Frequency and Other Effects on Diatonic Stress Shift[1)]

Ryuichi HOTTA

Summary

The paper addresses the historical development of the diatonic stress pattern in Modern English. Previous studies have revealed that the diatonic pattern, as exemplified by *récord* (n.) vs *recórd* (v.), appeared first in the late sixteenth century and have since grown to this day in a way that linguists often refer to as Lexical Diffusion. One question of theoretical importance about the diffusion of the diatonic stress pattern is why some words turn diatonic earlier than others. In her "Word Frequency and Lexical Diffusion," Phillips proposed the frequency effect as a contributing factor that determines the schedule of the diffusion, indicating that the least frequent words turned diatonic first. Following Phillips's proposal, in this paper I attempt to reevaluate the frequency effect on the diffusion, making three points in particular. Firstly, my independent survey on the historical growth of diatones with reference to the word frequency database *CELEX2* has made clear something unexpected: the most frequent words turned diatonic first. Secondly, in order to resolve the apparent contradiction, I argue that it is necessary to take account of the interplay between various effects including the phonetic effect and the prefix effect rather than assuming only the frequency effect as predominant. Thirdly, I point out theoretical problems that one must face when one examines a possible correlation between language change and the frequency effect.

Keywords

language change, lexical diffusion, stress, frequency, Modern English

1. Introduction

Over the last four centuries, Modern English has seen a noticeable growth of the diatonic stress pattern in disyllabic noun–verb homograph pairs. Diatones are stress–alternating pairs with the stress falling on the first syllable in the noun ("paroxytonic") and on the second in the verb ("oxytonic"). A glance at a few typical pairs would be sufficient to recognise the kind of accentual pattern in question: *cónvict* (n.) vs *convíct* (v.), *pérmit* (n.) vs. *permít* (v.), and *récord* (n.) vs *recórd* (v.). Examples abound, as listed fully in Appendix.

There have been several studies made on the subject from a phonetic, psycholinguistic, and diachronic point of view. The main interest of the present paper is diachronic, and it attempts to enlarge on the previous studies made by Sherman and Phillips, among others, that focused on the way that the diatonic stress pattern has developed over the Modern English period from about 1570 to the present. The pioneering work on the subject is Sherman's survey on the diffusion of diatones with reference to dozens of historical dictionaries. Inspired by Sherman's diachronic investigation, Phillips then developed a theoretical account of the way that diatones have grown. Her theoretical base is usage–based grammar and focuses on the role of word frequency in determining the schedule of linguistic diffusions. She has revised her frequency–based account as she has found fresh evidence that necessitates reconsideration, but the fundamental tenet, as applied to the diatonic stress shift, remains that the least frequent words developed the diatonic stress pattern first.

The purpose of this paper is to address three issues left largely unexplored

in Hotta's previous treatments on the subject ("Continuing Lexical Diffusion," "Nineteenth-Century Development," and "Diatonic Stress Shift"). Firstly, I wish to reevaluate Phillips's proposed frequency effect on the schedule of the shift since Hotta's quick survey, as reported in "Diatonic Stress Shift," suggested that the shift involved a more complex process than should be expected from Phillips's hypothesis. Secondly, as any linguistic change must be a result of several factors interacting with one another, so the diatonic stress shift must be attributed to various conditioning factors, of which the frequency effect should be only one. The question then is what factors may have been involved, besides the frequency effect, in the growth of the diatonic stress pattern. What must be assumed is, therefore, the interplay between various effects that likely motivated the stress shift. Thirdly, I would like to address general questions as to how frequency-based studies can be developed when applied to linguistic changes that show lexical diffusion. In particular, I will point out theoretical problems concerning word frequency lists and the complex nature of observing language change in progress.

2. Previous Studies

2.1 Sherman

Sherman's paper was a descriptive work on the historical growth of diatones in Modern English with reference to dozens of dictionaries and other lexicographical works published in the sixteenth to eighteenth centuries. Among the several findings of Sherman's, four observations are especially relevant to the present interest. Firstly, consulting Present-Day Standard English dictionaries reveals, perhaps unexpectedly, that the diatonic stress

pattern is *not* a dominant pattern with noun–verb homograph pairs. Table 1, based on Sherman's figures, shows counts of relevant pairs for four logically possible stress patterns of noun–verb homograph pairs, as printed in Hotta ("Diatonic Stress Shift" 2).

Table 1: Stress Patterns of Noun–Verb Homograph Pairs in PDE as Given in Hotta ("Diatonic Stress Shift" 2) Based on Sherman

Type 1–1: 950 (72.24%) (*prómise – prómise*)	Type 2–2: 215 (16.35%) (*resúlt – resúlt*)
Type 1–2: 150 (11.41%) (*récord – recórd*)	Type 2–1: 0 (0%) (no example)

Secondly, the distribution of the different patterns for PDE, as shown above, is a result of historical development. As Sherman remarks, "during the 17th and 18th centuries there were fewer noun–verb diatones than at the present time" (53). This is to say that diatones have been surely increasing in number since the seventeenth century at least.

Thirdly, Sherman noted, "the creation of stress alternation is more likely to occur as stress-retraction in an oxytonic pair than to occur as stress-advancement in a paroxytonic pair" (53). In other words, noun–verb homographs of the *resúlt* type (Type 2–2) have historically been the major source of innovative diatones.

Fourthly, as Sherman plotted the growth of diatones diachronically from the late sixteenth century onwards, he recognised a pattern reminiscent of Lexical Diffusion, a process in which language change starts slowly, speeds up at a "take-off" point, and then slows down again towards the end with a long tapering tail. At the end of his paper, Sherman even listed 213 noun–

verb pairs of Type 2–2 that he predicted would be turning diatonic in future, but his prediction as well as his mention of the nineteenth- and twentieth-century developments was no more than an extrapolation based on the facts from the preceding centuries and the mid-twentieth century, and therefore remained to be confirmed.[2)]

Sherman's empirical study invites questions of theoretical importance. One key question is how the diffusion is scheduled, on the assumption that the diatonic stress shift is an example of Lexical Diffusion. In other words, what factors play a significant role in determining the schedule of the diffusion, or which words turn diatonic first and which words next. Sherman only noted the gradual growth of diatones and failed to address the question about the schedule, while Phillips tried to make a detailed enquiry into the latter in her series of related studies.

2.2 Phillips

Following Sherman's work, Phillips examined the diatonic stress shift from a more distinctively lexical diffusionist point of view. She investigated it and several other language changes that seemingly proceeded in Lexical Diffusion and attempted to compare the schedules between different diffusions. The comparison led her to propose Frequency Actuation Hypothesis, which stated that "physiologically motivated sound changes affect the most frequent words first; other sound changes affect the least frequent words first" ("Actuation of Sound Change" 336). The revised version of the hypothesis is found in Phillips ("Lexical Diffusion" 231).

> [F]or segmental changes, physiologically motivated sound changes af-

fect the most frequent words first; other sound changes affect the least frequent words first. For suprasegmental changes, changes which require analysis (e.g., by part of speech or by morphemic element) affect the least frequent words first, whereas changes which eliminate or ignore grammatical information affect the most frequent words first.

Phillips considered the diatonic stress shift an example of the least frequent words changing first, reasoning that noun-verb distinction in such words should require morphosyntactic analysis on the part of the speaker. The critical discussion that Phillips made to support this view concerns a significant gap in the average frequency between diatones (Sherman's Type 1–2) and non-diatones (Sherman's Type 2–2). Phillips classified diatones and non-diatones according to the prefix that they had and compared the average word frequencies between diatones and non-diatones in each prefix class. The results were the same for all prefix classes that she considered: diatones showed lower average frequencies than non-diatones.[3)]

Thus far, Phillips's Frequency Actuation Hypothesis, or at least the part of it that proposes that the least frequent words change first, seems to be supported in terms of word frequency. The proposal, however, needs to be tested more carefully now. Firstly, the list of diatones as well as the frequency list that she depended upon may not be the latest or the best available today. She used Sherman's list of diatones, but an updated list of diatones in PDE would be much longer and make a more robust foundation for frequency-based survey. To this end I compiled an updated list of diatones in Hotta ("Diatonic Stress Shift"), as reprinted in Appendix. This is, I believe, the latest and most thorough list of diatones available at present.

As far as word frequency lists are concerned, rather than using Phillips's data source, I will consult the database named *CELEX2* (to be precise, the English component of *CELEX2*), which is grounded on a newer and larger corpus than many others available today.[4] Unlike the Phillips's data source, it gives POS-distinguished lemmata (i.e., nouns and verbs in homograph pairs being distinguished), which allows finer frequency analysis.[5] Although *CELEX2* is one of the best databases available for the present purpose, however, there remain problems that involve frequency-based studies with any frequency list, as I will discuss in Section 6.

The second problem with Phillips's hypothesis is that her account with the average frequencies is static while the object of the study is dynamic by nature. Comparing the average frequencies between diatones and non-diatones is, I must accept, a reasonable way to infer the schedule of the diffusion, but it must remain an indirect method that only leads to a very broad conclusion. A more direct method will be to take "snapshots" at distinct points in time and compare the average frequencies of innovative diatones from period to period. This method, more diachronically oriented than Phillips's, will allow a closer investigation such that it may be made clear which words became diatonic first and which next.

Thirdly, Phillips investigated the diatonic stress shift exclusively in terms of word frequency. This is understandable as her interests were mainly in developing Frequency Actuation Hypothesis in consideration of different language changes. There is adequate evidence, however, to suggest that the shift should be addressed not only in terms of frequency but from a multifactorial point of view that considers other linguistic factors as well. Frequency is arguably only one of many conditioning factors that combine to determine

the schedule. With this view in mind, I would like to propose two particular factors, the phonetic effect and the prefix effect, in Sections 4 and 5. They will give a more balanced account of the schedule of the shift.

Fourthly, Phillips adopted prefix-by-prefix classification of words in comparing the average frequencies, but her argument for assuming prefixes as natural classes is not convincing. Indeed she observes, "That the prefix should be a conditioning factor in these words is not so surprising when one considers that in the left-to-right processing of speech forms, prefixes trigger connections to all other words with that prefix" (*Word Frequency* 36), but I am not certain whether or not she is suggesting any effect that the prefix might have on the schedule of the shift when she makes this remark. Rather than assuming, as Phillips perhaps does, that the prefix effect works cooperatively under the general frequency effect, I propose that the prefix effect contributes independently, side by side with the frequency effect, to determining the schedule.

Before turning away from the review of previous studies to my own survey, I must add that in Hotta ("Continuing Lexical Diffusion" 52) I remarked "it is likely that the shift in stress patterns of words from Type B [Type 2–2] to Type C [Type 1–2] is not significantly motivated by frequency." In a later paper, "Diatonic Stress Shift," however, I addressed again the schedule of the diffusion in terms of frequency with the tentative conclusion that the frequency effect was there but in the opposite direction to Phillips's hypothesis, that is the most frequency words changing first. The analysis then remained limited and superficial, and therefore the present paper will take up the same issue once again, this time more closely.

3. The Frequency Effect

In this survey, word frequency information derives exclusively from *CELEX2*. *CELEX2* is a very large lexicological database of PDE based on *Oxford Advanced Learner's Dictionary* (1974) and *Longman Dictionary of Contemporary English* (1978) for lexical information and on COBUILD/Birmingham corpus of 17.9 million words for token frequency information. COBUILD/Birmingham corpus consists of written text (92.74%) and spoken text (7.26%), and 44 texts out of the 284 that constitute the written corpus are of American variety. The database is composed of eleven sub-databases, some lemma-based and others wordform-based, that each specialise in orthography, phonology, syllable structure, morphology, and syntax.

First of all, let us see whether Phillips's main findings can be supported with *CELEX2*. According to Phillips, the average word frequencies between diatones and non-diatones differed significantly, the former being less than the latter. A survey with *CELEX2* generally reconfirms this significant gap. The average frequency of 138 diatones (out of all the 235) for which both noun- and verb-token frequencies are given in *CELEX2* is 321.65 while that of 124 non-diatones is 426.96.[6)]

Next we will see whether Phillips's prefix-by-prefix analysis can also be supported with *CELEX2*. I calculated the average frequencies with diatones and non-diatones whose POS-distinguished token frequencies are given in *CELEX2* (i.e., 90 diatones out of my list and 86 non-diatones out of the words of Sherman's Type 2–2). The analysis is summarised in Table 2.[7)]

Table 2 makes it clear that in all the prefix classes considered, diatones are less frequent than non-diatones on the average, just as Phillips's findings

Table 2: Average Token Frequency of Diatones and Non–Diatones by Prefix Class

	Average Token Frequency	
Prefix Class	Diatones	Non-Diatones
aC-	204.14	654.94
coC-	355.08	535.37
de-	348.14	504.31
eC-	296.50	631.75
ex-	208.50	596.67
pre-	286.25	386.00
re-	285.57	490.78
suC-	713.13	1119.75

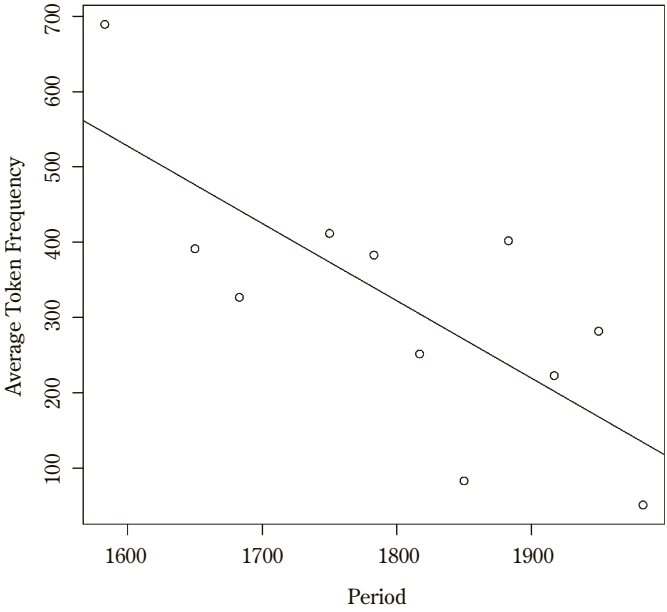
indicate. As mentioned in the last section, however, it is not clear to me how this prefix-based analysis can be interpreted as significant. The average frequency gap for each prefix class seems simply to represent the overall average frequency gap, which means that classification by prefix may not be meaningful because any random grouping should also represent the overall average frequency gap. I do not deny the relevance of prefixes, however, as I will address the prefix effect from a different viewpoint in Section 5.

The analysis of the average frequency with *CELEX2* generally concurs with Phillips's proposal: the gap between diatones and non-diatones seems to have the diachronic implication that less frequent words turned diatonic earlier. The gap, however, is merely an indirect index to the plausible schedule and must be corroborated with a direct, diachronic analysis. To this end, let us now take account of the periods when diatones at present were first attested as such in historical dictionaries. The consideration of the periods of first attestation in combination with token frequency information from *CELEX2* will make it clear whether or not there is any significant frequency effect on the schedule of the diatonic stress shift. I adopt periodisation by

Table 3: Average Token Frequency of Diatones from Period to Period

Period	Average Token Frequency
C16c	689.17
C17b	391.33
C17c	326.70
C18b	411.63
C18c	382.92
C19a	251.50
C19b	83.10
C19c	401.92
C20a	222.83
C20b	281.82
C20c	51.17

Figure 1: Average Token Frequency of Diatones from Period to Period



the third of a century, labelling the period 1601–33 as C17a, the 1734–66 as C18b, and the 1867–1900 as C19c, for example. Table 3 and Figure 1 show the average frequencies from period to period.

Contrary to expectations following Phillips's hypothesis, the figures indicate if anything that more frequent words turned diatonic earlier, with the coefficient of correlation between the period and the average frequency at −0.7564. The result is so contradictory to Phillips's proposal that one cannot help considering why it should be the case. One way to interpret the apparent contradiction is to imagine what may be called frequency zones. Let us assume, say, two frequency zones, the high and the middle-and-low. In general, middle-and-low frequency words tend to turn diatonic earlier than high frequency words, but within the middle-and-low group, more frequent words tend to turn diatonic earlier. In this scenario, the schedule of the diffusion is middle-frequency first, low-frequency second, and high frequency third. The problem with this proposal, however, is that although it is plausible that extremely frequent words can behave differently both from common frequent words and from lower frequency words, it is unclear where a dividing line between high frequency and middle-and-low frequency zones should be drawn.

An alternative solution would be to take the two end points for C16c and C20c out of consideration on the grounds that they show extreme figures. There are only six words for C16c and there are many neologisms, commonly with *re-*, for C20c, which period may well be underrepresented in the *CELEX2* based on the corpus compiled in 1971.

Yet another solution would be to imagine that other factors than the frequency effect are at work behind the scenes, causing an apparent contradic-

tion between the synchronically calculated averages and the periodised averages. I will propose two such factors, the phonetic effect and the prefix effect, which will be explored in the following sections.

4. The Phonetic Effect

I owe to Kelly the idea that phonetic conditions can be relevant to the diatonic stress shift. Kelly associates the rhythmic alternation that characterises the prosody of English with the different syntactic positions that disyllabic nouns and verbs tend to take. He argues that "disyllabic verbs were more likely than disyllabic nouns to receive an inflection that adds a syllable onto the word" and that "[b]ecause such syllables are weakly stressed, rhythmic alternation would be created if the disyllabic word received stress on the second syllable (e.g., 'suggesting') rather than the first ('promising')" (107). Thus disyllabic verbs tend to take inflectional suffixes such as *-ing*, *-ed*, *-(e)s*, the first of which invariably adds a syllable, while the second and third do so if the stem ends in a dental plosive (/t/ and /d/) or sibilant (/s/, /z/, /ʃ, /ʒ/, /ʧ/, and /ʤ/), respectively. On the other hand, the only inflectional suffix that can be added to nouns is *-(e)s* for the plural or possessive, the syllabic status of which depends on the same phonological environment as for its homophonic suffix to the verb.

Immediately relevant to the present discussion is Kelly's following remark: "In particular, noun-verb homographs are more likely to possess contrasting stress patterns if they end in the dental stops /t/, as in 'suspect,' and /d/, as in 'record' " (113). One supporting piece of evidence comes from my quick search of the MRC Psychological Database for all disyllabic noun-verb homograph pairs.[8] The search reveals that about one fifth of

them (585 out of 2,906 pairs) end in a dental plosive. On the other hand, of all the 235 diatones in the present study, those ending in dental plosives amount to 124 (52.76%). The propensity of pairs with dental plosive endings to show a diatonic contour is, therefore, synchronically confirmed.

Then one would be interested to know what diachronic evidence says on this matter. If, as Kelly claims, the stem-final /t, d/ encourages disyllabic verbs to assume (or maintain) an oxytonic contour, it should be supported by diachronic evidence as well. Table 4 shows the type count of diatones with a stem-final /t, d/ from period to period, with their average token frequency added for reference.

In every period, dental plosives that end the stem are abundant. What is more important, earlier diatonic innovators with a final /t, d/ are generally of higher frequency, or so at least until the mid-19th century. The coefficient of correlation between the period and the average frequency is −0.8064, and this suggests that the large class of diatones with a stem-final /t, d/ is

Table 4: /t, d/–Final Type Counts and Average Token Frequency from Period to Period

Period	/t, d/-Final Types	Total Types	Average Token Frequency
C16c	3	6	959.17
C17b	7	9	486.64
C17c	14	18	382.75
C18b	10	15	239.86
C18c	6	21	229.20
C19a	18	29	258.94
C19b	12	20	80.63
C19c	16	33	192.75
C20a	12	19	173.21
C20b	18	49	282.29
C20c	8	16	68.00

the main drive behind the propensity of their homograph verbs to remain oxytonic and possibly, in contrast, the propensity of their homograph nouns to turn paroxytonic, with the result of the unexpected schedule of more frequent words turning diatonic earlier. If one calculates with the rest of the diatones, excluding those with a final /t, d/, the coefficient of correlation is as weak as −0.09782. Phillips regards the diatonic stress shift as "a sound change that clearly did not have a phonetic basis" (332), but this assumption should by now be revised to accommodate itself to the likelihood that a final /t, d/ has an effect on the schedule of the diatonic diffusion.

5. The Prefix Effect

In Section 2.2, I made a critical comment on Phillips's treatment of prefix-by-prefix frequency, but I do not mean to reject the prefix effect *per se*. Rather, it is significantly relevant. Historical evidence shows that some prefixes contributed markedly to a growth of diatones in some periods. Table 5, reproduced from Hotta ("Diatonic Stress Shift" 9), represents how many types of diatone with particular prefixes came into being from period to period.

The table shows that up to C19a the most productive prefix for diatones was *coN-* (e.g., *con-*, *com-*, *col-*), while in C19b and C19c a significant expan-

Table 5: Prefixes that Attracted Diatonic Stress Shift as Given in Hotta ("Diatonic Stress Shift" 9)

	aC-	*coN-*	*de-*	*dis-*	*e(C)-*	*ex-*	*fore-*	*iN-*	*mis-*	*ob-*	*off-*	*out-*	*per-*	*post-*	*pre-*	*pro-*	*re-*	*suC-*	*tra-*	*trans-*	*un-*	*up-*
16c			1					1							1		3					
17b	1	5								1							1					
17c	1	8			1	1		1				1				1		1		1		
18b		5		2		1		2							1	1		1			1	
18c	1		3					2				2	2		2	1	2	1		2		1
19a	2	11			3		2	1				1			1	2	1	1	2			1
19b	1	3			1	1		6				1	1				2					2
19c	2	2	1	1	2			5			1	3		1		2	1	2		1		2
20a		1		2		2		2				1	1		1		2	3		1		
20b	5	1	3	2	2			3	1								26					2
20c				1				1								1	12					

sion is recognised for the prefix *iN-* (e.g., *in-*, *im-*, *il-*). The last century or so has in turn seen an explosion of diatones with the prefix *re-*. Popular prefixes have changed from period to period, but there has clearly been a kind of fashion. This "prefix effect" must be independent from the alleged frequency effect since the former, as a kind of fashion, is likely more of a temporary nature while the latter, as a general principle, is likely more of a stable nature. In particular, a number of *re-*neologisms emerging remarkably over the last decades suggest that the frequency effect plays less of a leading role than the prefix effect because their popularity, and their frequency by implication, may readily wax and wane within a short period of time.

In the above, I have proposed two distinct factors besides the frequency effect that seem to contribute in combination to pushing the diatonic stress shift forward. The frequency effect, as featured by Phillips, is just one of the factors and must be reinterpreted as such in trying to explain the diatonic stress shift diachronically. This leads me to a discussion of theoretical problems concerning frequency-based studies of language change.

6. Problems with Frequency-Based Studies of Language Change

Recent years have seen an increasing number of large-sized corpora and a growing interest in usage-based grammar. Under these circumstances it is small wonder that word frequency has been focused on as a key player in determining the schedule of language change. I do believe in frequency effects on language change, as can be clearly if broadly indicated by the significant gap in the average frequencies between diatones and non-diatones, but there are theoretical problems to be addressed when one considers the likely relationship between word frequency and the schedule of

language change.

One problem with frequency-based study lies in the fact that individual items often deviate largely from the expected behaviour for their fellow items of the same frequency-range. This is why one must often be contented with a broad generalisation derived from the average frequency and give up dealing with the frequency of individual items for finer analysis. It is not difficult to tell why this should be the case. As Phillips noted, "Word frequency, like most phenomena, varies with space and time: the frequency of a particular word will vary slightly from speaker to speaker and from speech community to speech community. It also varies across time . . . " (*Word Frequency* 336). On the other hand, Phillips noted elsewhere, more optimistically, when she investigated another stress shift in Late Modern and Present-Day English that "[t]he words' frequencies are based on present-day English, but the general pattern of relative frequencies probably holds for the English in our data base (1755–1993) as well" ("Lexical Diffusion" 225–26).

One practical solution to the problem about word frequencies ever changing and varying in time and space will be to compile any number of frequency lists that answer to target varieties of the language. In an age of large corpora of different varieties emerging rapidly, it will be increasingly possible to produce such frequency lists semi-automatically, although one must be cautious about the representativeness of corpora.[9)]

Another problem is that, as I have noted in the foregoing sections, the frequency effect on a language change is most likely among several factors that determine its schedule in combination. Too narrowed a focus on this particular effect is to run the risk of overevaluating it and eclipsing other po-

tential factors. In the diatonic stress shift, there seem to have been at least two other effects at work: the phonetic effect and the prefix effect. Considering the interplay between these (and possibly other yet unrecognised) factors would furnish a more cogent account of the schedule of language change and help to have a better-informed evaluation of the frequency effect on language change.

One needs be neither too pessimistic nor optimistic about frequency-based study of language change, however. As far as the diatonic stress shift is concerned, the general pattern suggested by the average frequency (i.e., the least frequent words changing first) is so clear that it can safely be accepted at least as a relevant factor that helps to account for the schedule of the diffusion. At the same time, an enquiry into individual items, even though their deviation from the average increases as analysis becomes finer, should not be neglected, so that we will recognise potential factors other than the frequency effect that should otherwise remain unseen.

7. Conclusion

In the present paper, Phillips's proposal that the least frequent items turned diatonic first was supported in general on the basis of synchronically comparing the average token frequencies between diatones and non-diatones. Nevertheless, when I considered historical data over the past four centuries from a diachronic point of view in consideration of the token frequencies of individual diatonic items rather than the average frequency, I came to the conclusion that the diatonic stress shift did not proceed in such a straightforward way as suggested by Phillips's proposal but even showed the opposite trend, that is the most frequent items turning diatonic first.

This apparently contradictory picture poses a challenging question, to which I reply with three possible suggestions. One interpretation is that while there lies a major division between high frequency words and middle-and-low frequency words, with the high frequency lagging behind in the diatonic stress shift, there lies another division within middle-and-low frequency words, with the low frequency lagging behind the middle frequency. Another way around the difficulty is to suspect that the two end points, C16c and C20c, may skew the overall statistics. The other way of clearing up the apparent contradiction is to suppose the involvement of factors other than the frequency effect that contribute to the situation looking confusing.

It is difficult to tell which suggestion is the best account, but I have tried to explore the third in particular in this paper. One factor that contributed partly to determining the schedule of the shift concerns a stem-final sound. There is evidence, synchronic and diachronic, that the dental plosive, /t, d/, in the stem-final encourages homograph verbs to remain oxytonic and, in contrast, homograph nouns to turn paroxytonic. Considering the phonetic effect and the frequency effect together reveals that more frequent items with a stem-final /t, d/ were among earlier diatonic innovators. Another factor that I recognised as having played a role in the diatonic stress shift is the prefix effect. The evidence shows that particular prefixes have gained popularity from period to period, for example, *coN-* until C19a, *iN-* in C19b and C19c, and *re-* in C20b and C20c. The ups and downs of the prefixes seem to be a matter of fashion and perhaps independent of the frequency effect.

Thus I argue that the frequency effect is only one of several factors that combine to determine the schedule of the diatonic diffusion and that it is necessary to pay attention to the interplay between them, namely at least be-

tween the frequency effect, the phonetic effect, and the prefix effect as they were shown to be relevant in this paper.[10)]

Finally, I addressed theoretical problems concerning frequency-based studies of language change. There remain a number of problems to solve before we can begin to understand how word frequency relates to language change: various frequency lists must be made available; finer analysis than looking at the average frequency must be sought; the interaction of frequency effect with other effects must be focused on. I hope that growing attention to usage-based grammar and its increasing interest in word frequency in recent years will push forward linguistic studies in this direction.

Appendix

The following is an updated list of 235 diatonic pairs in PDE, arranged diachronically, as it appears in Hotta ("Diatonic Stress Shift" 5–6). Words are assumed to be diatonic when at least one of the dictionaries consulted indicates their diatonic status, even secondarily. Periodisation is made by the third of a century, with C17a, C18b, and C19c representing the periods 1601–33, 1734–66, and 1867–1900, respectively, for example.

C16c (6 diatones): *desert, incense, present, rebel, record, refuse*

C17b (9): *accent, collect, compound, conduct, contract, convoy, object, relapse, torment*

C17c (18): *abstract, cement, compact, confine, conflict, conserve, consort, contest, converse, convict, essay, extract, ferment, insult, outcast, project, subject, transport*

C18b (15): *bombard, compress, concert, concrete, confect, contrast, discord, discount, export, import, impress, prelude, produce, survey, undress*

C18c (21): *affix, decrease, defile, descant, digest, increase, inlay, outleap, outwork, perfume, permit, prefix, presage, protest, purport, regress, reprint, surcharge, transfer, transverse, uprise*

C19a (29): *abject, aspect, augment, colleague, combat, commerce, complot, comport, conscript, console, content, context, convent, convert, efflux, entrance, escort, forecast, foretaste, impact, outlook, premise, progress, prospect, retail, surname, traject, traverse,*

upstart

C19b (20): *absent, commune, concord, consult, curvet, ensign, excerpt, imprint, impulse, infix, ingress, inset, instinct, outcry, pervert, reset, revise, turmoil, upcast, upset*

C19c (33): *ally, annex, congress, construct, costume, detail, dictate, dispatch, egress, eject, finance, gainsay, inflow, inlet, insert, invert, invite, levant, masthead, offset, outgo, outpour, outspread, placard, post–date, proceed, process, recount, sublease, surtax, transform, uplift, uprush*

C20a (19): *combine, discard, discourse, excise, exploit, foment, incline, indent, legate, mandate, outstretch, perfect, prefect, rebate, release, sub–let, surmise, suspect, transplant*

C20b (49): *addict, address, affect, alloy, assay, chagrin, control, decline, decoy, defect, discharge, dispute, employ, entail, implant, intern, intrigue, misprint, rampage, rebound, rebuff, rebuild, recall, recess, recoil, re–count, redo, redraft, redress, refill, refit, refund, rehash, reject, relay, remake, replay, report, rerun, research, retake, retouch, retread, reverse, rewrite, romance, sojourn, update, upgrade*

C20c and C21a (16): *dismount, humdrum, prolapse, recharge, recon, refan, regrind, rejig, relaunch, re–let, remould, resit, retard, rethink, retort, segment*

Notes

1) The present study is supported in part by a Grant-in-Aid for Young Scientists (B) for 2012 (No. 21720178) from the Japan Society for the Promotion of Science (JSPS). It was also partly financed by the tokutei-kadai-kenkyu (special research) 2012–13 by Chuo University. I wish to express my thanks to both organisations for their generous financial assistance.

2) For this purpose, Hotta furnished follow-up investigations about the nineteenth- and twentieth-century developments in "Nineteenth-Century Development" and "Continuing Lexical Diffusion," respectively.

3) Phillips's figures, given in her "Actuation of Sound Change" (336), are calculated on the *American Heritage word frequency book* published in 1971 which does not distinguish between noun-frequency and verb-frequency. Note also Phillips's following treatment: "Of those pairs which are already diatonic, I included only those which are known to have developed from final-stressed pairs" (333). For the *a*-prefixed group, the average frequency of diatonic words scores 7.4 while that of non-diatonic words scores 15.8. Likewise, the *con-/com-*, *de-*, *dis-*, *es-*, *ex-*, *pre-*, *re-*, and *sur-* groups each show a gap in the average frequency between diatones and non-diatones:

10.6, 37.1; 5.2, 8.0; 1.6, 4.9; 3.1, 10.6; 2.5, 22.0; 3.7, 8.1; 8.0, 11.1; and 5.5, 24.3, respectively.

4) To be fair, Phillips used the first version of the *CELEX* (1993) for another stress shift study in her "Word Frequency and Lexical Diffusion."

5) *CELEX 2* does come with frequency information of POS-distinguished lemmata, but it is to be noted that such distinct frequency counts are half theoretical in that they were calculated inductively on manual counts of up to a hundred random instances. See Section 3 below for more introductory notes about the database.

6) In this and the following analyses, the token frequency refers to a raw token count in the about 17.9 million-word corpus on which *CELEX 2* is based, not a normalised token count, say, per million words in the corpus. Note also that the noun and verb of each homograph pair are counted separately, whether by token or by type.

7) *C* in the prefix names represents any phonotactically possible consonant for the prefixes. coC-, for examples, represents *con-*, *com-*, *col-*, etc.

8) I made this search and reported the results in Hotta ("Diatonic Stress Shift" 13).

9) One such attempt that I made elsewhere was to compile an Early Middle English word frequency list out of *LAEME* text database. See Hotta's "Representativeness."

10) I should add that Ogura and Wang also emphasised "the interplay between word frequency and the phonological environments" (131) when they addressed the growth of *-s* in the third person singular present indicative of the verb in English.

References

Dictionaries and Corpora

CELEX 2. Ed. Baayen R. H., R. Piepenbrock and L. Gulikers. CD-ROM. Philadelphia: Linguistic Data Consortium, 1996.

LAEME Corpus = Laing, Margaret and Roger Lass, eds. *A Linguistic Atlas of Early Middle English, 1150–1325*. Available online at http://www.lel.ed.ac.uk/ihd/laeme1/laeme1.html. Edinburgh: University of Edinburgh, 2007. Accessed on 25 March 2013.

The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2

by M. D. Wilson. *Behavioural Research Methods, Instruments and Computers* 20. 1 (1988): 6 –11. (Database available online at http://www.psych.rl.ac.uk/ as of 25 March 2013.)

Secondary Sources

Hotta, Ryuichi. "Noun–Verb Stress Alternation: An Example of Continuing Lexical Diffusion in Present–Day English." *Journal of the Faculty of Letters: Language, Literature and Culture* 110 (2012): 36–63.

Hotta, Ryuichi. "Noun–Verb Stress Alternation: Its Nineteenth–Century Development and Its Earlier Historical Backgrounds." *Lexicon* 42 (2012): 79–94.

Hotta, Ryuichi. "Representativeness, Word Frequency, and Keywords in the *LAEME* Corpus." *Journal of the Faculty of Letters: Language, Literature and Culture* 112 (2013): 67–84.

Hotta, Ryuichi. "The Diatonic Stress Shift in Modern English." *Studies in Modern English* 29 (2013): 1–20.

Kelly, Michael H. "Rhythmic Alternation." *Cognition* 30 (1988): 107–37.

Kelly, Michael H. and J. Kathryn Bock. "Stress in Time." *Journal of Experimental Psychology: Human Perception and Performance* 14 (1988): 389–403.

Ogura, Mieko. "The Development of Periphrastic *Do* in English: A Case of Lexical Diffusion in Syntax." *Diachronica* 10 (1993): 51–85.

Ogura, Mieko and William S–Y. Wang. "Snowball Effect in Lexical Diffusion: The Development of *–s* in the Third Person Singular Present Indicative in English." *English Historical Linguistics 1994. Papers from the 8 th International Conference on English Historical Linguistics*. Ed. Derek Britton. Amsterdam: John Benjamins, 1994. 119–41.

Phillips, Betty S. "Word Frequency and the Actuation of Sound Change." *Language* 60 (1984): 320–42.

Phillips, Betty S. "Word Frequency and Lexical Diffusion in English Stress Shifts." *Germanic Linguistics*. Ed. Richard Hogg and Linda van Bergen. Amsterdam: John Benjamins, 1998. 223–32.

Phillips, Betty S. *Word Frequency and Lexical Diffusion*. Basingstoke: Palgrave Macmillan, 2006.

Rogers, Everett M. *Diffusion of Innovations*. 5 th ed. New York: Free Press, 1995.

Sereno, Joan A. "Stress Pattern Differentiation of Form Class in English." *The*

Journal of the Acoustical Society of America 79 (1986): S36.

Sereno, Joan A. and Allard Jongman. "Acoustic Correlates of Grammatical Class." *Language and Speech* 38 (1995): 57–76.

Sherman, D. "Noun-Verb Stress Alternation: An Example of the Lexical Diffusion of Sound Change in English." *Linguistics* 159 (1975): 43–71.

Sonderegger, M. and Partha Niyogi. "Variation and Change in English Noun/Verb Pair Stress: Data, Dynamical Systems Models, and Their Interaction." *Origins of Sound Patterns: Approaches to Phonologization*. Ed. A. C. L. Yu. Oxford: Oxford University Press, 2013. 262–84. (Preprint available online at http://people.linguistics.mcgill.ca/~morgan/ as of 7 March 2013.)

Wang, William S-Y. "Competing Changes as a Cause of Residue." *Language* 45 (1969): 9–25. (Rpt. in *Readings in Historical Phonology: Chapters in the Theory of Sound Change*. Ed. Philip Baldi and Ronald N. Werth. Pennsylvania: Pennsylvania State University Press, 1977. 236–57.)