Lexical Diffusion and Complex Adaptive System*

Mieko Ogura

Complexity is an inherently interdisciplinary concept that has penetrated a range of fields from physics to linguistics. Complex systems are made up of a large number of entities that by interacting locally with each other give rise to global properties that cannot be predicted or deduced from an even complete knowledge of the entities and of the rules governing their interactions. In many cases they are adaptive systems. Interesting principles have been proposed in an attempt to provide such an underlying, unified theory. These include selection, self-organization, scaling of the parameters, robustness, and networks of connections. In this study we show that lexical diffusion is the fundamental mechanism of language change, and it exhibits the underlying principles of complex adaptive system. First we discuss S-curve progress of language change, which is something like phase transition, and its snowball effect and relation to word frequency. Then we deal with linguistic selection, i.e., non-intentional functionally biased change and self-organization by language games, i.e., intentional, cooperative interactions of individuals in dynamic dialectology.

1. Introduction

The most explicit hypothesis on how sound change comes about was proposed by the Neogrammarians working in Germany in the mid-1870s. The Neogrammarian hypothesis has essentially two parts: lexical regularity and phonetic gradualness. We will not repeat the various arguments that have been offered to show that certain difficulties arise when we try to apply gradualism to all types of sound change (for details, see Ogura 1985, 1987). In view of the
unsatisfactory state of the mechanism that the Neogrammarians advocated, other processes for implementing sound change have been proposed. Empirical investigations over the past three decades on a variety of languages, using large amounts of data, have shown that there must be a process which is implemented in a manner that is lexically gradual, diffusing across the lexicon. This is an inevitable consequence of admitting phonetic abruptness; in his seminal article Wang (1969) called this process "lexical diffusion".

The chronological profile of lexical diffusion may be represented by the S-curve slope. When the change first enters the language, the number of words it affects may be small. The change gradually diffuses, going slowly at first. Then, as it spreads, it accelerates, picking up speed in mid-stream. In the most active period, the change moves quickly through a large number of lexical items. Then gradually, it slows down again, and tapers off at the end. The diffusion process is comparable to epidemics of infectious diseases, and the standard model of an epidemic produces an S-curve.

In lexical diffusion, diffusion from word to word in the language progresses from speaker to speaker in the community. The change catches on gradually, both within a language and when moving from speaker to speaker. Lexical diffusion model is defined as a 2-dimensional diffusion: diffusion from word to word in a single speaker, which we call W(ord)-diffusion, and diffusion from speaker to speaker of a single word, which we call Speaker-diffusion. When W-diffusion is slower than S-diffusion, the difference is greater between words. When W-diffusion is faster than S-diffusion, the difference is greater between speakers. W-diffusion may proceed so fast that it is difficult to observe it. This shows what is called the Neogrammrian regularity. Our 2-dimensional diffusion model operates not only in sound change but also in morphological, syntactic, semantic and lexical change. Figure 1 schematically shows the S-curve progress of 2-dimensional diffusion through time (t) when
Lexical Diffusion and Complex Adaptive System

W-diffusion is faster than S-diffusion ($W > S$), W-diffusion is slower than S-diffusion ($W < S$), and the rate of W-diffusion and S-diffusion is equal ($W = S$).

![Figure 1](image)

Figure 1 S-curve progress of 2-dimensional diffusion through time (from Ogura and Wang 1998)

Complexity is an inherently interdisciplinary concept that has penetrated a range of fields from physics to linguistics. Complex systems are made up of a large number of entities that by interacting locally with each other give rise to global properties that cannot be predicted or deduced from an even complete knowledge of the entities and of the rules governing their interactions. In many cases they are adaptive systems (Waldrop 1992, Vicsek 2002, Cangelosi & Parisi 2002).

Interesting principles have been proposed in an attempt to provide such an underlying, unified theory. These include selection,
self-organization, scaling (for example, power-law distribution, phase transition) of the parameters, robustness, and networks of connections. In this study we will show that lexical diffusion is the fundamental mechanism of language change, and it exhibits the underlying principles of complex adaptive system. We will discuss S-curve progress of language change, which is something like phase transition, and its snowball effect and relation to word frequency in section 2. In section 3 we will deal with dynamic dialectology and complex adaptive system. We will show that in linguistic selection, non-intentional functionally biased results of local actions of individuals tend to conspire to produce non-local universal patterns, while in self-organization by language games intentional, cooperative interactions of individuals tend to produce locally spread changes.

2. S-curve Progress, Snowball Effect and Word Frequency in Lexical Diffusion

The phase transition is familiar from such common sights as the bubbling water in a boiling kettle and the separate water molecules freezing into a block of ice. In this section we discuss S-curve progress of language change, which is something like phase transition, and its snowball effect and relation to word frequency, based on the development of periphrastic do and the development of -s in the third person singular present indicative in English. Furthermore we consider the correlation between word frequency and the direction of lexical diffusion.

2.1. The development of periphrastic do

Ogura (1993) examines the validity of the simultaneous equal activation scenario for the S-curve progress that Kroch (1989) proposes, and claims that changes in the different contexts initiate at different times and the later a change begins, the greater the rate of change becomes. Our data as well as Kroch's are based on Ellegård's
extensive and monumental study of 1953, which collects the periphrastic *do*-forms and non-periphrastic forms (called "simple forms") from prose texts from 1390 through 1710, showing their relative frequencies in several contexts and grouping the texts into 13 periods.\(^1\) The results are summarized in Table 1, and displayed graphically in Figure 2. For negative declaratives, negative questions, affirmative questions and negative imperatives, the numbers of tokens of the *do*-forms (abbreviated as do) and simple forms (abbreviated as s) are given, and for affirmative declaratives, the number of tokens of the *do*-forms and the total number of tokens of *do*-forms plus simple forms (abbreviated as n) are given.

We fit the data in Table 1 from period 0 through period 12 for affirmative declaratives, and from period 1 through period 12 for negative declaratives, negative questions and affirmative questions, and from period 2 through period 12 for negative imperatives to the

Table 1  The development of periphrastic *do* in various types of sentences (from Ellegård 1953)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>do  n</td>
<td>do  s</td>
<td>do  s</td>
<td>do  s</td>
<td>do  s</td>
</tr>
<tr>
<td>0</td>
<td>1390 - 1400</td>
<td>6 45000</td>
<td>0  —</td>
<td>0  —</td>
<td>0  —</td>
<td>0  —</td>
</tr>
<tr>
<td>1</td>
<td>1400 - 1425</td>
<td>11 4600</td>
<td>0  177</td>
<td>2  15</td>
<td>0  10</td>
<td>0  52</td>
</tr>
<tr>
<td>2</td>
<td>1425 - 1475</td>
<td>121 45500</td>
<td>11 892</td>
<td>2 23</td>
<td>6 136</td>
<td>3 279</td>
</tr>
<tr>
<td>3</td>
<td>1475 - 1500</td>
<td>1059 59600</td>
<td>33 660</td>
<td>3 24</td>
<td>10 132</td>
<td>0 129</td>
</tr>
<tr>
<td>4</td>
<td>1500 - 1525</td>
<td>396 28600</td>
<td>47 558</td>
<td>46 32</td>
<td>41 140</td>
<td>2 164</td>
</tr>
<tr>
<td>5</td>
<td>1525 - 1535</td>
<td>494 18800</td>
<td>89 562</td>
<td>34 22</td>
<td>33 69</td>
<td>0 101</td>
</tr>
<tr>
<td>6</td>
<td>1535 - 1550</td>
<td>1564 19200</td>
<td>205 530</td>
<td>63 21</td>
<td>93 114</td>
<td>0 72</td>
</tr>
<tr>
<td>7</td>
<td>1550 - 1575</td>
<td>1360 14600</td>
<td>119 194</td>
<td>41 7</td>
<td>72 56</td>
<td>4 39</td>
</tr>
<tr>
<td>8</td>
<td>1575 - 1600</td>
<td>1142 18000</td>
<td>150 479</td>
<td>83 45</td>
<td>228 150</td>
<td>8 117</td>
</tr>
<tr>
<td>9</td>
<td>1600 - 1625</td>
<td>240 7900</td>
<td>102 176</td>
<td>89 6</td>
<td>406 181</td>
<td>65 119</td>
</tr>
<tr>
<td>10</td>
<td>1625 - 1650</td>
<td>212 7200</td>
<td>109 235</td>
<td>32 6</td>
<td>116 24</td>
<td>5 16</td>
</tr>
<tr>
<td>11</td>
<td>1650 - 1700</td>
<td>140 7900</td>
<td>126 148</td>
<td>48 4</td>
<td>164 43</td>
<td>17 16</td>
</tr>
<tr>
<td>12</td>
<td>1710</td>
<td>5 2800</td>
<td>61 9</td>
<td>16 0</td>
<td>53 3</td>
<td>28 0</td>
</tr>
</tbody>
</table>

Aff. decl. = affirmative declarative sentence  
Neg. decl. = negative declarative sentences, main group  
Neg. q. = negative direct adverbial and yes/no questions  
Aff. q. = affirmative direct adverbial and yes/no questions  
Neg. imp. = negative imperatives, main group
Affirmative declaratives: lower solid line
Neg. declaratives, main group: lower broken line
Aff. questions, adv. and yes/no questions: upper solid line
Neg. questions, adv. and yes/no questions: upper broken line
Negative imperatives, main group: dotted line

Figure 2 The development of periphrastic do (from Ellegård 1953)
logistic function. Table 2 shows the slope and intercept parameters of the fits calculated by logit modeling in the Statistical Analysis System (SAS). When we compute the estimates for the parameters, we interpret the periods in Table 1 as the years from the reference point in time (t=0), in our case, the year 1175, when the first written example of all the sentence types appeared (see below).

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Slope and intercept parameters of logistic regressions on the data in Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Affirmative declaratives</td>
</tr>
<tr>
<td>slope</td>
<td>3.41</td>
</tr>
<tr>
<td>intercept</td>
<td>-23.61</td>
</tr>
</tbody>
</table>

Table 3 shows the estimates of the slope and intercept parameters obtained by Kroch (1989). He fixes the zero point in time at 1350, and uses a univariate version of the maximum likelihood fit in the variable rule (VARBRUL) program. He considers that there is a grammatical reanalysis in period 7, and cuts off the data after that (for criticisms of Kroch’s view, see Ogura 1993).

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Slope and intercept parameters of logistic regressions obtained by Kroch (1989)</th>
</tr>
</thead>
<tbody>
<tr>
<td>slope</td>
<td>2.82</td>
</tr>
<tr>
<td>intercept</td>
<td>-8.32</td>
</tr>
</tbody>
</table>

We take all the periods for which Ellegård provides data, and fit the data to the logistic curve. As shown in Table 2, it turns out that the contexts do have different slopes, i.e., different rates of change. When we look at the starting point of the changes, we find a clear correlation between the starting points and the rates of change. According to Visser (1963-73, §§ 1411-1476), the earliest dates of *do*
periphrasis in writing were: affirmative declaratives, c. 1175; negative declaratives, c. 1280; negative questions, c. 1370; affirmative questions, c. 1380; negative imperatives, c. 1422.

Our results in Table 2, together with the earliest dates of *do* periphrasis of five sentence types show that the later a change starts the sharper its slope becomes, i. e., the later a change starts the greater the rate of change becomes. This shows “snowball effect”, i. e., diffusion across more and more contexts at faster rates in later starting contexts.

The dates above cited from Visser are sporadic and they are the earliest dates at which the form is attested in the surviving written sample, which is not the same thing as the earliest dates of the actual change. Moreover, a linguistic phenomenon probably becomes noticeable in written texts long after it had become fairly wide-spread in spoken language. Even so, we may assume that these citations show the early appearances of periphrastic *do* which started undergoing the change at different times, based on the following statistical argumentation.

Suppose we assume with Kroch that the change began simultaneously in all environments, and proceeded with uniform slope. When the rate of occurrence of the change is very low in the earliest periods, it is only likely to show up in the surviving sample in those environments with lots of data. Thus, e. g., if the probability of the innovation was only .004, and we had one environment which occurred 10 times in our corpus, another with 100 occurrences, and another with 1000, we would be expecting .04, .4, and 4 examples of the innovation in respective random samples. Thus, at this period, it would be statistically quite reasonable to get a surviving example only in the third (highest frequency) environment. At some later period, when the probability of the innovation rises above .005, we would expect a case to occur in the middle frequency environment, and at still later times in the low frequency environment.
If the number of occurrences of tokens in each environment in the sample does not reflect the actual frequency, then the premise of the logical conclusion should be the simultaneous start of the changes. However, if the number of occurrences of tokens in each environment reflects the actual frequency, the situation is different. If we cannot get an example in the middle and low frequency environments at the earliest periods, it means that the actual change has not appeared in those environments. At some later period, when the probability of the innovation rises, an example occurs in the middle frequency environment, which means that the actual change has first appeared in the middle frequency environment. We may assume that the premise of the logical conclusion should be the sequential start of the changes.

The total number of occurrences of tokens in each environment in Table 1 reflects the actual frequency of tokens. Given the actual frequency of tokens in each environment reflected in Table 1 and the uniform rate of change, we can predict that the change in each environment started sequentially. Thus there is little probability for Kroch's assumption that the change began simultaneously in all environments and proceeded with uniform rate; note that the earliest dates in the Visser data are not the outcome of a common starting date.

We may gather from Table 1 that the changes appeared earlier in the contexts where the larger number of tokens occur. Affirmative declaratives, in which change initiated first of all contexts have the largest number of tokens. Then the change started in negative declaratives which have the second largest number of tokens, then in questions both negative and affirmative, and lastly in negative imperatives which have the smallest number of tokens.

Within each context, there is a significant tendency for the high frequency words to change late and therefore to have a sharp slope. Table 4 shows the development of the do-form in the say-group which
Mieko Ogura

consists of high frequency verbs *say*, *mean*, *do* and *think* and the main group which consists of the rest of the low frequency words in affirmative *wh*-object questions, based on the data presented in the form of page and line references and lexical information at the end of Ellegård's book. Blank means that the example has not appeared yet. As mentioned above, the early example in the main group appeared in c. 1380, though the data on the main group in Table 4 do not show the occurrence of *do* periphrasis in periods 1 and 2. The first occurrence of *do* periphrasis in the *say*-group in Table 4 is found in period 3. Thus the lag of the *say*-group is about 100 years.

Table 4 The development of the *do*-form in the *say*-group and the main group of affirmative *wh*-object questions

<table>
<thead>
<tr>
<th>Period</th>
<th>Date</th>
<th><em>say</em>-group</th>
<th>main group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>do</em></td>
<td><em>s</em></td>
</tr>
<tr>
<td>1</td>
<td>1400 - 1425</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1425 - 1475</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1475 - 1500</td>
<td>1</td>
<td>39</td>
</tr>
<tr>
<td>4</td>
<td>1500 - 1525</td>
<td>2</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>1525 - 1535</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>6</td>
<td>1535 - 1550</td>
<td>0</td>
<td>45</td>
</tr>
<tr>
<td>7</td>
<td>1550 - 1575</td>
<td>3</td>
<td>51</td>
</tr>
<tr>
<td>8</td>
<td>1575 - 1600</td>
<td>7</td>
<td>56</td>
</tr>
<tr>
<td>9</td>
<td>1600 - 1625</td>
<td>25</td>
<td>93</td>
</tr>
<tr>
<td>10</td>
<td>1625 - 1650</td>
<td>15</td>
<td>39</td>
</tr>
<tr>
<td>11</td>
<td>1650 - 1700</td>
<td>24</td>
<td>20</td>
</tr>
<tr>
<td>12</td>
<td>1710</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 5 shows slopes and intercepts for the *say*-group and the main group. We can say that the high-frequency words resisted the *do*-form, but once they started to change, the rate of change turned out to be greater than that of the low frequency words. We again find snowball effect within each context.

The above results can be confirmed in negative declaratives. Ellegård classifies the negative declaratives into 2 groups, the *know*-group which seems to have resisted the *do*-form such as *know*,
Lexical Diffusion and Complex Adaptive System

Table 5  Slope and intercept parameters of logistic regressions on the data in Table 4

<table>
<thead>
<tr>
<th></th>
<th>say-group</th>
<th>main group</th>
</tr>
</thead>
<tbody>
<tr>
<td>slope</td>
<td>10.49</td>
<td>6.82</td>
</tr>
<tr>
<td>intercept</td>
<td>−65.19</td>
<td>−41.33</td>
</tr>
</tbody>
</table>

do, doubt, care, list, fear, skill, trow and boot, and the main group which consists of the rest of the verbs. Ellegård's data in Tables 1 and 6 show that 22% of the total number of tokens for negative declaratives consists of the know-group verbs. He gives lexical information on the 1512 tokens of the total number of the 1620 know-verbs, and we get the frequency of the 9 know-verbs. The results are as follows: know, 869 tokens; do, 212 tokens; doubt, 229 tokens; care, 127 tokens; list, 15 tokens; fear, 47 tokens; skill, 5 tokens; trow, 3 tokens; boot, 5 tokens. Although Ellegård does not give lexical information on the main group verbs, we may assume that know, do, doubt, care, list and fear are the frequent words in negative declaratives. Three words in the know-group, i.e., skill, trow and boot are not frequent words. Though they are included in the know-group, they have no effect on the parameter estimates.

Table 6 shows the development of the do-form in the know-group. Blank means that the example has not appeared yet. As mentioned above, the early example in the main group appeared in c. 1280, though the data on the main group in Table 1 do not show the occurrence of do periphrasis in periods 0 and 1. The first occurrence of the know-group in Table 6 is found in period 3. Thus the lag of the know-group is about 200 years. We find that there is a significant tendency for the high frequency words to start to change late.

Table 7 shows slopes and intercepts for the know-group and the main group. The know-group, which started later, has a sharper slope, i.e., a greater rate of change.
2.2. The development of -s in the third person singular present indicative

The snowball effect and the interaction between word frequency and environments in lexical diffusion can also be found in the development of -s in the third person singular present indicative. Based on the data from the EModE section of the Helsinki Corpus, Ogura and Wang (1996) give the overall distributions of the -th and -s forms by sub-periods for the non-sibilant verbs which are divided into three groups according to word frequency as shown in Table 8. The percentages of the -s forms for the total tokens for each sub-period are given for each of the three groups of the non-sibilant verbs. For example, there are 33 types of non-sibilant verbs whose word frequency is from 1084 to 21. The -th and -s forms occur in 1103
tokens and 29 tokens respectively in EModE I for the 33 types. Thus
-s forms occur in 2.6% of the total tokens which are the sum of the
occurrences of the -th and -s forms (1103+29) in EModE I.

Table 8 The overall distributions of the -th and -s endings in non-sibilant
verbs

<table>
<thead>
<tr>
<th>freq</th>
<th>I-th</th>
<th>I-s</th>
<th>II-th</th>
<th>II-s</th>
<th>III-th</th>
<th>III-s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1084–21 (33 types)</td>
<td>1103 tokens</td>
<td>29 tokens (2.6%)</td>
<td>932 tokens</td>
<td>331 tokens (26.2%)</td>
<td>251 tokens</td>
<td>697 tokens (73.5%)</td>
</tr>
<tr>
<td>20–3 (176 types)</td>
<td>384 tokens (1.5%)</td>
<td>6 tokens</td>
<td>282 tokens</td>
<td>166 tokens (37.1%)</td>
<td>28 tokens</td>
<td>339 tokens (92.4%)</td>
</tr>
<tr>
<td>2–1 (262 types)</td>
<td>116 tokens (0%)</td>
<td>0 tokens</td>
<td>72 tokens (25.8%)</td>
<td>25 tokens</td>
<td>5 tokens</td>
<td>121 tokens (96.0%)</td>
</tr>
</tbody>
</table>

Within the non-sibilant verbs, most of the -s forms in EModE I
(1500–1570) occur in the most frequent 33 words, whose word
frequency is from 1084 to 21. Only 3 words have -s form in EModE I
among the 438 infrequent verbs. The change started slowly from a
handful of high frequency words. Holmqvist (1922) considers that
have, do and say are the laggars of the change, which has been an
established view so far. But our data show that have, do and say are by
far the most frequent words, and the most frequent verbs started to
change first.

But once the infrequent verbs got started, they changed more
quickly than the frequent verbs. Most of the less frequent 176 verbs
whose word frequency is from 20 to 3 show the -th forms in EModE I.
Many of them started to change and then quickly changed in EModE
II (1570–1640), and completed the change in EModE III (1640–1710).
The least frequent 262 verbs whose word frequency is 2 or 1 never
show the -s forms in EModE I and rarely in EModE II. Many of them
quickly changed into the -s forms and completed the change in
EModE III. The -th forms are rare in EModE III in the 438 infrequent verbs. On the other hand, the most frequent 33 verbs often show the -th forms still in EModE III. Especially the verbs do and have show the -th forms nearly 50% and more than 50% of the total tokens respectively in EModE III.

Figure 3 is an idealized diagram of the snowball effect in lexical diffusion. The abscissa shows the time and the ordinate shows the percentage of changed variants. Each S-curve represents the rate of change of each word, and the time interval between each word represents the rate of change through the lexicon. The later the change the more words change and later words change faster.

![Figure 3](image)

**Figure 3** An idealized diagram of snowball effect in lexical diffusion (from Ogura and Wang 1996)

### 2.3. Word frequency and the direction of lexical diffusion

Our next issue is what factors determine the directions for lexical diffusion with regard to word frequency. Hooper (1976) shows that reductive phonetic change like schwa-deletion in sequences -ary, -ery, -ory and -ury tends to affect high frequency words first, whereas analogical change like the weakening of strong verbs tends to affect
low frequency words first. Furthermore, she suggests that reductive sound change affecting high frequency words first have their source in production, whereas analogical leveling of the weakening of strong verbs is due to imperfect learning. Bybee (2002) further discusses the problem, based on the exemplars that are available in experience.

Phillips (1984) distinguishes phonological changes into physiologically motivated changes where the frequent words change first, and non-physiologically motivated changes where the infrequent words change first. She gives several instances of non-physiologically motivated changes which affect infrequent words first: the glide deletion after alveolar stops as in tune and duke, the unrounding of ME /ɔː(ː)/, and diatone formation.

We may add the development of periphrastic do and the Great Vowel Shift (GVS) to the non-physiologically motivated changes. Periphrastic do evolved as an adaptation to the cognitive constraint operating on the underlying structure (see note 1), and the change started in the infrequent words. The GVS was motivated by maximum perceptual contrast among the vowels in the system (see section 3.1). We assume that the initial impetus was a push-chain effect. Welna (2004) gives the earliest forms with /eː/-raising based on the OED and the MED (for the evidence of the push-chain effect from the present-day Northern England dialects, see Ogura 1987, 3.4). The words with i/y-spellings indicating /eː/-raising are found in infrequent words first especially before /r/ and a few other contexts. In the development of ME /iː/, the change moves quickly through words and there is no significant word frequency effect as stated below. Based on Ogura’s data for ME /iː/ (1995), Phillips (2006, 3.1.4) states that the change started from the frequent words, but her interpretation of the data is incorrect.

Phillips (2001, 2006) attempts to offer a further refinement of her hypothesis, namely that sound changes which require analysis – whether syntactic, morphological, or phonological – during their
implementation affect the least frequent words first; others affect the most frequent words first. However, the interaction between word frequency and environments operates not only in changes that affect the least frequent words first but also in those that affect the most frequent words first, as shown in the development of -s in the third person singular present indicative and the changes mentioned below. Thus her remarks are factually incorrect.

Ogura (1987, Ch. 5) examines the interplay of phonetic factors and word frequency in the snowball effect of shortening of Early Modern English /u :/, where the most frequent words change first. Ogura and Wang (1995) show that semantic change begins in the frequent words first among synonymous words. Semantic change is a pragmatic process, i.e., a context-dependent aspect. Thus it is in frequent words that semantic change is most likely to occur first.

We may synthesize these investigations and assume that:

a) Productively or physiologically motivated change, pragmatically motivated change, and socially motivated change occur in high frequency words first. Productively or physiologically motivated change, and pragmatically motivated change are the natural result of linguistic productions, and socially motivated change affects linguistic productions externally. Thus it follows that such changes are more advanced in productions that are more highly practiced, i.e., in high frequency words.

b) Perceptually motivated change and cognitively motivated change affect low frequency words first. Perceptually or cognitively unfavorable forms can be learned and maintained in their forms if they are of high frequency in the input. However, if their frequency of use is low, they may not be sufficiently available in experience to be acquired. Thus they may be subject to changes for perceptually or cognitively favorable forms.

Speakers always observe frequent words, and thus frequent words spread through interactions among people. When the change
starts from high frequency words, it takes a long time to complete the change because the unchanged variants of high frequency words are maintained, and thus frequent words also tend to become lagger, as shown in the development of -s in the third person singular present indicative. When the change starts from low frequency words as shown in the development of periphrastic do, speakers observe unchanged variants of high frequency words for a long time, and the high frequency words become lagger of the change.

Lastly we would like to explain the rapid mid-stream change of S-curve progress. Ogura (1995) shows, based on the development of ME /i:/ and ME /u:/words at 311 sites in England, that there is no significant ordering relation among words through which the change moves quickly in mid-stream, therefore no significant word frequency effect, and the order of the change of words varies among individuals. Gell-Mann (1992) was perhaps the first to suggest the relevance of Kolmogorov Complexity to the study of language evolution. Kolmogorov Complexity has led to methods for inductive inference, based on the search for the simplest interpretation of observed data, and has applied to representations of any kind: logical, linguistic, probabilistic or pictorial. When regularity exists in the observed data, the hypothesis will capture this regularity, when justified, and allow for generalization beyond what was observed. Thus we assume that the speakers, after they observe a small number of changed words, generalize the change into more and more words without observing all the relevant words, and thus the order of the generalization varies among individuals.

3. Linguistic Selection and Language Games in Dynamic Dialectology

Gell-Mann (1992) points out that the selective effect is the central feature of complex adaptive systems, and selection provides a bottleneck that induces adaptation. In the Origins of Species (1859),
Darwin noted that natural selection cannot directly promote altruistic acts where individuals reduce their own competitive ability but increase that of others. Yet cooperation is abundant in nature. Evolutionary game theory (Maynard-Smith 1982) has allowed biologists to analyze such dynamics. The minimal assumption is that the transmission of information between partners provides them with an advantage, for example, by exchanging know-how or coordinating their behavior, and that the advantage translates into more offspring, with similar communicative skills. In this section we discuss linguistic selection based on the GVS and the development of periphrastic do, and language games based on the development of -s in the third person singular present indicative and the development of West Germanic *a before nasals under Mercian domination.

### 3.1. Linguistic selection

Linguistic selection is unconscious functional selection between available variants by learners. Languages become adapted to the productive, perceptual and cognitive abilities of human beings in the transmission across generations. Languages tend towards uniformity, because every language will discover the same optimal functionally selected compromise.

The driving force of the GVS is maximum perceptual contrast among the vowels in the system. We assume that the developments of ME long vowels implemented themselves gradually across the lexicon for several centuries, maximizing dispersion and producing shifts in the direction that was originally determined by the triggering event.

Based on *A Linguistic Atlas of Early Middle English*, covering the period 1100–1350, Stenbrenden (2003) finds that of 139 tagged texts, twenty-three show ⟨ei/ey⟩ for ME /i:/, and rather more show other irregular spellings for ME /i:/.

The language of the MSS which have been localized belongs to the West Midlands and East Midlands mostly, though the North Midlands and the North are also rep-
Lexical Diffusion and Complex Adaptive System

resented. That is, there is some scattering of such forms throughout the country. With regard to late ME, *A Linguistic Atlas of Late Mediaeval English 1350–1450* (LALME) shows a wide scattering of ⟨ei/ey⟩ all over the country, though they are particularly frequent in the South-East and South-West Midlands and Yorkshire.

One might argue that the vowel shift did not proceed in the same way in the dialects as it is assumed to have done in Standard English. But this is not factually correct. Based on the *Survey of English Dialects* (1962–1971), Ogura (1990) and Ogura, Wang and Cavalli-Sforza (1991) quantitatively examine the present-day reflexes of ME long vowels as these are distributed in 311 sites in England. When we plot out the most frequent reflexes of 7 long vowels on a map of England, there emerge various clines emanating from several known population centers throughout England.

An analysis of the materials reveals that ME ⟨i :⟩ has 17 reflexes in England. On the basis of a variety of philological and phonetic considerations, we interpret the current reflexes to be related to each other diachronically as shown in Figure 4.

![Figure 4](image)

Figure 4 Relations among the reflexes of ME/i :/ (from Ogura 1990)
We next tabulate the frequencies of the 17 vowel reflexes for 39 ME /iː/ words. Map 1 shows the distribution of the first mode reflex, the most frequent vowel at a site, of 39 words. We can discern several regions where the change is more advanced. One such region is the coast of Essex. There the first mode reflex has reached 17(= /ɔi/). This focal area is surrounded by sites where the first mode reflex is 15 (= /ɔi/), which in turn borders with sites where the first mode reflex is 14(= /ʌi/) or 11(= /ai/). A similar situation can be seen around Oxford area. Here again, the first mode reflex has reached 17(= /ɔi/) in a focal area that is surrounded by sites where earlier reflexes occur. Another focal area revealed by the first mode reflexes is around
Manchester and Sheffield; yet another around Birmingham. We may refer to such regions as exhibiting a "gradient", which allows us to make inferences concerning historical developments. Typical of such gradients is the observation that the inner areas are more advanced in the change.

de Boer (2000) shows the self-organization of a vowel system through language games of imitation in simulation. A group of agents converge to a coherent vowel system by iterative interactions with each other. Livingstone (2002) extends de Boer's simulation, including a larger population of one hundred agents which spread across a spatial array. The results of the simulation show that the different vowel systems are used by the different subgroups of the population. Thus we maintain that the GVS was not implemented by the language games of imitation. If the vowel shift had been implemented by the language game, the vowel shift would not have proceeded in the same line as it is assumed to have done in Standard English. (For details, see Ogura and Wang 2004.)

Periphrastic do evolved as an adaptation to the cognitive constraints operating on the underlying structure. Ellegård (1953, pp. 56–75) gives early instances of affirmative declaratives from the 13th century verse of the Western texts, the 14th century verse of the Eastern texts, the 13th century verse of the South-Eastern texts, the 14th century verse of the Northern and North-Western texts, and the 14th century prose of the Central texts. The change probably started in the South-Western dialect and spread in most dialects in the 14th century and the simple forms coexisted with the new forms.

3.2. Language games

The changes arising from social factors spread by self-organization of language games, i.e., the cooperation in the interactions of the individuals. The size of the neighborhood determines the number of individuals that interact, and socially influential people have an
increased probability of being imitated by their neighbors. Hence, successful changes spread locally. Without socially influential people, a considerable number of innovators are necessary for the change to spread.

The third-person singular ending -s first emerged in the 10th century in Northumbrian texts. The change had not yet spread even to the northernmost part of the East Midlands at the beginning of the 13th century, but in the course of the 13th and 14th centuries, the -s form gradually displaced the old ending over the whole of Lincolnshire (Holmqvist 1922, Ch. III). In the century from 1350 to 1450, -th was the regular ending in Norfolk, but we can also find -s there (Holmqvist 1922, Chs. III & VI, McIntosh, Samuels & Benskin 1986, Vol. III). In the 15th century prose written in London English, the -s ending is very seldom found (Holmqvist 1922, Ch. VII), but the early instances of the -s form in prose appeared in the first half of the 16th century. We assume that the change must have propagated through interaction between ordinary people with the -th form and those with the -s form, and a considerable number of the latter would have been necessary for the propagation of the -s form. The change started in the North, and we may conclude that the -s form is an adoption from Old Norse (ON) in face-to-face interactions of ordinary English people with their Danish counterpart. In many places the Scandinavian population exceeded 50% (Bailey & Maroldt 1977).

Speakers model themselves on those with whom they wish to associate and identify, and with whom their aspirations are bound up (Labov 2001). Socially influential people are from time to time the bearers of new variants. Even rare variants get adopted and spread through entire communities in an S-shaped curve and convergence is rapid.

One example is the development of West Germanic *a before nasals under Mercian domination. Toon (1983) shows the progress of the sound change from a to o under Mercian domination. The change
Lexical Diffusion and Complex Adaptive System

probably began in the North and diffused through the Mercian speech community. The early 8th century *Epinal Glossary*, whose language may reflect an even earlier period (c. 700), has *a* consistently and is unaffected by the change. The other glossaries contain data mirroring the development of *a* to *o*. The data in the *Corpus Glossary* argues for the change by the turn of the 8th century. The *o* spelling predominates in Mercian charters from 736 on, and is found exclusively from 812 to 845.

Early West Saxon exhibits fluctuation between *a* and *o*. Of more global interest is the fact that Mercian political domination could effect linguistic change in Kent. The data from the Kentish charters before the period of Mercian control unanimously use *a* forms. But there is a dramatic change in the charters produced under Mercian domination. The data almost exclusively have *o* spellings between 803–824. During the first years of 833–850, after the beginning of decay of Mercian influence, there is fluctuation between *o* and *a* spellings. In the data for the second generation of 859–868, after the beginning of the Mercian exodus, *a* spellings predominate.

Lowe (2001) criticizes Toon’s dating of charters after 833, and Toon’s statistics should be revised. But we still find that charters dating from 833–850 and 859–868 show fluctuation between *o* and *a*: 833–850 in favor of *o* and 859–868 in favor of *a*. Thus Toon’s major point that linguistic change in Kent is concomitant with Mercian political domination holds true.

We may add that the change affects the frequent words first. Phillips (1984) considers that the change is motivated phonetically: the position of the tongue rises in anticipation of a following nasal. We assume that the change is motivated socially. If it were motivated phonetically, it would not have reversed with the decline of Mercian domination. (For more examples of linguistic selection and language games, see Ogura and Wang 2004, and the diffusion through social networks in linguistic selection and language games, see Ogura and
4. Conclusion

We have shown that lexical diffusion is the fundamental mechanism of language change and it exhibits the underlying principles of complex adaptive system, especially, phase transition, selection and self-organization.

Section 2 deals with S-curve progress of lexical diffusion, which is something like phase transition. Our focus has been on W-diffusion. We have discussed the snowball effect, i.e., the later the change, the more words change and the later words change faster. We have explained why productively, pragmatically, or socially motivated changes start in frequent words, while perceptually, or cognitively motivated changes start in infrequent words first, based on the exemplars available in experience in the language transmission across generations. We have also explained why there is a rapid mid-stream change from Kolmogorov Complexity.

Section 3 deals with linguistic selection and self-organization of language games. Our focus has been on S-diffusion. We have shown that in linguistic selection, non-intentional functionally biased results of local actions of individuals tend to conspire to produce non-local universal pattern, while in self-organization by language games intentional, cooperative interactions of individuals tend to produce locally spread changes in dynamic dialectology.

Labov (2001, Ch. 13) states that the transmission of change across generations has not been discussed or even recognized in the literature of historical linguistics. We have shown that language transmission across generations both in W-diffusion and S-diffusion play a large part in the complex adaptive system of lexical diffusion, or more generally in language change.
NOTES

* This is based on a paper presented at the Symposium on Lexical Diffusion at the 23rd meeting of Modern English Association on May 19, 2006. I wish to thank William S-Y. Wang for his stimulation to my progression from idea to data to knowledge on language evolution, and the reviewers for helpful comments. This research is supported by the grants from the Human Frontier Science Program and the Ministry of Education, Culture, Sports, Science and Technology of Japan.

1. There is a great cleavage of views on the origin of periphrastic do. Most of the views have been summarized by Ellegård (1953, Part I) and Visser (§§ 1411–1417); for a critique of Ellegård and Visser, see Denison (1985). We find that Kroch’s (1989) is the most plausible explanation of why the periphrastic construction arose in the history of English. He shows, on the basis of accounts in Robers (1985), that the crucial factor in the development of the modal auxiliary was the collapse of the subjunctive mood in the course of ME, which led to the use of modals in place of the subjunctive inflection. When the modals lost their status as main verbs and became syntactic auxiliaries, appearing in the INFL position, main verbs were no longer raised to the INFL position, and the tense marking of main verbs occurred via a transfer of the affix from INFL to the verb in its deep structure position. This transfer was blocked by the sentence negator not and periphrastic do was inserted to provide lexical support for the affixes in INFL in negative declaratives.

The appearance of do in questions is also motivated by the need for a lexical support for the affixes. The questions have been signaled by the fronting of INFL to the COMP(limentizer) position. In ME this fronting occurred subsequent to the verb raising to INFL; and in ModE V(erb)-to-I (NFL) raising was gradually lost in main verbs, and the appearance of the subject between INFL and the main verb blocked transfer of the affix, forcing do insertion.

Recently Van der Auwera & Genee (2002), McWhorter (2006) and Filppula (2006) propose the Celtic hypothesis. Ellegård (pp. 119–120) criticizes the Celtic origin but he thinks that it might have been a contributing factor. According to Van der Auwera & Genee (2002), the periphrastic construction is very prominent in Celtic and it is attested from
Old Irish (pre-tenth century) onwards. Since periphrastic *do* probably started in the South-West where Celtic influence was strong (Ellegård, pp. 47, 118–119, 208), we may assume that Celtic influence might have been a contributing factor to the change.

2. Several views have been presented on the origin of -s (for details, see Ogura & Wang 2004). We assume that the -s form is an adoption from Old Norse in face-to-face interactions of ordinary English people with their Danish counterpart (see section 3.2).

REFERENCES


— 48 —
Lexical Diffusion and Complex Adaptive System


Mieko Ogura


Lexical Diffusion and Complex Adaptive System


(Tsurumi University)

[ogura-m@tsurumi-u.ac.jp]

[Received September 15, 2006]