar (or identical) consonants:

s + s = s	Ex: let's sit, jus(t) sit
z + s = s	Ex: whose seat, Who's speaking?
sh + sh = sh	Ex: Spanish shoes, English sheep

This is simply insufficient, given the enormous importance of linking in English. As Cook (1991) points out, "If you speak word by word, as many people who learned 'printed' English do, you'll end up sounding mechanical and foreign." She advises learners to connect words to form sound groups: "Instead of thinking of each word as a unit, think of *sound units*. These sound units may or may not correspond to a word written on a page. [...] Sound units make a sentence flow smoothly..."

All of this means that the difference between the pronunciation of words in isolation (i.e. in their citation forms) and their pronunciation in a natural context (i.e. in rapid, connected speech) is much greater in English than in Japanese. Most learners of English do not receive sufficient instruction in recognizing and using reductions and rapid-speech forms, often due to instructors' pedantic beliefs that these are substandard. However, as Ladefoged (2001) cogently points out: "There is, of course, nothing slovenly or lazy about using weak forms and assimilations. Only people with artificial notions about what constitutes so-called good speech could use adjectives such as these to label [this] kind of speech. Weak forms and assimilations are common in the speech of every sort of speaker

in both Britain and America. Foreigners who make insufficient use of them sound stilted.”

### Orthographic Representation

Wong (1987) provides the following conversation, taken from a comic strip called “The Born Loser.” The spelling conventions employed are intended to reflect natural rapid-speech pronunciation, including reductions, weak forms, and linking. This conversation is immediately understandable to most native speakers of English, and it is obvious that the speakers (A and B) are both fishermen.

A: Hiya Mac.	B: Lobuddy.
A: Binear long?	B: Cuplours.
A: Ketchuneney?	B: Goddafew.
A: Kindarthay?	B: Bassencarp.
A: Enysizetoum?	B: Cuplapowns.
A: Hittinard?	B: Sordalite.
A: Wahchoozin?	B: Gobbawurms.
A: Fishanonboddum?	B: Rydonnaboddum.
A: Goddago.	B: Tubad. Takideezy.
A: Seeyarown. Gluk.	

Even a short conversation like this one exhibits a tremendous degree of reduction, as can be seen when compared with my standard written English version below, where deleted portions are enclosed in parentheses.

A: Hi-ya, Mac.	B: (Hel)lo, buddy.
A: (Have you) been (h)ere long?	B: (A) coupl(e of h)ours.
A: (Are you) catchin(g) any?	B: (I) got a few.
A: (What) kind ar(e) they?	B: Bass an(d) carp.
A: (Is there) any size to (th)em?	B: (A) coupl(e) o(f) poun(d)s.
A: (Are they) hittin(g) (h)ard?	B: Sort o(f) light.
A: What (are) y(ou) usin(g)?	B: (A) gob o(f) worms.
A: (Are you) fishin(g) on (the) bottom?	B: Right on (th)e bottom.

A: (I've) got to go.

B: (That's) too bad. Take it easy.

A: See you (a)roun(d). G(ood) luck.

In addition to weak forms of common words (*of, and, to, you*), common phonetic deletions here include word-initial [h], the “g” in “-ing” (i.e. [ŋ] > [n]), the [ð] in “them,” and the [d] in “pounds” and “around.” The “h” in “hours” is of course silent, so the omission of the “h” does not represent a true deletion; it is only omitted to more accurately reflect the actual pronunciation. The same is true of the “e” in “are.” In addition to phonetic deletions, this conversation also reflects the “telegraphic speech” often used in casual English, in which entire words are deleted, especially at the beginning of interrogative utterances: *have you, are you, are they, is there*, etc. Linking is reflected by the fact that each utterance is written together as one word. Particularly extensive examples of linking include: *binear, cuplours, hittinard, wahchoozin*, and *gluk*. Of these, the most complex is *wahchoozin*, which involves the following changes:

What are you using?

What are you usin'?

[ŋ] > [n]

What are ya usin'?

[u] > [ə] (weak form of “you”)

What ya usin'?

deletion of “are”

Whacha usin'?

[tʃ] > [tʃ] (affrication of palatalized [t])

Whachusin'? = Wahchoozin?

deletion of [ə]

Throughout the conversation, the intervocalic flap pronunciation [ɾ] of the phoneme /t/ is consistently spelled as “d”, in an attempt to portray its voiceless quality: *goddafew, boddum, goddago, sordalite, takideezy*. Similarly, other spellings are chosen because they reflect pronunciation more closely. The underlined portions of the following utterances are examples of this: *kindarthay, wurms, cuplours, enysize*. This is not always the case, however. *Binear* could have been left as *binere*, and *tubad* could have been left as *toobad*. In fact, *tubad* is a poor choice, since “u” is used in other words to represent the sound [ʌ]. Conversely, the author has opted to leave the word *few* as it is, instead of using a more phonetic spelling like *fyoo*. The letter “c” is left alone in *bassencarp* and *cuplours*, but inexplicably changed to “k” in *ketchuneney*. However, the “tch” is left alone, rather than

being simplified to “ch” as in *wahchoozin*. The vowel digraph “ou” (representing the sound [aʊ]) is changed to “ow” in two cases (*cuplapowns*, *seeyarown*), but not in *cuplours*. The sound [aɪ] is spelled in two ways (*sordalite*, *rydonnaboddum*), as is the sound [u]: *wahchoozin*, *tubad*. Finally, the schwa [ə] is spelled in no less than 4 ways: *ketchuneny*, *bassencarp*, *hittinard*, *fishanonboddum*.

This kind of irregularity and inconsistency is to be expected, given that the conversation appears in a comic strip, and is designed primarily for amusement and humorous effect. However, this raises the important question of how reductions and rapid speech in English should best be represented orthographically. Even in textbooks devoted entirely to pronunciation, there appears to be no standard, and the majority of textbooks employ an illogical mix of standard English spelling (which may or may not reflect actual pronunciation) and non-standard, invented spellings intended to reflect pronunciation, similar to what was observed in the conversation above. Even if we limit the discussion to the invented “pronunciation spellings,” there is often a conspicuous lack of logic and consistency in the conventions used. There are two major problems with this state of affairs. First, each book uses a different spelling system, and this lack of a standard places an unfair burden on learners. Second, most of the systems (if not all) are far from ideal, as we shall soon see. Here are four common words as spelled in four different textbooks:

	Rost, et al (2001)	Cook (1991)	Weinstein (2001)	Naumann (1986)
and	-n	'n	'n'	'n', 'nd
or	-r	er	er	er
to	tuh-	t' / d'	ta / da	t'
of	-uh	'v / [']	a	a' / uv

Even with small words such as these, there is a considerable degree of variation in spelling, including details like the use of the apostrophe. Other typical forms used by Rost (2001) include the following:

-n (and), -r (or)  
fe(r), tuh- (to), -uh (of)

gav-er, bought-ihm  
 guh-nuh, wuh-nuh, haf-tuh  
 ih-she, z-(th)ere, where-ze, wuh-zit  
 d-yuh, d-we, duh-they, (duh)-ze  
 why-(d)z(th)at, who-(d)ze  
 duhz-n(t), duh-zuhn-she  
 do(nt), do(n)-(t)they, do(n)-chuh  
 d(ih)-she, dih-they, (d)zit, dit, dthat, where-d(ih)-yuh  
 dih-n(t), dih-n-it, dih-n-(t)they, dih-n-e  
 I-av, she-as, ha-she, they-av, v-yuh, ze  
 ca-nyuh, cu-de, cu-yuh  
 wu-(d)yuh, wu-de, wu-she  
 wih(l)-yuh, wih(l)-she

Cook (1991) uses invented forms like the following in an exercise on colloquial reductions and liaisons:

Nā chet.  
 Whenju geddit?  
 Whyju tay kit?  
 I'll meechu layder.  
 Wouldjoo mindifai try dit?  
 Couldjoo spee di dup, pleez?  
 I shoulde toljoo.  
 Tellim I missim.  
 Whyncha getta job?  
 Jeet? No, joo?  
 I dunno, stoo hard.

However, one page earlier she uses a semi-phonetic notation in an exercise on liaisons and glides:

I shüd pi kæpän the<sup>(y)</sup>æmerika nintənash'n pæddern pridy<sup>(y)</sup>ezile, altho the<sup>(y)</sup>only

weidə geddidiz t' prækʰtɪs əlləv θh'time. I'vb'n təkɪŋ to<sup>(w)</sup> ə læddəvə merɪkən  
zla<sup>(t)</sup>ly, 'n they tell me the dai meazɪə to<sup>(w)</sup> ʌndəstæn<sup>(d)</sup>. Enyweɪ<sup>(y)</sup> I kʉd go<sup>(w)</sup> ə nə  
næn, bu<sup>(t)</sup> θhə<sup>(y)</sup> ɪmpɔr'n<sup>(t)</sup> θɪŋgɪz t' lɪsənweɪlən saʉn<sup>(d)</sup> gʉd.

She also includes a list of common words containing reduced sounds, in which she uses the apostrophe to indicate reduced vowels:

't (at / it)  
fer (for)  
frɪm (from)  
'n (an / and / in)  
er (are / or)  
w'n (one)  
θh' / [ɪ] (the)  
[ə] (a)  
k'n (can)  
h'd / 'd (had)  
w'z / wuz (was)  
w't (what)  
w'd (what do)  
w'j (what did you)  
θh't (that)

Most of the forms used by Weinstein (2001) are the same ones traditionally used to write reductions in material intended for native speakers (e.g. comics). A notable exception is *donno*, in contrast to traditional *dunno*:

yer, yers  
whadda(ya), whacha, whacher  
did ja, did jer  
wannə, gonnə, gottə, haftə, hastə, usetə, suppostə  
lemme, gimme

shoulda, coulda, woulda, musta, maya, mighta  
'bout, 'cause, c'mon  
'e, 'is, 'im, 'er, 'em  
kin (can), kant (can't)  
git (get)  
donno (don't know)  
shouldna, couldna, wouldna

Naumann (1986) uses a mix of traditional notation and spellings of his own invention. Like Cook, he makes liberal use of the apostrophe. This book differs from the others in that it is a conversation textbook, not a pronunciation textbook. It therefore contains far more reduced forms than the other three books, and the following are samples taken from the pronunciation glossary at the end of the book:

aboudit, 'ad, adall, adhome, ag'in, ain't  
anether, arncha, arn'they, 'as, 'ave  
beaudaful, bedder, bigger'n, b'n, 'bout  
'cause, cen'imediters, c'n, cook't, coudja  
didja, doin', don'ave, doncha, dudn'nit  
duz, duzit, duzn't, dya, dy'ave, dyuu  
eighdy, evenin', fam'ly, fav'rite, fergot, fronna  
gocha, goin', gotcha, gotta, gran'kids  
hafta, howbowcher, howbowchuu  
howbowt, hundred, husban'  
idiz, idizn't, idn'nit, i'duz, introdusha  
izit, izza, izzat, izzatsoo  
jist, kinda, lader, las', le', le'cha, liddle, lotta  
meecha, meechnu, mornin', nawchet, otta, outa  
perdy, priddy, p'tada, reca'nized, 'round  
'scuse, secont, se'dit, sorta, studn't  
takesha, takideazy, tamara, t'day  
th'n, thou, 'twas, twenny

'uzi', 'uzit, var'edy, vejies  
 wananether, wanna, wenta  
 wha', whacha, whacher, whachuuh  
 whadabowchuu, whadizit, whadja, whady  
 whadyuu, whatsa, whatsamadder, wha'wudja  
 whendja, when're, wheredja, wheredya  
 who'as, why'd, woudja

### Critique and Analysis

Of these four textbooks, the spelling conventions used by Rost (2001) differ the most from traditional ways of writing reduced forms. He uses spellings like *guh-nuh* (going to), *wuh-nuh* (want to), and *haf-tuh* (have to) instead of more common spellings like *gonna*, *wanna*, and *hafta*, respectively. However, the use of “h” after a vowel is a common convention in “pronunciation spellings” (used, for example, in foreign language phrasebooks for English speakers), and often indicates what are traditionally referred to as “short” vowels: *ih* [ɪ], *uh* [ʌ], *eh* [ɛ]. This style of spelling is quite familiar to most native English speakers, but not necessarily to learners of English as a second language, which should cause us to question its appropriateness. It is also a more lengthy and cumbersome way of writing these sounds.

Linking is indicated through the use of hyphens, which are often placed so that the final consonant sound of one syllable is effectively transferred to the beginning of the next: *wuh-zit* (“was it”), *ca-nyuh* (“can you”), *wu-de* (“would he”). In contrast to this, other textbooks often write linked utterances as one word each. Hyphenation undoubtedly enhances readability, but it may also lead to pauses in unnatural places, and the placement of the hyphens in forms like *dih-n-e* (“didn’t he”) seems somewhat arbitrary. Moreover, in Rost, sounds that may be omitted are enclosed in parentheses, which almost assuredly impairs readability. For example: *(d)z(th)at* (“does that”), *do(n)-(t)they*. As in other textbooks, standard English spellings which do not reflect actual pronunciation are mixed together with “pronunciation spellings.” For example, in *bought-ihm*, the verb “bought” is left alone and only the word “him” is changed.

The exact pronunciation of spellings like *dthat* (“did that”) would be difficult for even a

native speaker to accurately predict. Should it be pronounced [diðæt] or should the [d] and [ð] be coarticulated, as they often are by native speakers in this case? Clues may be found in the other reduced forms of “did” used by Rost: *dih-they*, *d(ih)-she*, and *dit* (“did it”). The form *dit* is the only one besides *dthat* that attaches “d” directly to the beginning of a word with no hyphen, but “it” begins with a vowel, which renders the pronunciation of *dit* unambiguous. The parentheses in the form *d(ih)-she* indicate a possible pronunciation of *d-she*, which is presumably [dəʃi]. The presence of the hyphen would seem to rule out the possibility of coarticulation, and the hyphen is the only thing that distinguishes *d-she* from *dthat* (in terms of how the reduced form of “did” is written).

Perhaps most significant is the fact that *they* and *that* both begin with the same phone [ð], so one could reasonably expect matching forms: either *dih-they* and *dih-that*, or *dthey* and *dthat*. The fact that they do not match (i.e. *dih-they* vs. *dthat*) would seem to imply a difference in pronunciation. Although the “ih” in *dih-* is likely intended to indicate [ɪ], the actual sound is almost certainly a more central sound like [i], and reduced vowel sounds like [i] and [ə] are sometimes represented in traditional spellings of reduced forms through the complete absence of a written vowel (or by simply a hyphen or apostrophe). Cook (1991), for example, makes extensive use of forms such as *th't* (that) and *frm* (from), as seen earlier. Therefore, *dthat* could be justifiably interpreted as [diðæt], meaning that the difference in spelling between *dih-they* and *dthat* is merely one example of the lack of logic and consistency mentioned earlier. This is actually the more likely scenario, since coarticulation is an advanced topic not usually broached in pronunciation textbooks.

Another problem with Rost’s notation is the use of “u” for the sound [ʊ]. This sound can be spelled in various ways in English (foot, push, would), but all of these spellings have other pronunciations: shoot [u], lush [ʌ], wound [u] / [aʊ]. It is therefore difficult to write the sound [ʊ] in such a way that the pronunciation is unambiguous, using only the letters of the Roman alphabet. However, Rost’s use of “u” in word-final position is a particularly poor choice, because it is likely to be interpreted by learners as [u]. Rost uses forms like *cu-de* [kɒdi] (“could he”) and *wu-(d)yuh* [wʊ(d)jə] (“would you”), as opposed to the more traditional spelling used for expressions with “could” and “would” in the other three textbooks:

couldjoo, wouldjoo (Cook, 1991)

coulda, woulda (Weinstein, 2001)

coudja, woudja (Naumann, 1986)

Actually, traditional spellings like these are not ideal either, but one advantage is that retaining the original spelling makes the words more immediately recognizable. It should be noted that Naumann diverges slightly from the traditional spelling by omitting the “l”.

As seen earlier, Cook uses at least two distinct notation systems, which is unnecessarily confusing to learners, in and of itself. Moreover, there are problems with both systems, and her method of representing linking is erratic and inconsistent. She sometimes writes utterances as one word (e.g. *mindifai* = “mind if I”), but also sometimes takes the same approach as Rost (2001) in attaching the final consonant of one word to the beginning of the next, albeit without the use of hyphens (e.g. *spee di dup* = “speed it up”). Unfortunately, Cook moves the final consonant even when it is not only entirely unnecessary, but when it results in a consonant cluster (“zl”) that is more difficult to pronounce than it would have been otherwise: *läddəvə merikan zla<sup>(t)</sup>ly* (“lot of Americans lately”).

In one of her notation systems, she uses traditional “pronunciation spellings” like “ay” for [e], “ee” for [i], and “oo” for [u], but not consistently. For example, she uses two different spellings for [u], even for the same word (“you”): *whenju*, *wouldjoo*. Also, she retains the pronoun “I” in most cases, but inexplicably changes it to “ai” in *mindifai*. There are also problems with the other notation system she uses, which is semi-phonetic in its use of the symbols [ə] and [æ]. She uses an older spelling convention (“ä”) for [a], but in at least one instance she adds an “h, resulting in “äh”. Of course, “äh” is another traditional way of writing [a], but to combine two spelling conventions like this is unnecessary and potentially confusing. Learners might reasonably wonder whether there is a difference in pronunciation between “ä” and “äh” , since they are spelled differently.

The diaeresis (or *umlaut*) is used over one other vowel (“ü”) to represent the sound [ʊ]. The use of a diacritic is understandable here, since, as discussed earlier, this sound is difficult to represent unambiguously using only conventional Roman letters, but whether “ü” is the best choice is another matter. Another convention employed by Cook is the use of superscript letters in parentheses to show one of two things: either the glides [j] and

[w] (*the*<sup>(v)</sup> *only*, *to*<sup>(w)</sup> *a*), or sounds which can be optionally omitted (*zla*<sup>(l)</sup> *ly*, *soun*<sup>(d)</sup>).

Cook, like others, mixes phonetic or “pronunciation” spellings (ə, æ, ä, ü) with traditional spelling. For example, the traditionally “short” [ɛ] is represented by “e” before a double consonant (*tell*, *lisənwellən*, *geddidiz*), whereas other occurrences of “e” represent the traditionally “long” sound [i] (*ezile*, *the*<sup>(v)</sup>, *me*). Even this is inconsistent, however, since “e” represents [ɛ] in the following words: *əmerika*, *enyway*. Also, “e” has been left alone in traditional spellings where it represents the schwa [ə]: *pæ:ddern*, *the*, *understæn*<sup>(d)</sup>. Of course, *weidə* and *they* are exceptions to all this, as they involve diphthongs. Even here there is inconsistency, since both contain the sound [ei] (or more accurately [eɪ]), and *they* retains its traditional spelling, whereas *weidə* replaces the “ay” in *way* with the phonetic [ei]. To make matters worse, this sound is written yet a third way: *nintənash'n*.

There are other inconsistencies as well. The “d” is doubled in *pæ:ddern*, presumably an analogous retention of the double “t” in the original spelling, but this is not the case with *pridy* (“pretty”). Note that Cook employs the traditional convention of using “d” to represent the voiceless quality of the intervocalic flap pronunciation [ɾ] of the phoneme /t/. Reduced vowels are alternately represented by either the schwa [ə] or simply by an apostrophe. It could be argued that the apostrophe indicates a syllabic consonant, but this only applies in certain cases, as with the nasal [ŋ], for example (*nintənash'n* [nɪntənɛɪʃŋ]), as opposed to a case like *th'time*, which is almost certainly [ðətɑɪm] or [ð̥tɑɪm]. Adding to the confusion, the apostrophe is also retained in its traditional role of indicating a contraction, resulting in sequences like *I'vb'n*.

The orthographic conventions used by Weinstein (2001) and Naumann (1986) have similar problems, as can be observed by examining the lists of forms provided earlier. Naumann’s forms, in particular, are highly inconsistent and idiosyncratic, and the phonetic notation used in the pronunciation glossary at the end of his book is full of errors. Both [ʊ] and [u] are represented by [u], making it impossible for learners to distinguish the two, which completely defeats the purpose of phonetic transcription. Similarly, [ɪ] and [i] are both represented by [i]. Throughout the glossary, the symbols [a], [e], and [ə] are used instead of the correct [ɑ], [ɛ], and [ʌ], respectively. (The correct [ʌ] is only used in one word: *yup*.) The symbol [æ] is only used in two words (*hafta*, *hassle*), and elsewhere this sound is incorrectly rendered as [ə], or in one case as [a]. The symbol [ə] is therefore used to represent no less than three sounds: [ə], [ʌ], and [æ]. Other errors include the following:

<u>incorrect</u>	<u>correct</u>	
[fevrit]	[feɪvrɪt]	fav'rite
[kɔuz]	[kʌz]	'cause
[dʒə]	[djə]	dya ("do you")
[frɒnə]	[frʌnə]	fronna ("front of")

For the sake of clarity, I have omitted the accent marks used by Naumann to indicate syllabic stress. It should be noted that "did you" can be reduced to either [dʒə] or [djə], but the pronunciation [dʒə] should be spelled accordingly (as *ja*, for example, not as *dya*).

### Advantages of IPA

It is clear that there are numerous problems with most of the existing methods of orthographically representing reductions and rapid speech in English. Traditional "pronunciation spelling" conventions often reflect native-speaker intuitions, and are therefore less useful to learners of English as a second language. When using a pronunciation textbook, a learner should ideally be able to accurately determine the correct pronunciation of anything on the written page, without having to resort to consulting with a teacher or native speaker. The only way to achieve this is to insist on a one-to-one correspondence between sound and symbol. If each sound is written in only one way and each symbol has only one pronunciation, there is of simply no room for ambiguity. The International Phonetic Alphabet (IPA) is of course the best-known example of such a system. Since variations of the IPA are widely used in dictionaries to show pronunciation, learners are often familiar with such notation to a greater or lesser degree, but they are not usually accustomed to reading more than one word at a time. Compare the following three reading passages:

(1)

αἰῤῥοῦ πικρὰν δἰαμερἄκα νἰτἑνεἰῤῥἑν πἄρἑν πρἰρι ἰζἑλι, αἰῥο δἰονλι weiῤῥἑ  
 ḡεῤῥἰζ τἑπρακτἰς αἰῥἑταιμ. αἰῤῥἑν takintue larəvə mḗrəkənz leitliəndḗi  
 telmiḗῥai miziə tuanəstən. eniwei aikud goanənan, baḗi imporrənθἰḡἰz telἰsən  
 wḗlən saungud.

(2)

ɑɪ ʃʊd pɪk ʌp ən ði əmɛrəkən ɪntənɛɪʃən pærən prɪrɪ ɪzəli, ɔldʒo ði ɒnli weɪ tə  
gɛt ɪt ɪz tə prækɪs ɔl əv ðə taɪm. ɑɪv bən takɪŋ tu ə lɑt əv əmɛrəkənz leɪtli  
ən ðeɪ tɛl mi ðæt ɑɪm ɪziə tu ʌnəstænd. ɛniweɪ ɑɪ kʊd go ən ən ɑn, bʌt ði  
ɪmpɔrɪnt θɪŋ ɪz tə lɪsən wɛl ən saʊnd gʊd.

(3)

ɪ shʊd pi kəpən the<sup>(y)</sup> əmɛrɪkə nɪntənəʃ'n pæ:ddɛrn prɪdɪ<sup>(y)</sup> ezɪle, ɔlθo the<sup>(y)</sup> ɒnli  
weɪdə geddɪdɪz t' prækɪs ɔləv th'tɪm. ɪ'v b'n təkɪŋ to<sup>(w)</sup> ə læddəvə mɛrɪkən zla<sup>(t)</sup> lɪ, 'n  
they tɛl mi ðeɪ mɛəzɪə to<sup>(w)</sup> ʌndɛrstænd<sup>(d)</sup>. ɛnɪweɪ<sup>(y)</sup> ɪ kʊd go<sup>(w)</sup> ə nə næn, bu<sup>(t)</sup>  
the<sup>(y)</sup> ɪmpɔr'n<sup>(t)</sup> θɪŋgɪz t' lɪsənweɪlən saʊnd<sup>(d)</sup> gʊd.

The top two passages are my phonetic transcriptions of the third passage, Cook's (1991) example of semi-phonetic notation examined earlier. (For the sake of simplification, I have not included syllabic stress or vowel length in my transcription.) Comparing the passages reveals just how far Cook's notation is from an actual phonetic transcription. In the first passage, I have grouped words to show linking, whereas in the second passage each word stands alone. The second passage is easier to read than the first one, but both are quite a challenge for most learners. However, this is not because phonetic transcription is inherently more difficult to read. Quite the contrary, in one sense it is easier to read than traditional English spelling in that each symbol can be pronounced only way. The only reason that learners find it difficult is because they are unaccustomed to reading it.

The majority of the IPA symbols used to transcribe spoken American English are unmodified letters of the Roman alphabet, and 16 consonants are pronounced the same way they are in traditional English spelling: b, d, f, g, h, k, l, m, n, p, r, s, t, v, w, z. In addition, five vowel symbols are used for the same sounds that they sometimes represent in traditional spelling: [ɑ] in *father*, [e] in *they*, [ɪ] in *penniless*, [o] in *go*, [u] in *rule*. The symbol [j] is pronounced as in German or Swedish, rather than as [dʒ] as in English, and this leaves only four letters of the Roman alphabet (c, q, x, y), all of which are used in the IPA for writing sounds not found in the English language. Only 11 additional symbols are needed to write the remaining sounds of English, and six of them differ only slightly from

the familiar letters of the Roman alphabet: [ɪ] (small capital “i” as in *SIT*), [ʊ] (small capital “u” as in *PUSH*), [ɛ] (actually Greek epsilon, but could be treated as a variation of capital “e” as in *LET*), [æ] (symbol includes “a” as in *cat*), [ŋ] (combination of “n” and “g” as in *sing*), and [ʃ] (elongated “s” as in *sure* or as part of the digraph “sh”). Technically, the correct IPA symbol for the standard English “r” is the approximant [ɹ], whereas [r] represents a trill as in Spanish, but this phonetic detail is usually ignored. Since English does not have a trill sound, there is no reason not to simply use [r].

This leaves only six IPA symbols for learners to memorize: [ʌ], [ə], [θ], [ð], [ɜ], and [ɹ]. The symbols [θ] and [ð] are particularly useful, since the other notation systems examined earlier fail to distinguish between these sounds, as does traditional English spelling. The symbol for an alveolar flap [ɹ] enables us to accurately render the intervocalic allophone of the phonemes /t/ and /d/, unlike the other notation systems, which resort to “d” in spite of the fact that [ɹ] is a different sound from [d]. There are of course at least four additional IPA symbols that can be used in transcribing English: [ɔ], [ɪ], [ɜ], and [ɜ̄]. However, [ɔ] does not appear in the pronunciation of all speakers of American English, so there is absolutely no problem with learners pronouncing “law” as [lɑ] rather than [lɔ]. The reduced vowel [ɪ] is close enough to [ə] that this finer phonetic distinction can be safely ignored in materials designed for learners of English, and the two symbols [ɜ] and [ɜ̄] which have the diacritic indicating rhoticity can both be replaced by [ɜr] for simplicity.

Of course, the symbols [ɜ] and [ɜ̄] could be regarded as more accurate because in each case the entire vowel sound is rhotacized, as opposed to a linear sequence of [ə] (or [ɜ]) + [r]. Harder to overlook is the fact that [ə] in English represents a maximally reduced vowel, and therefore by definition should appear only in unstressed syllables. Hence, the fine phonetic distinction between [ɜ] and [ɜ̄], as in a word like “murder” [mɜdɜ̄], which is stressed on the first syllable. Another option, therefore, is to use the symbol [ɜ], but eliminate the rhoticity diacritics: [mɜrdɜr]. Pronunciation dictionaries, in fact, differ on precisely this point. The *Longman Pronunciation Dictionary* (Wells, 2000) uses [ɜ], whereas the *English Pronouncing Dictionary* (Jones, 1997) uses [ɜr]. However, because [ɜ] is the mirror image of [ɛ], the latter approach can easily lead to confusion. Compare the names *Murray* [mɜri] and *Mary* [mɛri], for example. This is a good reason for using the rhotacized versions, since the rhoticity diacritic is never attached to [ɛ], thereby making it easy to distinguish it from [ɜ] at a glance. Personally, I prefer the rhotacized symbols,

because they effectively ascribe a special status to these sounds, which are so characteristic of American English, and serve to emphasize the unitary nature of the sounds, which are so challenging for many native Japanese learners.

Two of the most common sounds of English, both affricates, are traditionally written as digraphs in IPA notation: [tʃ] and [dʒ]. Unitary non-IPA symbols with a hacek (or *caron*) can also be used for these sounds: [č] (used in languages like Czech, but can be thought of as a modified portion of the English digraph “ch”) and [ǰ] (pronounced just like “j” in English). Of course, similar symbols exist for the fricatives [ʃ] and [ʒ], namely [š] and [ž]. One disadvantage to using [ǰ] is the fact that it could be confused with [j], as discussed below.

## Conclusion

Clearly, IPA transcription of English should present no particular difficulty for students, provided they are afforded sufficient practice in reading it, and it is my belief that the benefits of using such a system of orthography in pronunciation textbooks outweigh the disadvantages. Given sufficient exposure, students can quickly become as familiar with spellings like *wɔjə* as they are with spellings like *wouldja*. Using a modified form of IPA notation could make it even easier for students to read. For example, the symbol [y] could be used in place of [j], since the vowel sound represented by [y] in the IPA is never used in English, and the symbol [j] can be confusing, since it is pronounced as in German or Swedish, rather than [dʒ] as in English (as mentioned earlier). This would free the symbol [ǰ] to be used to represent the sound [dʒ] if so desired. On the other hand, since [j] is widely used in phonetic notation in dictionaries, it might be best for students to become better accustomed to it. In any case, it could be confusing to use both the symbols [j] and [ǰ], so if the IPA symbol [j] is retained for indicating the “y” sound in “yes,” then the standard IPA digraph [dʒ] should probably be retained for the “j” in “jam.” I prefer to avoid [j] altogether and to use [y] and [ǰ] for the sake of clarity and simplicity.

Here then are some basic English conversation phrases written as linked utterances in the phonetic notation that I personally suggest:

haʊzɪtɡoɪn?

How's it going?

wʌtʃəneɪm?

What's your name?

wɛryəfrʌm?	Where (are) you from?
koʃhɛlpmi?	Could you help me?
wanəgræbəbaɪt?	Want to grab a bite? (= Shall we go eat?)

Again, I have not included syllabic stress or vowel length in my transcription. A narrow transcription is entirely unnecessary, since the purpose of the notation is pedagogical in nature. It should be noted that hyphens could also be inserted for greater legibility: wanə-græbə-baɪt?

Ironically, the very fact that IPA notation (or a modification thereof) is initially so challenging for learners to read is potentially its greatest advantage. This is because it diverges from traditional English spelling to such a degree that it forces students to focus on pronunciation rather than spelling, which could in turn assist them in discarding fossilized preconceptions about pronunciation arising from overexposure to written English and katakana orthography, thereby enabling them to approach spoken English from a completely fresh perspective.

## REFERENCES

- Baker, Ann and Sharon Goldstein. *Pronunciation Pairs: An introductory course for students of English*. Cambridge: Cambridge University Press, 1990
- Cook, Ann. *American Accent Training: A guide to speaking and pronouncing American English for everyone who speaks English as a Second Language*. Barron's Educational Series Inc., 1991
- Jones, Daniel. *English Pronouncing Dictionary*, 15th ed. (edited by Peter Roach and James Hartman) Cambridge: Cambridge University Press, 1997.
- Ladefoged, Peter. *A Course in Phonetics*, 4th ed. Orlando: Harcourt, Inc., 2001.
- Naumann, John P. *Speaking English the American Way (Saa Amerika-go o Hanasô)*. Tokyo: Eiyûsha, 1986.
- Rost, Michael, Robert Stratton and Sakae Onoda. *Check It Out! New Listening in the Real World*. Tokyo: Pearson Education, 2002.
- Shinmura, Izuru. *Kôjien*. Iwanami Shoten, 1998.
- Weinstein, Nina. *Whaddaya Say? Guided Practice in Relaxed Speech*. New York: Prentice Hall Regents, 2001.
- Wells, John. *Longman Pronunciation Dictionary*. Harlow, U.K.: Longman, 1990. 2nd ed. Harlow: Pearson Education, 2000.
- Wong, Rita. *Teaching Pronunciation: Focus on English Rhythm and Intonation*. New Jersey:

Prentice Hall Regents, 1987.