A USAGE-BASED MODEL FOR A UNIVERSAL MECHANISM OF PHONETIC CHANGES

IZUMI YASUI

University of Tsukuba*


Keywords: usage-based and frequency, exemplar model, phonetically gradual sound change, lexical variation, morphologically-conditioned phonological alternation

1. Introduction

Joan Bybee, author of the book Phonology and Language Use (henceforth PLU), is also the author of several other books and articles on phonology, morphology, language universals, and linguistic change. Her approach to phonetic and phonological phenomena and her research perspectives are totally different from those proposed and discussed by universal grammarians in the framework of generative grammar and phonology, though her main goal is not different from that of generative phonology, which is to describe the phenomena shown in (1a–d):

(1) (a) the relations among similar but physically distinct sounds that are nonetheless taken to be ‘the same’ in some sense (allophonic relations);
(b) the relations among variants of morphemes as they occur in different contexts;
(c) phonological units of various sizes—features, segments,

* I would like to extend my deepest thanks to David John Nathan, James Andrew Elwood and two anonymous reviewers for English Linguistics for reading earlier versions of this review article and providing me with many helpful suggestions and invaluable comments on both content and exposition.
syllables, feet, and so on; and
(d) language-specific and universal properties of these relations and units. (PLU: 19)

In fact, Bybee proposes a new theory of phonetics that determines where sound changes occur or where they are more or less encouraged. Traditional descriptions are subsumed under the constraint-based phonological rules that specify the conditions under which a particular sound change will occur. These conditions tend not to be ‘if and only if’ conditions, but ‘only if’ conditions. If a string of sounds meets the condition specified by a rule, it is not necessary for a change to occur. Rules only say that sound changes can occur under certain conditions, but sound changes such as assimilation or vowel reduction are not inevitable. In her model, the acceptability of a phonological form is gradient, and its acceptability correlates with frequency. The more frequent forms show more acceptability, that is, the most acceptable forms are the most frequent ones. The gradient constraint model based on usage, Bybee argues, provides a better account of gradient phonotactic data than traditional categorical linguistic constraints.

Chapter 1 provides a brief introduction of Bybee’s stance and perspectives, of which a basic assumption is ‘that the cognitive and psychological processes and principles that govern language are not specific to language, but are in general the same as those that govern other aspects of human cognitive and social behavior’ (PLU: 17). In Chapter 2, she introduces her model of representation, as proposed for morphology in Bybee (1985), that lexical representations are built and compiled based upon the language users’ experience, and that, according to their compilations, lexical items are stored and manipulated as phonological representations by language users.

This leads Bybee to a frequency-based, or usage-based, model of phonetic change. This usage-based model can be outlined as a system where high frequency items tend to have strength in representations and be easier to access than little-used items which tend to become more difficult to access. In a usage-based model sound changes are not considered to be a binary phenomenon in which a certain sound change is well-formed or ill-formed, but a continuum exhibiting a gradience. Bybee points out that more frequent constructions systematically show higher likelihoods of more reduced pronunciation, and she demonstrates that the frequency effects come to be responsible for sound changes and particularly for determining the possible, likely, or probable cases in
which sound changes are more or less motivated.

Chapter 3 offers crucial evidence supporting the need for phonetic details in lexical and other stored representations. Bybee argues that certain pairs of interdependent features, such as vowel length and nasal-ity, are necessarily present in stored representations. As for phonetic change, she describes the fundamental tendency for high-frequency words to change faster than low-frequency words (PLU: 35).

Chapters 4 and 5 demonstrate how the phonological phenomena treated and explained by traditional phonological theories are accounted for in the usage-based model. Constraints on sound change and phonological patterns in the lexicon are treated in Chapter 4, and Chapter 5 discusses the interaction of morphology with phonology including the differences between morphologically-based patterns and phonologically-based patterns.

Chapter 6 demonstrates that the mechanism of sound changes observed at the word-level can work equally well within a string of words. Longer sequences with high frequency are stored in memory as chunks and processed in the same manner as words.

Chapter 7 studies more varied sized units, including sandhi alternations and French liaison. It is demonstrated that French liaison persists longer in high-frequency contexts, while in low-frequency contexts it tends to be lost. It can be said that French liaison shows the same regularization as sequences of morphemes or words do. French liaison consonants in constructions of higher token frequency are more conservative, while those in constructions with lower token frequency tend to be easily regularized, showing a gradient phenomenon.

In the final chapter, Chapter 8, Bybee delves into the theoretical status of universal tendencies of sound change. She argues that a diachronic component is essential to any attempt at elucidating linguistic structures or revealing universals of language.

Throughout the book, Bybee explores the potential of the usage-based model to untangle the difficult issues of sound changes, productivity, regularization, and historical change in various levels of phonetic phenomena.

2. Discussion

2.1. Usage-Based Model

Some arguments will be developed in this section about which
aspects of a usage-based model of phonetic change are promising and in which aspects it encounters examples that it cannot adequately explain.

The constraint-based theory of grammar, phonetics and phonology and the usage-based model show very different stances toward language. If logic-based theories are one side of a coin, then usage-based models will be the other. These two frameworks can never meet, nor can we maintain them simultaneously, because one focuses on logic-based aspects of language, while the other is based on statistic aspects of language usage. What Bybee is enterprisingly trying to do is to find out where these two polarities can meet and explicate something about their object of description (i.e. language)—not about themselves (i.e. theories).

The usage-based model proposed by Bybee is quite distinct from any version of generative theory. In generative theory, grammatical structures constitute an autonomous formal system and linguistic data should be explicated in terms of minimalist or module-based account in which linguistic knowledge is explicated by 'the complex array of theoretical apparatus featuring specialized devices for the various 'components' of the linguistic system' (Langacker (2000: 288)). In the usage-based model it is assumed that 'many structures are learned as established units even when they also follow from general principles' (Langacker (2000: 262)). As for nouns a singular form and its plural form are both conventional linguistic units and one is not considered to be derived from the other. The usage-based model takes the maximalist account and a non-reductionist viewpoint for explaining linguistic knowledge along the line suggested in cognitive grammar. The rules and constraints should be determined in a bottom-up manner, not in a top-down manner, unlike in generative theory. In phonetics in the framework of the usage-based model, a string of words is analysed as it is, as a unit or gestalt in cognitive grammar.

Nowadays, corpus linguistics is making a great contribution by providing various kinds of statistical data about language usage. However, it has not made as much contribution to logic-based linguistic theories as we should expect, though several points have been clarified in regard to the relation between the two types of theory. For example, in the field of morphology, analogical leveling is resisted by high frequency items: the present and the past forms of high frequency words such as *keep/keeped and *sleep/slept are never changed into
*sleep/slept. On the other hand, words with low frequency, such as weep/weep, creep/creep, and leap/leap show a tendency to be regularized to weeped, crepeed, and leaped, respectively. In the field of syntax, high frequency discourages grammatical and analogical change, while low frequency renders items less conservative in these changes (PLU: 12).

On the other hand, in the field of phonetics and phonology, the situation is totally different from that in morphology, syntax, and semantics. In fact, the opposite tendency is observed. 'High frequency encourages phonetic change' (PLU: 12), or, 'phonetic change is more advanced in high-frequency words' (PLU: 42). In addition, it is discouraged in low frequency words. This means that language usage is a crucial factor that conditions phonetic behavior; that is, whether the occurrence of a certain sound change is more or less likely. Sound changes are controlled and determined by the muscular activity or muscular gestures taking place during the pronunciation of a certain word or a particular string of words.

2.2. I Don’t vs. I Dunn

It is generally known that I don’t is sometimes changed into I dunn, though we are not sure in which context the phonetic change is more likely or less likely to occur. Bybee’s answer to this question is that it is collocational frequency that determines it (PLU: 161). I found this to be one of the most intriguing examples she adduced for the usage-based model. It is argued that when I don’t occurs together with know, it is very likely that I don’t becomes I dunn, since know occurs with the highest frequency in this collocation. Bybee points out that the verbs occurring with dunn include, in order of frequency, know, think, have, want, like, mean, feel, and care. This order of frequency is a little different from the order that I found in the BNC corpus. The BNC corpus gives the following order:

<table>
<thead>
<tr>
<th>Verb</th>
<th>Total Number of Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>I don't know</td>
<td>11918</td>
</tr>
<tr>
<td>I don't think</td>
<td>6995</td>
</tr>
<tr>
<td>I don't want</td>
<td>3247</td>
</tr>
<tr>
<td>I don't like</td>
<td>1540</td>
</tr>
<tr>
<td>I don't have</td>
<td>743</td>
</tr>
<tr>
<td>I don't care</td>
<td>611</td>
</tr>
<tr>
<td>I don't mean</td>
<td>309</td>
</tr>
</tbody>
</table>
I don’t feel

It can be said, as Bybee points out, that phrases or a particular string of words used more frequently will more freely undergo vowel reduction. They behave phonologically as if they constitute a single word. Moreover, it is certain that far less frequent collocations such as I don’t recommend / comprehend / dislike in (3) are unlikely to undergo the vowel reduction in question.

(3) a. And so I don’t recommend you do it.
   b. I don’t comprehend the reason either.
   c. I don’t dislike food.  (BNC)

2.3. Rhythm Changes in Phrases

Bybee claims, along the line suggested by Hammond (1999), that ‘low-frequency words tend not to undergo rhythm changes in phrases’ (PLU: 43). There seem to be some examples that defy this claim. As for determining the possibility of occurrences of rhythm changes, let us consider whether or not it is adequate and tenable to associate rhythm changes with token frequencies. Or, to what extent is the frequency-based approach to the Rhythm Rule adequate? May we conclude that what is responsible for sound and rhythm changes is the frequency of the strings of words or muscular gestures that occur during the pronunciation of particular expressions?

It is true that the Rhythm Rule usually applies to the phrases in (a) and changes them into (b) in (4)–(6).

(4) a. Japanese
   b. Japanese students

(5) a. He is well-known.
   b. He is a well-known man.

(6) a. The stove is red-hot.
   b. a red-hot stove

However, it is well-known that while the stress in words like terrific, comic, and problematic is determined by the suffix -ic, which assigns stress to the penultimate syllable, there are exceptions such as catholic, heretic, lunatic, politic, and rhetoric which have stress on the first syllable. They cannot undergo the Rhythm Rule. It is true that frequency considerably affects these rhythm changes but it is not the sole factor that controls them.

Let us consider (7a) and (7b):

(7) a. analytic thought
b. analytical thought
In the framework of Liberman and Prince (1977), (7a) and (7b) are assumed to follow the Rhythm Rule because they both involve a stress crash as their rules define it. However, \textit{analytic thought} undergoes the rule to get eurhythmy (\textit{\'{a}nalytic th\'ought}), while the Rhythm Rule does not apply to \textit{analytical thought}. We have \textit{an\'{a}lytical th\'ought} but not \textit{\'{a}nalytical th\'ought}.

Hayes (1984) proposes two rules in terms of the assumption that the number of the desired syllable interval is four and incorporates data as in (7) in the tree-and-grid theory in phonology. One of the two rules is the Quadrisyllabic Rule:

\begin{equation}
\text{(8) Quadrisyllabic Rule}
\end{equation}
A grid is eurhythmic when it contains a row whose marks are spaced close to four syllables apart. (Hayes (1984:46))

The Quadrisyllabic Rule is proposed as a principle of grid scansion: a grid should be ‘searched for the row that best fits the rule’s description, and eurhythmty is proportional to how close the marks come to the desired four-syllable interval’ (Hayes (1984:46)).

The other rule is the Disyllabic Rule:

\begin{equation}
\text{(9) Disyllabic Rule}
\end{equation}
The domains delimited on the level of scansion should be divided evenly by a mark on the next lower grid level. (Hayes (1984:48))

The two rules guarantee that rhythmic intervals should always be equally retained both for the level of scansion (defined by the Quadrisyllabic Rule) and for the level below it as well (the Disyllabic Rule). Hayes argues that the Rhythm Rule does not apply to \textit{analytical thought} because if it did, the number of the syllable interval of the stressed syllables of its output would be more than four, which is not the desired syllable interval.

In the usage-based model, Bybee would resort to the difference of token frequency of the adjectives \textit{analytic thought} and \textit{analytical thought} to explicate the difference in their phonetic behavior, and might explain that \textit{analytical thought} tends not to undergo rhythm changes in phrases because \textit{analytical thought} might show lower token frequency than \textit{analytic thought}. In fact, a KwicOnGugle simple search, contrary to expectations, shows that \textit{analytical thought} has a higher token frequency than \textit{analytic thought} does:

\begin{equation}
\text{(10) a. analytic thought} \quad 1,590 \quad \text{(cf. analytic 1,200,000)}
\end{equation}
b. analytical thought 5,680 (cf. analytical 3,670,000)

Hayes presents a further example which his framework can accommodate:

(11) a. hundred thirteen men
   b. one thirteen Main street

The Rhythm Rule applies to *hundred thirteen mén* in (11a), yielding *hundred thirteen mén*. In (11b) it does not apply to *thirteen* since the stress of *Main* remains as it is because the desired intervals are not created. Thus, in (11a), both the Disyllabic Rule and the Quadrisyllabic Rule apply because it meets the Disyllabic Rule, while (11b) undergoes the relevant adjustment only once because it does not meet the Disyllable Rule (Hayes (1984: 49–50)).

Gussenhoven (1991) presents counterexamples to Hayes’ proposal, one of which is the following.

(12) with ethnic Chinese backing

In (12), *Chinézé* would undergo the Rhythm Rule yielding *éthnic Chinézé bácking*, but actually the Rhythm Rule cannot apply to (12). The proper stress pattern of (12) is not *with éthnic Chinézé bácking* but *with éthnic Chinézé bácking*.

Further lexical variation can be observed. Some lexical items tend to resist the application of the Rhythm Rule, while others easily allow it. For example, words like *morning páper, Easter hóliday, garden séát, lawn tén尼斯, paper nápkin, pound nóte, queen móther, ill-treat*, and *befórehand* usually preserve their original stress pattern, while words like *eagle-éyed, two-fáce, four-légged, absent-mínded*, and *blue-éyed* tend to allow the application of the Rhythm Rule. The applicability of the Rhythm Rule at least partly depends on which lexical item the rule applies to.

The frequency-based phonology in the usage-based model should also explain these examples, as should the other theories.

2.4. Syntax vs. Frequency

2.4.1. Contraction

Is there any case in which the frequency approach based on muscular gestures is blocked? Let us examine some examples of contraction.

Auxiliaries can easily undergo contraction as the sentences in (13)–(20) show:

(13) a. He is a teacher.
    b. He’s a teacher.
(14) a. It is nice.
   b. It’s nice.

(15) a. Harry is at work in the afternoon.
   b. Harry’s at work in the afternoon.

(16) a. There is this much wine in the bottle.
   b. There’s this much wine in the bottle.

(17) a. I am ready to help you.
   b. I’m ready to help you.

(18) a. Bill is rich these days.
   b. Bill’s rich these days.

(19) a. The concert is in Royce Hall this evening.
   b. The concert’s in Royce Hall this evening.

(20) a. Though John is rich, I still like him.
   b. Though John’s rich, I still like him. (Yasui (1992))

However, the sentences in (21)–(26) show that this is not always the case. The following examples show representative environments in which the contraction rule is prohibited:

(21) a. Sam’s at work in the morning and Harry is in the afternoon.
   b. *Sam’s at work in the morning and Harry’s in the afternoon.

(22) a. I don’t know how much wine there is in the bottle.
   b. *I don’t know how much wine there’s in the bottle.

(23) a. Ready I am to help you.
   b. *Ready I’m to help you.

(24) a. Sam’s richer than Bill is these days.
   b. *Sam’s richer than Bill’s these days.

(25) a. Tell Harry where the concert is this evening.
   b. *Tell Harry where the concert’s this evening.


In (21)–(26), the auxiliaries and the verb be have stress, while they do not have stress in (13)–(20). In (21)–(26) why do these items receive stress in the relevant position? Why is the application of the contraction rule blocked and why cannot we use the contracted (or weakened or enclitic) form in the position in question? The auxiliary that would otherwise be weakened is immediately followed by a zero substitute (an empty category or a trace, denoted as $\emptyset$). For example, Harry’s in (21b) is followed by an empty category because $\text{*Harry’s } \emptyset$ in the
afternoon is derived through *Harry is at work in the afternoon* by a deletion rule and, in (21b) *I don’t know how much wine there’s \( \emptyset \) in the bottle* is derived by wh-movement from the underlying sentence *I don’t know [there’s (HOW MUCH) wine in the bottle]*. In these cases the auxiliaries and the verb be have stress on them, which tells us that they are followed by an empty category or a trace. The unpronounced traces left behind after the application of the front-shifting of an interrogative, exclamatory or relative expression (including relatives incorporated into than clauses or as in comparative constructions) are responsible for blocking the contraction. Thus a purely phonological, frequency conditioning explanation is not sufficient to cover the facts of the permitted contractions in (13)–(20) and the prohibited contractions in (21)–(26). The restriction on the use of contracted forms of auxiliaries is determined by phonetic factors coupled with syntactic factors (see Yasui (1992); King (1970)), and not solely by usage and frequency. Moreover, syntactic factors do not constitute the necessary condition for the non-occurrence of contracted forms. The contracted forms can be allowed in the examples in (27) in which an intervening gap (\( \emptyset \)) exists between a lexical category and an auxiliary.

(27) a. The man you met \( \emptyset \)’s just arrived. (has)
   b. The man you met \( \emptyset \)’s making an awful fuss. (is)

(Kaisse (1983:97))

As Zwicky (1970) points out, not all auxiliaries can undergo reduction in the same way. They exhibit a different ease of reduction and show lexical variations. For example, in (28)–(29), *is* and *has* are more freely reducible and the modals or other persons and tenses of verbs *be* and *have* cannot attach phonologically to a preceding word in a relative clause.

(28) a. The man who ate’s been feeling sick. (has)
   b. The man who ate’s going to get sick. (is)

(Kaisse (1983:97); Kaisse (1985:44))

(29) a. *The people who cry’ve been there. (have)
   b. *The people who cry’ll be there. (will)
   c. *The people you know’re there. (are)

(Kaisse (1983:98); Kaisse (1985:44))

Kaisse (1985) observes several other syntactic constraints we should resort to for blocking contracted forms. The sentences in (30) allow auxiliary reduction as shown in (31).

(30) a. Our star reporter has been speaking tonight.
b. The entire remaining crew is standing on the dock.
c. The first-string quarterback is looking on.

(31) a. Our star reporter’s been speaking tonight.
b. The entire remaining crew’s standing on the dock.
c. The first-string quarterback’s looking on.

However, neither has in (30a) nor is of the sentences in (30b, c) will undergo auxiliary reduction in the subject-aux inverted versions of (32) derived by Participle Preposing applying to (30):

(32) a. Speaking tonight has been our star reporter.
b. Standing on the dock is the entire remaining crew.
c. Looking on is the first-string quarterback.

(33) a. *Speaking tonight’s been our star reporter.
b. *Standing on the dock’s the entire remaining crew.
c. *Looking on’s the first-string quarterback.

(Kaisse (1985: 47))

Kaisse (1985) proposes the Government Condition (34), the violation of which explains the ungrammaticality of (33).

(34) The Government Condition:
Auxiliaries may cliticize only onto a constituent that they govern. (Kaisse (1985: 47))

The sentences in (35) can undergo auxiliary reduction.

(35) a. Jack is/’s considered easy-going not even in New York.
b. It is/’s possible to get KUOW on my radio only at night.
c. Louis is/’s not only smart, he’s also a great guy.
d. He is/’s never to darken my doorstep again.
e. He has/’s never been known to do such a thing.

(Kaisse (1985: 51))

However, the application of auxiliary reduction is impossible in the sentences of (36) in which the negated constituent is moved to sentence-initial position by the application of the Negated Constituent Preposing.

(36) a. Not even in New York is Jack considered easy-going.
b. Only at night is it possible to get KUOW on my radio.
c. Not only is Louis smart, he’s also a great guy.
d. Never is he to darken my doorstep again.
e. Never has he been known to do such a thing.

(37) a. *Not even in New York’s Jack considered easy-going.
b. *Only at night’s it possible to get KUOW on my radio.
c. *Not only’s Louis smart, he’s also a great guy.
d. *Never’s he to darken my doorstep again.
e. *Never’s he been known to do such a thing.

(Kaisse (1985: 51–52))

(37) reminds us of the NP-Host Condition discussed in Kaisse (1983). However, the NP-Host Condition cannot explain the well-formedness of the sentences in (38) and (39). We have a further condition on the host NP, that is, the NP must c-command an auxiliary, since the NPs embedded in prepositional phrases in (38) and (39) do not license auxiliary reduction as has been fully argued in Kaisse (1983).

(38) a. Under this slab is/*’s buried Joan of Arc.
b. Under this slab has/*’s been buried a fabulous treasure.
c. On the wall is/*’s hanging a particularly ugly picture.
d. On the wall has/*’s been hanging a picture of Amy.

(Kaisse (1985: 48))

(39) a. How likely is/*’s it to rain?
b. On which day is/*’s John leaving?
c. In which spot has/*’s he been keeping his money?
d. For what reason is/*’s he doing that?

(Kaisse (1985: 52))

Thus, we need to have the syntactic condition in (40):

(40) Auxiliaries may cliticize only onto a c-commanding NP.

(Kaisse (1985: 53))

It will safely be envisioned that syntactic as well as lexical constraints should be evoked to filter out the non-occurrence of contracted forms and that the usage-based model will be valid for determining the likelihood of optional occurrences.

2.4.2. Want to VS. Wanna

Another example that apparently accords with the usage-based model but also needs syntactic consideration is the change of want to into wanna as in (41):

(41) a. Who did John want to visit?
b. Who did John wanna visit?

However, we do not have a wanna version for (42a):

(42) a. Who did John want to visit Sue?
b. *Who did John wanna visit Sue?

In both uncontracted versions in (41)–(42), want is immediately followed by to, but in (42a) the trace of wh is intervening between want and to, which prevents want to from becoming wanna. Thus, there are
some cases in which the generative process should be considered in order to make the phonological rules operate in a way that will yield the correct result.

The Trace Mechanism Proposal will explain that (43a) has two readings as in (43a-i) and (43a-ii), while (43b) has only one, that is (43a-ii) (, which are first noted by L. Horn (see Lakoff (1970)) and are cited in Postal and Pullum (1978: 3)).

(43) a. Teddy is the man I want to succeed.
   (i) I want Teddy to succeed. (succeed = achieve what one aims or want to do (NOED))
   (ii) I want to succeed Teddy. (succeed = take over a throne, inheritance, office, or other position from someone (NOED))

b. Teddy is the man I wanna succeed.

(42)-(43) and related examples are given a clear-cut explication in terms of the Trace Mechanism Proposal by Chomsky (1976), Lightfoot (1976), Chomsky and Lasnik (1977) and others.

However, wanna is not the only contracted form in English. Postal and Pullum (1978) show an exhaustive list of the verbs that allow a contracted form in English.

(44) a. want + to = wanna
    b. going + to = gonna
    c. have + to = hafta
    d. ought + to = oughta
    e. used + to = usta
    f. got + to = gotta
    g. supposed + to = supposta

(Postal and Pullum (1978: 2))

Although Postal and Pullum (1978) claim that the list of (33) is exhaustive, we have other forms like those in (45), in which the infinitival to is phonologically merged with a verb. They are derived through a process, called to-encliticization, which is at least superficially similar to wanna-formation.

(45) a. tryin' + to → trynna
    b. plan + to → planna

(Suiko (1978: 312))

Moreover, we have other serious examples in which eligibility of encliticization differs between the rhyming pairs of predicates.

(46) a. You ought to (oughta) be a secret agent.
b. You are thought to (*thoughta) be a secret agent.

(47) a. I want to (wanna) play the bagpipes.
b. I am wont to (*wonna) play the bagpipes.

(48) a. The beast used to (usta) provide amusement for the
masses.
b. The beast refused to (*refusta) provide amusement for
the masses.

(49) a. He was supposed to (supposta) go out for a walk.
b. He proposed to (*proposta) go out for a walk.

(Postal and Pullum (1978: 2))

A simple search on the KwicOnGugle website reveals that the constructions which allow encliticization do not always show higher frequency than those which do not. The number of each construction we have from the search is:

(50) a. ought to 1,870,000
    b. thought to 2,360,000
(51) a. want to 8,420,000
    b. wont to 180,000
(52) a. used to 5,870,000
    b. refused to 1,960,000
(53) a. supposed to 247,000
    b. proposed to 1,370,000

It is true that (51) and (52) show the high frequency effect, but (50) and (53) show a reverse result.

Syntax and lexical variations will be the necessary condition for its occurrence. The usage-based model in which the difference of eligibility is explained in relation to frequency, (that is, encliticization is more likely to occur in constructions with higher frequency,) faces some difficulty in giving an elegant solution to the data given above.

2.5. Palatalization

Word level palatalization as in (54) would be a good example for the usage-based model proposed in the book.

(54) ocean, mission, nation, etc.

There also are words including intellectual, individual, and educate for which palatalization is optional depending on the context in which they occur. A usage-based model would give us an optimal explanation for accommodating these examples.

Apparently, for phrasal level palatalization, the usage-based model
also seems to work well. In (55a) want you is very frequently palatalized to yield wanchu [wəntʃu].

(55)  
   a. I want you to do that.  
   b. I wanchu to do that.  

However, palatalization will be blocked if a speaker says (55a) while pointing at the hearer and pronouncing the pronoun you with a contrastive or a deictic stress. It would not be a convincing augment in the framework of the usage-based model that emphatic stress would be less frequent than non-emphatic stress and that this would be the reason for blocking. To explain these facts the model should incorporate pragmatic factors to determine the blocking of palatalization. The usage-based model cannot accommodate these pragmatic problems as it is.

2.6. Metanalysis

Metanalysis is a term to describe the phenomenon of a historical miscategorization (Jespersen (1909–49 Vol. 2: 141–142). The historical sound changes are exemplified in (56):

(56)  
   a. a napron > an apron  
   b. an ekename > a nickname  

Apron used to be a word beginning with the consonant n (i.e. napron) and taking the indefinite article a. It underwent a historical change in which the n came to be taken as the last part of the indefinite article an. Similarly ekename changed into nickname, although an apple didn’t change into *a napple, and a navel orange did not change into *an avel orange. A usage-based model should explain these historical sound changes caused by metanalysis even though no such sound changes occurred in history in terms of frequency. It may not safely be concluded that the different degrees of historical sound-changeability of various words are equally subsumed under the usage-based model of sound changes.

2.7. Articles

In English a vowel-initial noun takes the indefinite article an and a consonant-initial noun takes the indefinite article a. Why is the final [n] of the indefinite article deleted before a consonant, while it is retained before a vowel? Bybee argues that ‘the retention of [n] before a consonant, was due to the frequency processing of an + a vowel-initial noun as a single unit’ (PLU: 166).

If the frequency of an + a vowel-initial noun as a single unit is the
reason for the retention of [n] before a vowel initial noun, why is [n] before a consonant-initial noun deleted even if it shows almost the same frequency as an + a vowel-initial noun? Why is an + a consonant-initial noun unable to become a single unit, while an + a vowel-initial becomes a single unit?

In the beginning all the indefinite articles were an, which is a reduced form of the numeral one and an older and fuller form of a. The indefinite article an was followed by a noun with an initial consonant as well as a noun with an initial vowel. An began to sink to a by 1150, yet it was often retained before a consonant initial word as late as 1340. Its history shows a gradual suppression of the an before consonants of all kinds of nouns. (OED, a, a²; an, a³).

Frequency does not seem to be an adequate reason for explaining the historical facts, since there is no reason for preventing an from co-occurring with a consonant-initial word. There are other reasons for the retention and deletion of the [n] of indefinite articles.

I would like to point out that the retention and deletion of the [n] of indefinite articles before nouns observe the tendency to avoid hiatus. When two vowel sounds meet within a word or between words, in English and many other languages, it is often desirable to avoid hiatus using one of several euphonic techniques. Insertion of [r] between there and are in there are, for example, is considered to be one of these euphonic techniques.

The insertion of [r] after [-high] vowels is not the only hiatus breaker in English. The glides [j] and [w] systematically appear to resolve hiatus. After high front vowels such as [i:], [eI], [aI], and [ɔI], the glide [j] is inserted, and after high back vowels such as [u:], [œu], and [au], the glide [w] is inserted. Thus, typical examples of this type of hiatus resolution in English are:

(57) a. The key is [kiːˈjɪz]
    b. The show is [ʃəʊˈwɪz]

The retention and deletion of the [n] of indefinite articles is motivated by the avoidance of hiatus in English in general. The higher frequency processing proposal may have some difficulties in explaining these examples.

2.8. Metathesis and Spoonerism

Does ‘frequency’ always encourage a string of sounds to become a single unit? If it is a fixed and robust tendency, how can we explain
metathesis and spoonerism? Metathesis is the transposition of phonemes; for example, OE hros > ModE horse, OE bridd > ModE bird, and OE thridde > ModE third.

We have another transposition called spoonerism. This is the term for speech errors or sound reversals that involve interchanging parts of (or initial sounds of) two words in a phrase and is named after Reverend William Archibald Spooner (1844–1930), Dean and Warden of New College, Oxford, who is claimed to have said: ‘You have hissed my mystery lectures’ for ‘You have missed my history lectures.’ Other popular attributions to Dr. Spooner are ‘a well boiled icicle’ for ‘well oiled bicycle’ and ‘our shoving leopard’ for ‘our loving shepherd.’

If frequency makes a string of words or sounds become a single unit, these kinds of transpositions cannot even occur, or, even if they occur, they cannot come to be standard usage. However, this is not the case. The usage-based model has to accommodate these sound changes in some way.

2.9. Muscular Gesture

Articulation is a unifying phonological process that incorporates the temporal relations of productions. An abstract characterization of an articulatory event forms the gesture. In other words, gestures are ‘abstractions away from tokens of articulatory events’ (PLU: 83). From the viewpoint of gestures, each word of a language utilizes the same set of muscular events, and is not made up of random combinations of muscular activity (PLU: 83). More pronounceable linguistic forms come from more frequently pronounced clusters of words or sounds. These series of muscular gestures produce segments as emergent units that are stored as units, which leads us to find the reason for the fact that ‘more familiar strings are viewed as more acceptable’ (PLU: 93). This is why English speakers can accurately judge the degree of Englishness of nonce words as well as existing English words.

If sound changes are governed by muscular gestures in a unified way, it may also be claimed that sound changes across languages are subsumed under the same mechanisms, that is, sound changes across languages might be the same or nearly be the same. However, that is not the case.

Considering devoicing of vowels, Japanese vowels can be devoiced in some contexts, for example, [i] in kisha ‘train, reporter,’ kiku ‘hear, chrysanthemum,’ [u] in hukai ‘deep, unpleasant,’ kusuri ‘medicine’ ‘both
[u] in kusuri are devoiced, 'arimasu 'there is,' and [a] in kakato 'heel' (both [a] in kakato are devoiced). In contrast, English vowels are usually not devoiced in any context. If particular sound changes occur as a result of the same muscular gesture, then devoicing in various languages should more or less show the same tendency in some fashion. However, the pictures we have are different from what might be expected. Sound changes including vowel reduction, and, particularly, vowel devoicing, do not seem to accord with a muscular-gesture based model in a unified way. We should consider to what extent a usage-based model is valid if all other conditions are the same.

Let us consider another example that defies a unified solution. The English adverb often, which is derived from oft [ɔft] by adding en, can be pronounced either with or without [t], while the English verb soften, which is also an extended form of soft [ɔft], i.e. soft+en, is nearly always pronounced without [t]. A usage-based model should also give an explanation of this discrepancy, which might need a less elegant solution in Bybee's book.

2.10. French Liaison

Bybee skilfully discusses French liaison in chapter 7, which advances an elaborate and interesting argument along the line suggested in this book and introduces a brand-new approach to its account in the usage-based perspective. French liaison is the sounding of a consonant that is normally silent at the end of a word because the next word begins with a vowel and is an example of sandhi, that is, the process in which the form of a word changes as a result of its position in utterance. Sandhi can be divided into two kinds, internal sandhi that occurs within words and external sandhi which occurs across words. French liaison and the change from the English indefinite article a to an before a vowel are included in external sandhi.

Different examples of French liaison have defied an elegant solution so far and various necessary conditions for establishing and maintaining French liaison have been pointed out based on close observations. The crucial contribution of morphology, syntax, and lexicon to the account of French liaison has been considered to be indispensable and widely recognized for typical French liaison examples as in (58):

(58) a. peti[t] ami 'girlfriend'
    b. les[z] arbres 'the trees'
    c. vous[z] amiex 'I love you.'
d. un bon[n] élève ‘a good student’
e. J’ai vingt[t] ans. ‘I’m twenty’

However, the important and crucial role of frequency in establishing and maintaining French liaison that Bybee proposes has seldom been taken up seriously to explicate its behavior.

Bybee considers French liaison as ‘the deletion of word-final consonants’ before the initial consonant of another word and ‘only a specific instance of the more general deletion of syllable-final consonants’ (PLU: 168), which can be entirely explicable in terms of phonetic conditions. She also points out the occurrence or non-occurrence of French liaison is not a binary phenomenon but its productivity and entrenchment show degrees of gradience,—in fact, French liaison occurs more or less freely according to different constructions.

The important factor in the establishment and maintenance of liaison is the frequency of co-occurrence of certain words. The mechanism of phonetic changes observed in word-level phonology and phrase-level phonology is also applicable to the level of the construction which can emerge via frequent co-occurrence of more than one word.

The liaison consonant does not consistently appear in certain contexts. For example, the verb être ‘to be’ can appear sometimes with the liaison consonant and sometimes without, as tabulated below.

(59) Table: Number of Instances of Liaison for the Forms of the Verb être ‘to be’

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>NL</th>
<th>Total</th>
<th>Percentage of Liaison</th>
</tr>
</thead>
<tbody>
<tr>
<td>est (3rd Sg. Pres. Ind.)</td>
<td>2,591</td>
<td>77</td>
<td>2,668</td>
<td>97%</td>
</tr>
<tr>
<td>sont (3rd Pl. Pres. Ind.)</td>
<td>242</td>
<td>38</td>
<td>280</td>
<td>86%</td>
</tr>
<tr>
<td>étant (Pres. Part.)</td>
<td>22</td>
<td>7</td>
<td>29</td>
<td>76%</td>
</tr>
<tr>
<td>était (3rd Sg. Impf.)</td>
<td>272</td>
<td>95</td>
<td>367</td>
<td>75%</td>
</tr>
<tr>
<td>êtes (2nd Pres. Ind.)</td>
<td>24</td>
<td>10</td>
<td>34</td>
<td>71%</td>
</tr>
<tr>
<td>étaient (3rd Pl. Impf.)</td>
<td>36</td>
<td>21</td>
<td>57</td>
<td>63%</td>
</tr>
<tr>
<td>sommes (1st Pl. Pres. Ind.)</td>
<td>43</td>
<td>31</td>
<td>74</td>
<td>58%</td>
</tr>
<tr>
<td>suis (1st Sg. Pres. Ind.)</td>
<td>65</td>
<td>74</td>
<td>139</td>
<td>47%</td>
</tr>
<tr>
<td>serait (3rd Sg. Fut.)</td>
<td>17</td>
<td>24</td>
<td>41</td>
<td>41%</td>
</tr>
<tr>
<td>soit (3rd Sg. Pres. Subj.)</td>
<td>22</td>
<td>32</td>
<td>54</td>
<td>41%</td>
</tr>
<tr>
<td>j’étais (1st Sg. Impf.)</td>
<td>6</td>
<td>23</td>
<td>21</td>
<td>21%</td>
</tr>
</tbody>
</table>

(PLU: 180)
The quite close correspondence between token frequency and percentage is pointed out with further elegant explanations for apparent counterexamples of *étant* and *suis*. The high percentage of *étant* in spite of its low token frequency is due to its fixed phrasability; in addition, *je suis* 'I am' [ʒe si] will often undergo further reduction to yield [ʃi] in spite of having few instances of liaison.

The close correspondence shows us the tendency of high frequency constructions to retain liaison, while low-frequency forms do not preserve liaison as high frequency does—that is, low-frequency forms will be apt to go undergo regularization. Bybee argues this is the same tendency that is observed within words, that is, the irregular past form of a low-frequency verb is more easily regularized than that of a high-frequency verb is, thus we have regularization for *weep/wept(weep)* but not for *feel/feelt(*feeled)*.

As for the probability of optional liaison, Bybee's usage-based model works tenably and elegantly. The liaison consonant seems to be almost obligatory in examples as in (58). It can be explicated in the framework Bybee proposes. How about (60)?

(60)  a. Il est × une heure. 'It's one o'clock.'
     b. C'est le × onze avril. 'It's April 11.'

The numerals do not occur with a liaison consonant. The usage-based model does not seem to accommodate the numeral examples. In these cases, the semantic explanation based on focus or new information has to be introduced to explain the inapplicability of liaison rules. Frequency cannot explain why *sud-est* 'south-east' and *sud-ouest* 'south-west' undergo the liaison and *nor-est* 'north-east' and *nord-ouest* 'north-west' do not, instead undergoing enchainement.

Consider optional liaison examples as in (61).

(61) cinq francs (five francs)

*Cinq francs* can be pronounced either without a liaison consonant *cinq × francs* or with a liaison consonant *cinq [k] francs*. However, when the non-liaison pronunciation of *cinq × francs* may be mistaken for other expressions like *cent francs* 'a hundred francs,' it is preferable to pronounce it with a liaison consonant. In this case, a pragmatic consideration should be incorporated for determining its occurrence and non-occurrence.

Even if there are some aspects other than frequency in the liaison phenomena, Bybee's arguments are more or less valid in that the frequency of co-occurrence is the most important factor in the likelihood
of maintenance of liaison in French. The fact that the appearance of the liaison consonant even under the condition in which either a pause or a hesitation syllable separates the two words of a construction tells us that phonological analysis will be the best way to solve the recalcitrant behavior of liaison.

The observation of liaison in French shows that words, phrases, and constructions all share the same property because word-level phonology can also occur within frequent phrases and constructions, which suggests that they are storage and processing units. Again, Bybee gives us a convincing suggestion that storage units with high token frequency are conservative, while those with low token frequency tend to be regularized, that is, to be pronounced without a liaison consonant.

3. Conclusion

The usage-based model proposed by Bybee has given us a stimulating and intriguing way of revealing an answer to the recalcitrant problem of why some sound changes are more likely to occur than others. Until now the problem, except for those examples where a logical and theoretical approach can provide a fuller explication, has not been seriously engaged or has been put aside because of its difficulty. We face having to work out a solution that accounts for the phenomena handled by logical and statistic approaches. Bybee’s enterprising book intends to solve this challenging problem.

Bybee’s usage-based model in cognitive linguistics has shown that linguistic constructions should not be reduced to several modular components but should be captured as they are in a bottom-up manner because the whole behavior of language cannot be captured as the array of the components. High frequency constructions will tend to undergo sound change and to retain the change as it is, while low-frequency constructions will be apt to undergo analogical regularization or leveling which high frequency items tend to resist.

Cognitive phonology and phonetics are a less advanced area in cognitive linguistics compared with other areas in the field such as cognitive semantics, syntax, and morphology. Bybee has given an outstanding contribution to the less advanced area in cognitive linguistics. Bybee’s argument in this book is on the right track to the solution in many areas of sound change but, as has been pointed out, there still remain some facts that the framework will need to accommodate in the near
future. However, the contribution of Bybee’s book to the new field of phonology more than addresses my concern.

REFERENCES


Dictionaries

Corpora
British National Corpus World Edition [BNC]
KwicOnGugle (http://163.136.182.112/xyz01/ by Hiroaki Sato)

Humanities and Social Sciences
Modern Cultures and Public Policies
University of Tsukuba
Tennodai 1-1-1, Tsukuba
Ibaraki 305-8571
email: yasui@sakura.cc.tsukuba.ac.jp