# Tongue Twister Effect in L2 Silent Reading

Chizuru Mori

## 1. Introduction

The role played by inner speech during reading has received considerable attention, as it is a central issue in understanding written language. One of the claims that has often been made in L1 (English) studies is that inner speech aids higher-order comprehension such as syntactic parsing and semantic integration. The reason for this is that phonological codes involved in inner speech are more stable in working memory[1], and hence they are useful when a reader refers back to earlier information (cf. Rayner & Pollatsek, 1989). If so, it is possible to infer that inner speech is beneficial for L2 learners as well. In the present study, I will investigate the role of inner speech in the reading processes of advanced L2 learners. More specifically, I will explore the role of phonological recoding in L2 reading. Here, it is necessary to clarify what "phonological recoding" means. According to the definition of Rayner & Pollatsek (1989), *phonological recoding* refers to "the mental representations of speech that can give rise to the experience of hearing sounds, " whereas *subvocalisation* refers to "activity in the speech tract (either muscle movement or articulatory process)" (p. 189).

There is an argument that even when subvocalisation is

---

[1] Baddeley (1986) defined the term "working memory" as follows: "The term working memory implies a system for the temporary holding and manipulation of information during the performance of a range of cognitive tasks such as comprehension, learning and reasoning" (p.34). It is generally agreed that "working memory is a oncept which replaced short-term memory" (cf. Daneman 1987, p.59).

suppressed by a certain technique, still phonological recoding takes place. This implies that phonological recoding is a deeper level activation than subvocalisation. Perfetti (1985) suggested that in normal silent reading of native speakers of English, phoneme level activation occurs rather than speech muscle activation. It seems sensible, therefore, to do research on this deeper level activation (i. e. phonological recoding) in L2 reading rather than subvocalisation because phonological recoding is considered to be the essential constituent part of inner speech. The main concern in the present study is to explore the role of phonological recoding in advanced L2 reading and see if there is any similarity between L1 reading and L2 reading.

It should be also noted that L2 learners in the present study are native speakers of Japanese who have been learning English as a foreign language in school settings. If it is found that they benefit from phonological recoding as native speakers of English do, it raises an interesting question: how can logographic language users such as Japanese speakers come to be proficient at reading an alphabetic language such as English? This topic will be discussed as a related issue of the role of phonological recoding in L2 reading, because it will be of great value for pedagogical purpose.

## 2. Literature Review

According to Rayner & Pollatsek (1989), there are three principal techniques that have been used to measure phonological recoding in L1 reading: Homophonic Phrases, Tongue Twister and Phonemic Similarity. Among these, the first two techniques have been widely used. One of the influential studies using Homophonic Phrases is Doctor & Coltheart (1980). In this study, children (age 6-10) were given phrases of the following types:

a. Meaningful sentences containing no homophones
(e. g. He ran <u>in</u> the street)

b. Meaningful sentences containing homophones
(e. g. He ran through the street)
c. Meaningless all-word sentences that sound correct
(e. g. He ran threw the street)
d. Meaningless all-word sentences that sound wrong
(e. g. He ran sew the street)

They found that meaningless sentences that sound correct (e. g. "He ran threw the street") produced more incorrect responses than did meaningless sentences that sound wrong (e. g. "He ran sew the street"). It was concluded that very young readers rely extensively on phonological recoding when reading for meaning, and as they grow older, this reliance decreases.

The second technique is called "Tongue-Twister." The most comprehensive study involving tongue twisters is McCutchen & Perfetti (1982). They criticised the concurrent articulation paradigm (cf. Kleiman, 1975; Levy, 1977; Slowiaczek & Clifton, 1980), saying that "a model of speech processes in reading does not need to assume the operation of some general all-purpose 'speech mechanism' " (p. 673). Instead, they assumed that speech codes may be specific for speech segments and hypothesized that an abstract phonological representation containing information about the word-initial phoneme would be useful in reading, particularly for the integration processes of comprehension:

> ... word-initial phonetic information could provide a concise index by which to reaccess specific words, if that became necessary during comprehension.
>
> (McCutchen & Perfetti 1982, p. 673)

Their rationale was that if their hypothesis was correct, phonetically similar representation such as tongue twister sentences caused the kinds of similarity confusions often observed in memory tasks. They constructed several tongue twister sentences that repeated initial consonants across several words (e. g. The detective discovered the danger and decided to dig for details), and asked the subjects to judge the acceptability of these sentences. They found that the subjects took significantly more time on tongue twister sentences than on control sentences, which means that the subjects could not make use of the word-initial information for tongue twister sentences because of similarity confusions. This result led them to conclude that specific phonetic features get activated in the reading process and this phonological information might be useful postlexically in working memory.

As Rayner & Pollatsek (1989) summarised, the main role of phonological recoding at a sentence or discourse level is considered to be to aid higher-order comprehension processes in reading:

> Although the meaning of individual words can be determined without recoding written language into speech, phonological codes appear to be activated for most words we read and this phonological information is held in working memory and is used to comprehend text.
>
> (Rayner & Pollatsek 1989, p. 216)

Not many studies have been done on the phonological processes involved when people read a foreign or a second language. Segalowitz & Hébert (1990) explored phonological recoding in reading by highly skilled bilinguals (L1=French, L2=English). Their main research question was to see if differential phonological effects would be obtained in the processing of sentence comprehension in L2 as opposed to single word recognition. They conducted experiments in lexical decision and sentence verification, using sentences which contained homophones. In the lexical decision task, an English list and a French list were constructed. Each list consisted of 36 homophones (e. g. poll), 36 control words (e. g. moss), 36 pseudowords (e. g. grean) and 36 control nonwords (e. g. trean). The subjects' task was to judge whether the letter string presented was a real word or not. The results were as follows. For error rate, there were no significant homophone or pseudoword effects. For reaction time, there was a significant pseudoword effect, reflecting the slower responses to pseudowords than to control nonwords. They concluded that these bilinguals did not rely on sound so much when they read a single word.

In the sentence verification task, there were four types of sentences which contained homophones:

(1) Homophone sentences that contained a homophone
(e. g. She said the weather was <u>fair</u> outside.)
(2) Control sentences that did not contain a homophone
(e. g. She said the weather was <u>nice</u> outside.)
(3) Congruent sentences that sounded meaningful
(e. g. She said the weather was <u>fare</u> outside.)
(4) Incongruent sentences that sounded meaningless
(e. g. She said the weather was <u>hair</u> outside.)

They obtained the following results. For error rate, there was a significant homophone effect, in which more errors were made for sentences containing homophones (Type 1) than control sentences (Type 2). Also, there was a significant congruency effect, in which there were more errors with congruent sentences (Type 3) than with incongruent sentences (Type 4). On the other hand, no significant effects were found in reaction time. In sum, they found that the subjects did not produce any significant homophone effect in L2 in the lexical decision task, whereas a homophone effect was obtained in the sentence verification task. This finding led them to conclude that subjects have to depend on sound more in the sentence verification task than in the lexical decision task because they have to draw upon phonological codes postlexically in working memory to integrate the sentential information.

While Segalowitz & Hébert (1990) was concerned with second language reading of fluent bilinguals, Mori (1992) investigated whether or not L2 learners carry out phonological recoding when they read for meaning. L2 learners in this case meant Japanese speakers who have been learning English as a foreign language in school settings. There were two assumptions: (1) L2 learners would take the same course as L1 readers, in which they rely on phonological recoding first, and gradually come to make use of visual information, (2) L2 learners would not at all depend on phonological recoding because in most of the cases, they might be more familiar with visual representation of words rather than the sounds. Replicating Segalowitz & Hébert (1990), she used a sentence verification task containing homophones, though the adjustments were made for the number of test sentences and the length of each test sentence. As a result, she obtained a significant congruent effect (i. e. a significant difference between Type 3 and Type 4), but was not able to find any significant homophone effect (i. e. no difference between Type 1 and Type 2) in error rate. There were no significant effects in the analysis of reaction time. Based on the result of significantly greater number of errors in meaningless sentences that sound correct, she concluded that advanced L2 learners might go through the same developmental process as L1, pointing out that her findings (i. e. a significant congruent effect and no homophone effect) were perfectly consistent with those of Doctor & Coltheart (1980).

## 3. Research Questions and Experimental Methods

It has already been made clear in Mori (1992) that Japanese people depend on sound to a certain extent when they read English sentences. In the present study, I would like to expand Mori (1992) and investigate if phonological recoding is useful particularly for sentence comprehension (but is not necessarily so for word recognition) in L2 reading. As mentioned earlier, it is generally agreed that phonological codes are held in working memory, and are used to comprehend text. Perfetti (1985) and McCutchen, Bell, France & Perfetti (1991), for example, argued that phonological codes are more stable than visual codes and they are superior *reference securing* and *order securing* codes (cf. McCutchen et al. 1991, p. 89). Thus, phonological information stored in working memory is useful especially for higher-order processing such as syntactic parsing and semantic integration. I will explore this issue by investigating phonological recoding at a lexical level as well as a sentence level, employing Tongue-Twister Technique. If phonological recoding takes place only at a sentence level and not at a lexical level, it can be concluded that phonological recoding is useful particularly for higher-order processing in sentence comprehension.

The present study basically replicates the experimental methods employed by McCutchen et al. (1991). They proposed the reference-and order-securing hypothesis which elegantly integrate specific phonological codes with working memory:

> On the basis of the tongue-twister effect, we propose a model of reading in which sentence comprehension is the result of integrating phonetically indexed lexical representations in a temporary working buffer.... Together, these features provide a unique reference to the lexical items in working memory and aiding in retention of its place in the temporal order of the items, thus facilitating reaccess to specific lexical items when necessary.
>
> (McCutchen et al. 1991, p. 89)

They tested this model by using a lexical decision task and a sentence verification task which contained tongue twister sentences. While in McCutchen & Perfetti (1982), most of the test sentences repeated the same word-initial consonant (e. g. only /d/ or only /p/), McCutchen et al. (1991) used the words which began with either /t/ and /d/, or /s/ and /z/ to lessen the visual similarity of tongue twister sentences. In Experiment 1, there were three sets of syntactically parallel sentences for the sentence verification task. In one set of sentences, all the content words began with word-initial alveolar stops. Another set of sentences contained word-initial alveolar fricatives. A third set was control sentences which contained various word-initial phonemes. Each of these acceptable control sentences was matched semantically with one of the tongue twister sentences. Here are the examples of a syntactic frame and the three acceptable sentences:

> Syntactic frame:
> DETERMINER +NOUN +VERB +DETERMINER +NOUN +MODIFIER +PREPOSITION +DETERMINER +NOUN
>
> Control:
> The cabs hauled the visitors straight to the restaurant.
> Alveolar stop (tongue twister):
> The taxis delivered the tourists directly to the tavern.
> Alveolar fricative (tongue twister):
> The sparrow snatched the spider swiftly off the ceiling.
> (McCutchen et al. 1991, p. 91)

Unacceptable sentences, on the other hand, were constructed by changing content words across sentences within a given consonant type, though the original syntactic frame was maintained. Thus, these sentences were semantically unacceptable but the syntactic structure was the same as acceptable sentences: e. g. The telegram delighted the tent differently to the tutor.

Experiment 2 was designed to decide whether the tongue twister effect was observed during sentence comprehension or during lexical access. McCutchen et al. (1991) hypothesized that there would be no tongue twister effect in a lexical decision task which does not require comprehension and memory. They used a modified lexical decision task on lists of tongue twisters. Three sets of word displays were constructed: a stop set of words with initial alveolar stops (/t/ and /d/), a fricative set of words with initial alveolar fricatives (/s/ and /z/), and a control set of words with various initial phonemes. They also varied the number of words from 4 to 8 in order to observe any interference in longer displays. There were positive trials which contained a nonword, and negative trials which did not contain any nonwords. Nonwords were constructed based on actual words by substituting a single letter with a letter which is similar in its shape. For example, the nonword *seire* was constructed from the real word *seize.* Here is an example of the display of words with initial alveolar fricatives (positive trial): *sound serve start secede says stay seire settle* (p. 98). The subjects' task was to judge whether there was a nonword or not in the word strings.

The results were as follows. In experiment 1, they obtained a significant tongue twister effect: i. e. the reaction time for tongue twister sentences was significantly longer than for control sentences. On the other hand, in Experiment 2, there was no significant tongue twister effects. That is, the subjects did not take a longer time for tongue twister displays than for control displays. Based on the results of these two experiments, McCutchen et al. (1991) concluded that the tongue twister effect was truly phonological, and also initial consonants were part of phonological codes. In addition to this, their reference-and order-securing hypothesis was again supported. They confirmed that phonological recoding has its effect considerably in the comprehension processes, not in the processes involved in lexical access.

## 4. Experiments

In the present study, I adopt both the sentence verification task containing tongue twister sentences, and the lexical decision task containing tongue twister word strings.

## 4.1. Experiment 1

### 4.1.1. Method

*Materials*

I replicate the syntactic frame in McCutchen et al. (1991), using alveolar stops (/t/ and /d/), and velar stops (/k/ and /g/) in order to reduce the visual similarity of tongue twister sentences (e. g. The detective discovered the danger and decided to dig for details). The reason why /k/ and /g/ are used instead of /s/ and /z/ is because the /k/ sound has more variety of letters (c, k, q) than the /s/ sound (c, s), so that it will lessen the visual similarity of tongue twister sentences.

It should be remembered here that there is another problem in tongue twister sentences besides visual similarity. Perfetti & McCutchen (1982) pointed out that "the control sentences (of McCutchen & Perfetti, 1982) turned out to be more easily processed for some semantic reason" (p. 255). This may be because tongue twister sentences tend to be rather awkward with their phonetic restrictions, whereas control sentences are semantically more natural. In order to rule out semantically bizarre sentences, McCutchen et al. (1991) collected pilot acceptability ratings. In the present study, four native speakers of English were asked if some of the tongue twister sentences (both acceptable and unacceptable) were bizarre or not, and the bizarre sentences or ambiguous sentences were eliminated before the experiment.

There remained 42 sentences in total, and there were 14 sentences for each sentence type: /t/ and /d/ tongue twister, /k/ and /g/ tongue twister and control sentences. Half of them (7) were acceptable and the other half were unacceptable (see Appendix 1). All the sentences were in the same syntactic frame as described in the previous section. The acceptable control sentences were matched with the tongue twister sentences in terms of the familiarity of vocabulary and semantic acceptability. The number of syllables in each sentence varied from 8 to 13, but the total number of syllables in each sentence type were made almost parallel: 72 syllables for the acceptable alveolar stop sentences, 72 syllables for the acceptable velar stop sentences and 71 syllables for the acceptable control sentences.

### *Subjects*

Sixteen native speakers of Japanese (age 24-40) who had been studying at a postgraduate level at Lancaster University for nearly 10 months served as subjects. All of them had studied English as a foreign language in school settings for more than ten years. Their degree levels and majors are: M. A. in International Relations (1), M. A. in Linguistics (3), M. A. in Women's Studies and English (2), M. B. A. (2), MSc in Systems and Information Management (1), Diploma in Historical Studies (1), Diploma in Business Analysis (2), Diploma in International Relations (4). Their reading comprehension scores in IELTS test administered by the British Council were 6.0-7.5. These intermediate and advanced level learners were considered to know at least the basic structures of English and vocabulary selected for the experiment.

### *Procedure*

Subjects were tested individually, and each subject was given an explanation about the experiment before he/she began. Instructions were given in Japanese. First of all, the experimenter told him/her that their task was to judge the acceptability of sentences which would appear one by one on the computer screen. Then, it was explained that "acceptability" meant "semantic acceptability," not "syntactic acceptability," showing them examples written on a card. The following are the examples for acceptable and unacceptable sentences:

Acceptable

*The man takes pleasure in taking pictures of flowers.*

Unacceptable

*The roof takes pains in reading clothes of mine.*

This pair of sentences were shown one by one in this order, confirming that even though the two sentences were syntactically identical, the latter was obviously unacceptable because of its semantic anomaly. The subjects were also asked to judge acceptability in an "ordinary sense": i. e. they should not be so imaginative as to accept "The roof took pains ... "

Subjects were instructed to press the left button on a computer mouse if the sentence was meaningful and the right button if it was not. They were instructed to make their responses as quickly as possible, while maintaining accuracy. Each subject received 4 practice sentences on the computer screen. The test sentences were presented in the same order for each subject. Stimulus presentation was controlled by WINDOWS microcomputer software equipped with a real-time clock. The sentences were presented individually in lower case letters in the upper part of the screen.

### 4.1.2. Results

The mean reaction times (RTs) and corresponding mean percentage error rate for the six categories are presented in Table 1.

**Table 1**

Mean Reaction Times in Seconds (and Percentage Correct Scores) for Acceptable and Unacceptable Sentences

| Acceptable | | | Unacceptable | | |
|---|---|---|---|---|---|
| /t/ /d/ | /k/ /g/ | control | /t//d/ | /k/ /g/ | control |
| 5.21<br>(90.2) | 5 .72<br>(86.6) | 4.18<br>(94.6) | 6.09<br>(94.6) | 6.75<br>(90.2) | 5.90<br>(91.7) |

The RT data were submitted to 2 × 3 analysis of variance (ANOVA), with sentence acceptability and sentence type (alveolar stop tongue twister, velar stop tongue twister, and control) as the two independent variables.

The RT analyses were based on correct responses only. The ANOVA for RTs revealed significant main effects of sentence acceptability and sentence type. Judgments were made significantly more quickly for acceptable sentences [ $F(1, 36) = 26.10$, $p < 0.01$ ]. Also, there was a significant tongue twister effect: judgments were made more quickly for control sentences than for alveolar stop or velar stop tongue twisters [ $F(2, 36) = 8.47$, $p < 0.01$ ].

The interaction between acceptability and sentence type was not significant [ $F(2, 36) = 1.17$, *n. s.* ]. Post hoc pairwise comparisons with Tukey's procedure further revealed that the tongue twister effect was more clearly observed in acceptable sentences. For acceptable sentences, there was a significant difference between control and /t/ & /d/ sentences ($p < 0.05$), and also a significant difference between control and /k/ & /g/ sentences ($p < 0.05$). No significant difference was found between the two types of tongue twister sentences. For unacceptable sentences, on the other hand, a tongue twister effect was observed only between control and /k/ & /g/ sentences ($p < 0.05$).

The error data were submitted to analysis of deviance. There was no significant difference between sentence types [change in deviance is 2.93, cf. $x_2^2$ , *n. s.* ], nor between acceptability [change in deviance is 0.30, cf. $x_1^2$ , *n. s.* ].

### 4.1.3. Discussion

The results of Experiment 1 provide evidence that there was a tongue twister effect in the reading processes of Japanese students. Analysis of the reaction times for acceptable sentences showed a significant main effect for sentence type: there was a significant difference between the control sentences and the two sets of tongue twister sentences, which did not differ from each other.

This result is perfectly consistent with McCutchen et al.'s original study.

As for unacceptable sentences, however, there is a discrepancy between the original study and the present study. Whereas a tongue twister effect was observed between the velar stop sentences and the control sentences in the present study, McCutchen et al. (1991) did not find this difference. This result of McCutchen et al. (1991) was contradictory to that of McCutchen & Perfetti (1982), and they attributed this difference to the various response strategies brought about by different instructions. In McCutchen et al. (1991), their instructions to subjects stressed accuracy over speed, so that the subjects used "a more conservative response strategy" (p. 96).

This explanation might be applicable to the present study. Even though there was a tongue twister effect in unacceptable sentences in the present study, it was partial (only between /k/ & /g/ sentences and control), and the difference was not very dramatic. It appears that the tongue twister effect in unacceptable sentences is not so reliable as that in acceptable sentences. It depends on what strategy people use. If people use a conservative strategy and read the same sentence several times, the tongue twister effect might be less robust.

The tongue twister effect in acceptable sentences in Experiment 1 suggests that Japanese students may make use of word-initial phonemes (consonants) when they read English sentences, just as native speakers of English do. Of course, as McCutchen & Perfetti (1982) suggested, word-initial phonemes are not the only information they can use. Yet, it might be reasonable to infer that such consonants are one part of the phonological code for the reading processes of second language learners as well.

## 4.2. Experiment 2

### 4.2.1. Method

*Materials*

A modified lexical decision task on lists of tongue twister word strings was used in Experiment 2. Two sets of word displays were constructed: a set of words with initial alveolar stops (/t/ and /d/) and a control set of words with mixed initial phonemes. There were 5 words in each set. This display size is chosen because it is neither too short nor too long. The display size is not varied, as McCutchen et al. (1991) reported that there was no interaction between the length and any of the variables. The words in the alveolar stop set and the control set were matched for part of speech and number of syllables.

Trials consisted of positive trials (i. e. a nonword was present) and negative trials (i. e. all real words). Nonwords were constructed in the manner of McCutchen et al. (1991). Each positive trial was linked to a corresponding negative trial, with word sequences matched for part of speech and number of syllables. Nonwords were designed to occur equally often in each serial position.

There were 32 trials, half of which (16) were positive trials and the other half negative trials. There were 8 trials for each of four categories (/t/ & /d/ positive, /t/ & /d/ negative, control positive, control negative) (see Appendix 2 ).

*Subjects*

The same 16 Japanese in Experiment 1 served as subjects.

*Procedure*

Experiment 2 followed Experiment 1 in the same session.

Subjects were tested individually and instructions were given in Japanese. Subjects were instructed to press the left button if the word string did not contain any nonwords and the right button if it did. The experimenter gave this instruction to the subjects, showing them the following examples written on cards:

Negative trial
*get* *hold* *imagine* *marry* *obey*
Positive trial
*cat* *bird* *attitude* *pudzej* *usage*

Each subject received 2 practice displays on the computer screen. The word displays were presented in the same order for each subject. Stimulus presentation was controlled by WINDOWS microcomputer software equipped with a real-time clock. Words were displayed in lower case letters in a single row.

### 4.2.2. Results

The mean reaction times (RTs) and corresponding mean percentage error rates for the four categories are presented in Table 2.

**Table 2**

Mean Reaction Times in Seconds
(and Percentage Correct Scores for Word Displays)

| Negative | | Positive | |
|---|---|---|---|
| /t/ /d/ | control | /t/ /d/ | control |
| 2.58<br>(98.4) | 2.71<br>(99.2) | 2.27<br>(96.9) | 2.18<br>(97.7) |

Since the error rate was very low in Experiment 2 (less than 4 %), RTs were the major focus of analysis. The RT analyses were based on correct responses only. The RT data were submitted to 2 × 2 ANOVA with word type (alveolar stop and control) and word/

nonword (negative and positive) as independent variables. The ANOVA for RTs yielded a significant main effect of word/nonword [$F(1, 28) = 5.59$, $p<0.05$], but did not yield any significant effect of word type [$F(1, 28) = 0.01$, *n. s.*]. There was no significant interaction between word type and word/nonword. In other words, judgments were made more quickly for positive trials than for negative trials, but there was no significant difference between the two word types.

### 4.2.3. Discussion

Unlike Experiment 1, Experiment 2 did not show any tongue twister effect. The significant difference was observed only between negative and positive trials. Post hoc pairwise comparison with Tukey's procedure revealed that there was a significant difference between the negative and positive trials of control word displays ($p<0.05$). It is reasonable that judgment for negative trials took longer than for positive trials. For negative trials, subjects had to read all the words until the end, whereas for positive trials subjects could make a judgment when a nonword appeared. If a nonword occurred in the first position of the string, for example, judgment might have been made very quickly.

Of key interest is that there was no tongue twister effect. This result is consistent with McCutchen et al.'s original study. It suggests that the phonological processes reflected in the tongue twister effect are localised in the comprehension and memory processes in L2 reading, not in lexical access. Even though the word strings began with /t/ or /d/ such as *tail*, *difficult*, *dull*, *trivial*, *dirty*, this phonological similarity did not affect the Japanese subjects' performance this time, because subjects did not have to connect these word strings into a meaningful sentence. Thus, they did not draw on phonological information when they read a single word. To put it in another way, phonological information is most effective during the memory and comprehension processes involved in sentence processing, not during processes involved in isolated word reading.

## 5. General Discussion and Pedagogical Implications

In the present study, it was revealed that there was a tongue twister effect in the sentence verification task, whereas there was no such effect in the lexical decision task. This result indicates that phonological information is not especially effective for lexical access, but is useful for sentence comprehension in L2 reading. This finding is consistent with the general view in L1 studies that the main role of phonological recoding at a sentence level is considered to be to aid higher-order comprehension processes in reading.

How phonological information is related to working memory has long been investigated in L1 studies. Baddeley (1966), for example, investigated how acoustic, semantic and visual similarity affected memory task. He discovered that there was a large adverse effect of acoustic similarity on the memory task both in visual and auditory presentation. In other words, memory-related task is rather sensitive to phonetic similarity:

> ... subjects show remarkable consistency and uniformity in using an almost acoustic coding system for short-term remembering of disconnected words.
>
> (Baddeley 1966, p. 364)

(The concept of working memory had not been formed at this time, and therefore Baddeley used the word "short-term remembering.")

In fact, there is a line of evidence that working memory may rely on acoustic coding (i. e. phonological information) (Conrad, 1964; Conrad, 1972). Then, it is conceivable that those who are poor at phonological recoding cannot be very good readers, because these people cannot make use of phonological information in working memory to reinterpret an earlier part of a sentence in light of words that occur later in the sentence. In the field of L1 acquisition studies, a lot of research have been done on phonological recoding and working memory in terms of reading comprehension ability (i. e. good readers and poor readers). Now that it was made clear that phonological recoding is most effective in the comprehension processes of L2 reading as it is in L1 reading, it would be relevant to gain insights from the L1 studies.

Shankweiler et al. (1979) investigated how good and poor readers were differentiated in terms of recall of random letter strings. Their subjects were good readers, marginal readers and severely backward readers. All of them were school children. There were two kinds of stimuli: rhyming consonants (e. g. B, C, D, G, P, T, V, Z) and nonrhyming consonants (e. g. H, K, L, Q, R, S, W, Y). They discovered that the recall performance of good readers was more affected by phonetic confusability (similarity) than that of marginal readers. Severely backward readers showed a weaker effect of confusability. They concluded that good and poor readers differ in their use of phonetic codes in working memory, suggesting that memory-related problems of poor readers may be manifestations of deficiencies in phonological recoding.

Related to this, Katz et al. (1981) examined whether poor readers' problem with order memory was related to an underlying deficiency in the use of phonetic codes. They differentiated in their ability to order stimuli that could be easily recoded as phonetic form, but not in their ability to order nonlinguistic stimuli that could not be phonetically recoded. Their stimuli were nonsense drawings which were "phonetically unrecodable" and drawings of common objects, which they called "phonetically recodable". As they expected, they found that good readers' performance was significantly better than poor readers' in ordering stimuli that were phonetically recodable. They concluded from these findings that phonological recoding was effective for the retention of order information, and poor readers' problem with order memory was related to deficiencies in their use of phonetic codes.

From the pedagogical point of view, it appears that ability in phonological recoding is a key issue for efficient reading. Shankweiler & Crain (1986) discussed the role of phonological recoding in working memory in more detail. According to their interpretation, working memory has two working parts concerning language processes: a storage buffer and a central executive. In the storage buffer, rehearsal of phonetically coded material can take place. A second component is a "control mechanism" which is capable of integrating information from the phonological, syntactic and semantic parsers:

> It (the control mechanism) facilitates the organization of the products of lower-level processing by relaying information that has undergone analysis at one level to the next higher-level. The first duty of the control mechanism is to transfer phonologically analyzed material out of the buffer and push it upwards through the higher level parsers, thus freeing the buffer for succeeding material.
>
> (Shankweiler & Crain 1986, p. 150)

Based on this model of working memory, they supposed that language related problems of poor readers might be due to two properties of working memory system: (1) limitations of the working memory system supporting the analysis of input and (2) the dependence of higher-level processing on lower-level (orthographic and phonological) analysis of the contents. Thus, they assumed that unless the phonological analysis of letter strings was pursued efficiently, higher-level analysis would be inhibited altogether. In the end, they reached the conclusion that the problem of learning to read is located in the process of phonological analysis of orthographic input, and this process demands "the construction of algorithms" for relating orthographic structure to phonological structure:

> To construct this interface is an intellectual task, which

> requires overt attention and metalinguistic knowledge that doesn't come free with language acquisition. Until an entire set of analytic metaphonologic strategies are practiced enough to become largely automatic, higher-level processing will be curtailed because working memory is overloaded.
> (Shankweiler & Crain 1986, p. 164)

They emphasised that automatic decoding at lower-level analysis would be crucial for efficient reading. In fact, the importance of automatic decoding for skilled reading had been verified by Perfetti & Hogaboam (1975). They examined the difference between good and poor readers in vocalisation latencies to single printed words. The subjects were third-grade and fifth-grade children. They were told to say each word presented on a slide as quickly as possible, but at the same time they were told that they should not attempt to say the word until they "know what it was." Perfetti & Hogaboam (1975) found that good readers were significantly quicker for pseudowords and low frequency English words. There was no significant difference in vocalisation latency for high frequency words. Also, they discovered that knowledge of word meanings might be a less significant factor for good readers than for poor readers. They suggested that poor readers might have failed to develop automatic decoding skills. Perfetti & Hogaboam (1975) explained the different reading processes between good readers and poor readers as follows:

> ... If immediate recognition fails, as it does with less familiar words, true code breaking is engaged. It is here that good readers and poor readers are maximally different. The good reader does this quickly and automatically on the basis of well-learned skills that take advantage of letter and sound redundancies, while the poor reader does this with effort and not automatically. (Perfetti & Hogaboam 1975, p. 468)

It seems that this automatic decoding skill is also crucial for L2 reading, because the reading processes of advanced L2 learners are considered to be quite similar to the processes of native speakers of English, and therefore, it is reasonable to infer that the problems of poor readers in L1 might be comparable to L2 reading. Then, the point is to explore how L2 learners become good at automatic decoding. As Shankweiler & Crain (1986) pointed out, phonological analysis of orthographic input is an intellectual task which requires metalinguistic knowledge. It implies that conscious learning is needed for the acquisition of spelling-sound correspondences.

In the field of L1 studies, it was pointed out that "phoneme awareness" is the most important prerequisite for reading, because it helps children gain access to spelling-sound correspondences. Treiman & Baron (1983) attempted to demonstrate that phoneme-analysis training would help children take advantage of spelling-sound correspondences. In their experiments, preschool children participated in a phoneme-analysis condition and a control condition. It was revealed that those who were trained to segment phonemes performed better in reading words specifically using the spelling-sound correspondences related to the trained segments. Bradley & Bryant's study (1983) is along the same line. They also investigated children's ability to analyse syllables into their constituent phonemes. They reported that a group of children who were given the phoneme-analysis training at a pre-reading stage were consistently better on tests of reading and spelling three years later. They provided clear evidence that making young children aware of phonemes helps them learn to read. Since some researchers claimed that phoneme awareness would benefit from learning to read, as well as vice versa (cf. Ehri, 1983; Perfetti et al., 1987), it is now agreed that "phonological awareness and the development of reading skill (for alphabetic writing systems) go hand in hand" (Rayner & Pollatsek 1989, p. 346).

As it is difficult for English speaking children to have phoneme awareness and discover the alphabetic principle, it might be also hard for non-native speakers to be familiar with phoneme segmentation and alphabetic principle. As far as Japanese speakers are concerned, the fact that their mother tongue is not alphabetic may make matters even worse. Read et al. (1986) confirmed that this phoneme segmentation skill does not develop with congnitive maturation or non-alphabetic literacy, but does develop in the process of learning to read and write alphabetically.

Mann (1986) examined phoneme awareness of Japanese and American children. Since Japanese Kana is syllabic, Mann (1986) reasoned that Japanese children should be aware of syllables, whereas American children should be aware of both phonemes and syllables. Using a syllable counting and a phoneme counting test, she found that in contrast to first-grade children in America, who were aware of both syllables and phonemes, almost all first graders in Japan were aware of syllables, but were not aware of phonemes. She further investigated the syllable and phoneme awareness of Japanese children in the later elementary grades. The data revealed that the majority of Japanese children were able to manipulate both syllables and phonemes by the fourth grade. She inferred that this phoneme awareness of Japanese children might be promoted by the experience of learning Kana, which is a phonological orthography.

As Mann (1986) pointed out, it might be possible that Japanese Kana promotes phoneme awareness implicitly. The reason for this is that Japanese children learn Kana by means of a matrix in which all the characters in a row share the same vowel and all the characters in a column share the same consonant. It is possible that Japanese children notice that か ([ka]) and さ ([sa]) share the similar sound by referring to the matrix, and thus become aware of the initial phoneme.

It appears that Japanese speakers are in a better position than Chinese speakers to develop phoneme awareness in that Japanese has not only logographic Kanji but also syllabic Kana.

However, from the pedagogical point of view, it seems that explicit instruction is still needed particularly for beginning English readers in Japan. There are two reasons. Firstly, it is conceivable that whether or not children can develop phoneme awareness from Kana depends on how sensitive a child is to sound. It is possible that some children are very sensitive to sound structure, while others are not. Therefore, explicit instruction is needed particularly for those who are not sensitive to sound. The second reason is concerned with cognitive maturation. In Japan, English language teaching starts at a secondary school level. Students are 12 or 13 years old when they start to learn English. By this age, they are considered to have reached cognitive maturation to cope with an intellectual task which requires metalinguistic knowledge. They might benefit from the explicit instruction of spelling-sound correspondences in a less painful way than English speaking children do. It would be helpful at this stage if teachers introduced pronunciation of words in such a way that students can be more aware of spelling-sound correspondences.

Some researchers in Japan have introduced the phonics type of instruction at the very beginning of English teaching. Yamada et al. (1988) advocated the use of the Roman alphabet at a secondary school level to let students understand the spelling-sound correspondences of English words more easily. The Roman alphabet is taught at the end of fourth grade of elementary school to familiarise children with alphabetic letters. (In the Japanese language, alphabetic letters are used particularly for units of length, weight and so on, such as *cm* and *kg*.) The Roman alphabet represents the sound of each Kana. Since the Roman alphabet representations are basically syllabic, Yamada at al. (1988) first chose English words which can be segmented into a syllabic pattern of CV or V: e.g. *camera*, *banana*, *America*. Then, they segmented the word into syllables (ca/me/ra, ba/na/na, A/me/ri/ca), and pronounced them syllable by syllable, such as [kæ] [mə] [rə], so that students could recognise that each syllable corresponds with the sound. If students are familiar with the Roman alphabet, it is much easier for them to relate the segmented syllables with the sounds. Yamada et al. (1988) reported that the knowledge of the Roman alphabet coupled with their syllable-based instruction dramatically improved students' reading and spelling abilities.

The method of Yamada et al. (1988) is considered to be on the right track because they developed this method based on an awareness that automatic decoding skill is crucial for skilled reading. The significance of teaching sounds as a first step to automatic decoding should be reevaluated and probably should be more emphasised.

## 6. Conclusion

In the present study, I have explored the nature of phonological codes, and how they support comprehension in L2 reading. In Experiment 1, Japanese speakers took longer for the tongue twister sentences, which indicated that they activated phonemes when they read English sentences silently. Also, it was suggested that phonetic similarities might cause the kinds of similarity confusions in working memory, and thus the comprehension suffered. In Experiment 2, they did not show any tongue twister effect for the tongue twister word displays. This result demonstrated that when they read a single word, this phonetic similarity did not affect their performance. It was concluded that phonological information is particularly effective during comprehension processes involved in sentence processing, not during processes involved in a single word reading. It is assumed that when they read sentences, they have to integrate all information stored in working memory, and during this process, phonological information is of great help. This seems to be comparable to L1 reading processes.

Related to these findings, the problems of poor readers were discussed. Gaining insights from L1 studies, I supposed that one of the L2 readers' problems might be poor phonological recoding skill. In order to be good at phonological recoding, it is necessary for L2 learners to know the spelling-sound correspondences of the English alphabets. Since "phoneme awareness" is considered to be important to gain access to the alphabetic principle, the key issue is how to develop the "phoneme awareness" of logographic language users. It was suggested that Japanese Kana might be beneficial, and the Roman alphabet might be even more so.

From a theoretical point of view, there is a related issue which needs further investigation. The issue is the nature of inner speech in general as employed by L2 learners. As the main concern of this study was "phonological recoding," and not "subvocalisation," I did not touch upon it in the discussion. However, it might be also relevant to explore whether L2 learners subvocalise or not when they read English sentences. It is reasonable to infer that less skilled learners will depend more on subvocalisation, because it produces better copies which are more similar to actual speech (cf. Perfetti 1985), and thus makes the comprehension easier. If it is discovered that less skilled L2 learners depend more on subvocalisation, and skilled learners depend less, it might be clear evidence that the more learners become advanced, the less they depend on subvocalisation, and the more phoneme level activation takes place.

## References

Baddeley, A. D. (1966), "Short-Term Memory for Word Sequences as a Function of Acoustic, Semantic and Formal Similarity," *Quarterly Journal of Experimental Psychology*, 18, 362-365.

Bradley, L. & Bryant, P. E. (1983), "Categorizing Sounds and Learning to Read — A Causal Connection," *Nature*, 301, 419-421.

Conrad, R. (1964), "Acoustic Confusions in Immediate Memory," *British Journal of*

*Psychology*, 55, 75-84.

Conrad, R. (1972), "Speech and Reading," In J. F. Kavanagh & I. G. Mattingly (Eds.) *Language by Ear and by Eye: The Relationship between Speech and Reading* (pp. 205-240). Cambridge, Mass. : MIT Press.

Daneman, M. (1987), "Reading and Working Memory," In J. R. Beech & A. M. Colley (Eds.) *Cognitive Approaches to Reading* (pp. 57-86). New York: Wiley.

Doctor, E. A. & Coltheart, M. (1980), "Children's Use of Phonological Encoding When Reading for Meaning," *Memory & Cognition*, 8, 195-209.

Ehri, L. C. (1983), "Influence of Orthography on Phonological and Lexical Awareness in Beginning Readers," In J. Downing & R. Valtin (Eds.) *Language Awareness and Learning to Read*. New York: Springer-Verlag.

Katz, R. B., Shankweiler, D. & Liberman, I. Y. (1981), "Memory for Item Order and Phonetic Coding in the Beginning Reader," *Journal of Experimental Child Psychology*, 32, 474-484.

Kleiman, G. M. (1975), "Speech Recoding in Reading," *Journal of Verbal Learning & Verbal Behavior*, 14, 323-339.

Levy, B. A. (1977), "Reading: Speech and Meaning Processes," *Journal of Verbal Learning & Verbal Behavior*, 16, 623-638.

Mann, V. A. (1986), "Phonological Awareness: The Role of Reading Experience," *Cognition*, 24, 65-92.

McCusker, L. X., Hillinger, M. L. & Bias, R. G. (1981), "Phonological Recoding and Reading," *Psychological Bulletin*, 89, 2, 217-245.

McCutchen, D., Bell, L. C., France, I. M. & Perfetti, C. A. (1991), "Phoneme-Specific Interference in Reading: The Tongue-Twister Effect Revisited," *Reading Research Quarterly*, 26, 1, 87-103.

McCutchen, D. & Perfetti, C. A. (1982), "Visual Tongue-Twister Effect: Phonological Activation in Silent Reading," *Journal of Verbal Learning & Verbal Behavior*, 21, 672-687.

Mori, C. (1992), "Phonological Recoding in L2 Reading," *Manuscript*.

Perfetti, C. A. (1985), *Reading Ability*. New York: Oxford University Press.

Perfetti, C. A., Beverly, S., Bell, L. C. & Hughes, C. (1987), "Phonemic Knowledge and Learning to Read: A Longitudinal Study of First Grade Children," *Merrill-Palmer Quarterly*, 33, 283-319.

Perfetti, C. A. & Hogaboam, T. W. (1975), "The Relationship between Single

Word Decoding and Reading Comprehension Skill," *Journal of Educational Psychology*, 67, 461-469.

Rayner, K. & Pollatsek, A. (1989), *The Psychology of Reading*. London: Prentice Hall International.

Read, C., Zhang, Y., Nie, H. & Ding, B. (1986), "The Ability to Manipulate Speech Sounds Depends on Knowing Alphabetic Spelling," *Cognition*, 24, 31-44.

Segalowitz, M. & Hébert, M. (1990), "Phonological Recoding in the First and Second Language Reading of Skilled Bilinguals," *Language Learning*, 40, 4, 503-538.

Shankweiler, D. & Crain, S. (1986), "Language Mechanisms and Reading Disorder: A Modular Approach," *Cognition*, 24, 139-168.

Shankweiler, D., Liberman, I. Y., Mark, L. S., Fowler, C. A. & Fischer, F. W. (1979), "The Speech Code and Learning to Read," *Journal of Experimental Psychology: Human Learning & Memory*, 5, 6, 531-545.

Slowiaczek, M. L. & Clifton, C. (1980), "Subvocalization and Reading for Meaning," *Journal of Verbal Learning & Verbal Behavior*, 19, 573-582.

Treiman, R. A. & Baron, J. (1983), "Phonemic Analysis Training Helps Children Benefit from Spelling-Sound Rules," *Memory & Cognition*, 11, 382-389.

Waters, G. S. & Seidenberg, M. S. (1985), "Spelling-Sound Effects in Reading: Time-Course and Decision Criteria," *Memory & Cognition*, 13, 557-572.

Yamada, J., Matsuura, N. & Yanase, Y. (1988), *Eigo Gakuryokusa Wa Dokokara Shojirunoka* [Where does Individual Difference in Learning English Come From?]. Tokyo: Taishukan.

# Appendix 1

Tongue Twister and Control Sentences for Experiment 1

/t/ and /d/

Acceptable

1. The teacher taught the text dramatically in the tutorial.
2. The taxis delivered the tourists directly to the town.
3. The daughter tried the dress delightedly in the dress shop.
4. The daughter dropped the toy deliberately in the toilet.
5. The typist typed the document twice at the desk.
6. The doctor did the task tolerantly in the dark.
7. The diplomat told the truth tonight to the detective.

/k/ and /g/

Acceptable

1. The guardsman closed the gate quietly for the conference.
2. The cook carried the glasses cautiously to the guest.
3. The clerk confirmed the quality confidently to the customer.
4. The grandfather greeted the guest gloomily in the corridor.
5. The conductor guided the car cautiously into the garage.
6. The grandmother gave the cakes gladly to the gardener.
7. The kid kicked the can carefully to the corner.

Control

Acceptable

1. The actress played the role impressively on the stage.
2. The politician told a lie halfway through the interview.
3. The lecturer taught the novel interestingly in the class.
4. The cab took the visitors straight to the hotel.
5. The man gained a reputation steadily with the work.
6. The girl broke the vase intentionally at the party.
7. The scientist continued the work patiently through the night.

/t/ and /d/

Unacceptable

1. The driver typed the tourists directly in the tunnel.
2. The doctor told the toy down at the desk.
3. The typist dropped the daughter delightedly in the tutorial.
4. The diplomat did the dress tolerantly to the town.
5. The task delivered the truth twice in the dress shop.
6. The taxis taught the teacher dramatically in the toilet.
7. The text tried the detective deliberately to the traffic.

/k/ and /g/

Unacceptable

1. The grandfather closed the cakes confidently in the quiz.
2. The clerk gained the gate carefully in the corridor.
3. The grandmother kicked the quality gradually to the guest.
4. The guest confirmed the cook gladly to the corner.
5. The can carried ground quietly to the gardener.
6. The competitor gave the conductor cautiously for the conference.
7. The car greeted the guardsman cautiously to the customer.

Control

Unacceptable

1. The girl played the vase steadily to the hotel.
2. The politician baked the role interestingly at the party.
3. The scientist told cookies straight on the stage.
4. The reputation taught the lady impressively through the night.
5. The cab broke a lie patiently with the work.
6. The novel gained the lecturer halfway for the children.
7. The actress continued the visitors occasionally in the class.

# Appendix 2

Tongue Twister and Control Word Displays for Experiment 2

/t/ and /d/

Positive Trials

1. tea <u>toagm</u> dance deal door
2. damage temple district <u>dewazu</u> travel
3. <u>tarhej</u> dozen terminal two dust
4. dog date <u>teh</u> tent tear
5. tempt deny <u>digfeh</u> digest turn
6. tank taste town tulip <u>dankeup</u>
7. decide develop try dominate <u>tegu</u>
8. <u>deob</u> terrible tall delicate double

Negative Trials

1. death doubt tree trick tool
2. telephone tragedy debt tomorrow doll
3. trip terror tonight degree test
4. draft drill train task trade
5. temper today twenty data despair
6. trouble detail departure tone text
7. trust twelve tooth duty triumph
8. tough difficult dull trivial dirty

Control

Positive Trials

1. <u>nihkt</u> fact home meal bed
2. middle novel problem <u>repabl</u> sister
3. feeling <u>brozkeh</u> criticism bear fight
4. bread chair <u>sgk</u> net face
5. pay become <u>rezemd</u> finish send
6. bank fault hat kettle <u>lankeagu</u>

7. cultivate satisfy shut recognize peajk
8. bzacg convenient red favorite heavy

Negative Trials
1. bread mouth god fate coin
2. candle paper station record neighbor
3. monkey respect conclusion place stuff
4. phone key grape rain joy
5. pick succeed depend complete walk
6. gold wall tray member window
7. procedure government soap reference knife
8. hard popular short beautiful nervous