

At-lexical, articulatory interference in silent reading: The “upstream” tongue-twister effect

DANIEL H. ROBINSON and ANDREW D. KATAYAMA

Mississippi State University, Mississippi State, Mississippi

In two experiments, we investigated the interpretation and boundary conditions of the tongue-twister (TT) effect in silent reading. Previously, McCutchen, Bell, France, and Perfetti (1991) observed a TT effect when students made semantic acceptability judgments on sentences, but not when they made lexical decisions on lists of words. Using similar methodology in Experiment 1, along with two changes (using “better” TTs and longer word lists), we observed a TT effect in a lexical decision task. In Experiment 2, a memory span task revealed that students recalled fewer words from TT lists than from control lists. These results suggest that the basic mechanism of the TT effect may be articulatory, rather than working-memory, interference that occurs during lexical access and resurfaces post-lexically, inhibiting efforts to maintain the temporal order of several words.

The debate over the functions of phonology in adults' silent reading has raged on over the past several years. It has become “the single dominant theoretical issue in the psychology of reading” (Crowder & Wagner, 1992, p. 157). Recently, Berent and Perfetti (1995) stated that the most important function of phonology is assisting sentence comprehension. Phonology facilitates this process by helping retain the temporal order of words in working memory so that they can be syntactically and semantically analyzed (Johnston, Thompson, Fletcher-Flinn, & Holligan, 1995). Another possible function of phonology is in the identification of isolated words, by activating phonological representations to access lexical meanings. Thus, phonology may serve two functions in reading: (1) identification of the printed word and (2) temporary storage of several words while sentence comprehension takes place. Although researchers are in agreement regarding the latter function, the former function remains controversial (Zhang & Perfetti, 1993).

The Tongue-Twister Effect

One window that may shed some light on this issue is the tongue-twister (TT) paradigm. The rationale here is that if silent reading is dependent on phonology, readers will take longer to read sentences that contain words that share similar initial phonemes (e.g., “The Swiss wrist-watch strap shop shuts soon.”) than they would to read sentences that contain a natural mix of initial phonemes (e.g., “The brown bearskin rug man left town.”) because the former are more difficult to articulate (Haber & Haber, 1982). This TT effect has been demonstrated in studies in which subjects simply read sentences silently (Haber & Haber, 1982), in which subjects read and were required to comprehend the sentences (Ayres, 1984), and in which

various groups made semantic acceptability judgments on sentences (American students who were deaf, Hanson, Goodell, & Perfetti, 1991; Chinese students, Zhang & Perfetti, 1993; and American students with normal hearing, McCutchen & Perfetti, 1982).

One possible explanation for the TT effect, first offered by Baddeley and Lewis (1981), is that readers simply lose their places when reading TT lists due to the visual similarity of the initial letters of the words. Whether the TT effect results from visual-grapheme similarity among words or phonemic similarity has been addressed by McCutchen and Perfetti (1982) and McCutchen et al. (1991), who partially controlled for the possible confounding of graphemes and phonemes by mixing letters that represented voicing contrasts (e.g., /s/ and /z/) within TT sentences. They found evidence that the TT effect results from phonemic, rather than visual, interference. This visual versus phonemic interference debate was unequivocally resolved by Zhang and Perfetti (1993), who found a TT effect using Chinese characters, completely controlling for the confounding between visual and phonemic similarity. In each of the cited studies, TT effects were demonstrated using sentences as stimulus materials, providing evidence that phonological activation during sentence comprehension causes interference among similar phonological representations, resulting in longer reading times, more errors, or both.

There has been less evidence concerning whether phonological activation during individual word recognition may cause similar interference. To address this question, in a second experiment, McCutchen et al. (1991) had students make lexical decisions on lists that contained either TTs or control words. No TT effect was found, and they concluded that although phonology is automatically activated as part of lexical access, the function of this activation appears “downstream” during sentence comprehension, and not during individual word recognition. They explained the TT effect by proposing a model of reading

Correspondence should be addressed to D. H. Robinson, School of Education, University of South Dakota, Vermillion, SD 57069 (e-mail: drobinso@sundance.usd.edu).

in which phonemically indexed representations are stored in working memory. This information helps to distinguish words from one another and also aids in keeping a word in its proper temporal order in the sentence. These newly established representations can then be used in the higher level process of comprehension. When persons are required to read TT sentences, the phonemic representations in working memory are very similar, and the ability to differentiate the words is weakened—leading to difficulty in comprehension (McCutchen et al., 1991).

The important point here is that McCutchen et al. (1991) claimed that the TT effect occurs only within working memory and not in isolated word reading. They argued that because accessing one word at a time in the lexicon does not require maintenance of each word's reference and order within working memory, phonological activation during lexical access does not reach a level high enough to cause interference. This level, they proposed, is reached later, during the memory and comprehension phases of reading. The present study represents another attempt to investigate whether the TT effect occurs during individual word recognition. This study was motivated by a concern that perhaps McCutchen et al. did not find a TT effect with a lexical decision task because of two shortcomings in the word lists they used.

“Better” Tongue Twisters

The first shortcoming may have been the type of TTs McCutchen et al. (1991) used. The following is an example of a TT list used in their study: *sound serve start secede says stay seize settle*. According to *Webster's New World Dictionary*, a TT is a phrase that is hard to speak rapidly. Reading a list of words that have the same initial phonemes would seem to take longer than reading a similar list of words with different initial phonemes. However, consider how much longer it would take to read the following list, which contains words with similar initial articulatory, rather than phonemic, units: *brackets blasting brief bloom breeze blinked brittle*. Compare the rate at which you can read the previous list with the rate at which you can read the following list: *thunder widow proof flash prize glanced function*. It would appear that articulating some lists of words may twist the tongue more than others. Rather than simply confusing the reader with words containing the same initial phoneme, the present study used words with similar initial articulatory units (two phonemes). Alternating their placement throughout the lists would make it difficult for readers to articulate them quickly. It was believed that these lists would be “better” TTs than those used by McCutchen et al. on the basis of the assumption that they take longer to articulate.

Recently, Berent and Perfetti (1995) proposed a two-cycles model of phonology assembly. It assumes that there is always an initial, at least partial, activation of phonology even for high-frequency words. This model states that activation of consonant sounds comes first, followed by a slower, more controlled activation of vowel information. Accordingly, it would seem probable that

lists like the ones used in the present study, which alternate words with two similar initial phonemes, would be more sensitive to articulatory interference than would lists with one similar initial phoneme.

Longer Lists

A second shortcoming of McCutchen et al.'s (1991) study may have been the length of the word lists. They used lists containing four and eight words to see whether phonological activation would build up over successive lexical accesses (i.e., reading a longer list of phonemically similar words would seem to be more difficult than reading a shorter one). However, if phonemic interference is cumulative, lists containing only four and eight words may have been too short to allow for detection of any interference in reading times. To avoid this potential problem, the present study used lists containing 10 words.

By incorporating these two changes, we believed that we would be better able to detect articulatory interference, although we did not include any conditions that matched the methodology used by McCutchen et al. (1991); thus we cannot make a direct comparison. Our purpose here was *not* to identify potential methodological weaknesses in a previous study, but rather to incorporate a few changes to determine whether the TT effect occurs during lexical access. That is, does phonology play an “upstream” role in reading?

EXPERIMENT 1

The main purpose of Experiment 1 was to demonstrate a TT effect in a lexical decision task, thereby challenging McCutchen et al.'s (1991) conclusion that “the TT effect is not sensitive to such lexical-level activation” (p. 101). We felt that the null results in their second experiment were probably due to their use of bare phonemes rather than articulatory units in their TT lists. In other words, we predicted that a TT effect would appear in a lexical decision task as a result of articulatory, rather than phonemic, interference. This interference would result in increased response times in lexical decisions.

Method

Participants. Forty-two students enrolled in an undergraduate educational psychology course at a large state university volunteered for course credit. Students were assigned to one of two counterbalancing conditions on the basis of order of appearance. Each student participated for about 15 min.

Materials. Six hundred words were selected that had frequencies (per million words) ranging from 1 to 50 (Carroll, Davies, & Richman, 1971). Two sets of 60 ten-word lists (Sets A and B) were constructed using the same 600 words. List size consisted of 10 words, which was the maximum possible because of the width of the Macintosh SE30 computer screen (7 in.). Each set contained 30 control lists containing words with mixed initial articulatory units and 30 TT lists containing words with similar initial articulatory units placed alternately in a list.

Because the task was to judge whether there was a nonword present in each of the lists, both positive lists (which included a nonword) and negative lists (which did not include any nonwords) were

constructed. Positive lists were fully counterbalanced so that nonwords occurred equally often in each serial position (1–10). Nonwords were constructed from actual words by substituting a single letter with a shape similar to that of the original letter. For example, the nonword *stumber* was constructed from the real word *slumber* by substituting *t* for *l*. Thus, all nonwords had the same orthographic shapes as those of actual words. Moreover, all nonwords were orthographically legal and therefore pronounceable. The following are examples of the lists used.

1. Positive list, control: *flared slim freight braces frosling pregnant sheets glasses shovel blasts*.

2. Positive list, TT: *frequencies flank freebom fleet frenzy flocked frogs fliers frantically flat*.

3. Negative list, control: *shores freshman glossy shearing frowned slang fling chest plaster grapes*.

4. Negative list, TT: *flagpole fractional flopping fragrant flick fresh flask frantic flier frown*.

Each 60-word set contained 15 negative controls, 15 negative TTs, 15 positive controls, and 15 positive TTs. The two sets of word lists were parallel in the sense that each contained the same 600 words. Thus, differences among words regarding part of speech, number of syllables, frequency, and temporal duration were controlled. For example, Lists 1–15 in Set A (negative controls) contained the same 150 words that were used in Lists 16–30 in Set B (negative tongue twisters). The only difference between TTs and controls in the two sets was the sequential arrangement of the words. The logic is that if equal numbers of students were randomly assigned to receive Sets A and B, and there were no differences in performance for the two sets, then the data could be collapsed across this counterbalancing condition and students' performance on TT and control lists could be directly compared.

Design and Procedure. The study consisted of one within-subjects factor (list type: TT vs. control) and one between-subjects factor (counterbalancing condition: Set A vs. Set B). Students were tested in a computer lab in groups of about 10. All materials were presented and data were collected via microcomputer (Costin, 1988). Students were informed about the nature of the study through a series of instructions that appeared on the screen. Word lists were displayed in lowercase, in a single row, from left to right, with one space between each word. The students' task was to read each list and to press the *z* key if a nonword was present and the */* key if there was no nonword present. Students were instructed to work as quickly as possible, although accuracy was stressed as being more important than speed. Students were given eight practice trials with corrective feedback before they began the exercise. Response times and proportion correct scores were recorded.

Analysis. Only data from negative trials were used because they permitted the cleanest interpretation (students had to view the entire list before making a decision). In the positive trials, TT and control lists were not matched in terms of the number of syllables that preceded the nonword; thus we could not be sure that differences in response times would be solely due to differences in word lists. Outliers in the response times data set were treated using the following procedure. Means and standard deviations were calculated for each student for both TT and control lists. Within each student–list type combination, response times that were more than 2 *SD* from the respective mean were removed. Outliers accounted for less than 5% of all responses.

Independent-samples *t* tests were first conducted separately on response times and proportions correct scores to test for the effect of counterbalancing condition. No significant differences were found; thus data were collapsed across this condition (there were equal numbers of students in each group). For each dependent measure (response times and proportions correct), two scores (control and TT) were obtained for each student. These two scores were contrasted within subjects via related-samples *t* tests. For all statistical tests, an alpha of .05 was used. Table 1 presents response times and proportions correct scores on tongue-twister and control lists.

Table 1
Means and Standard Deviations for Response Times and Proportion Correct Scores in Experiment 1 (*n* = 42)

List Type	Response Time		Proportion Correct Score	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Tongue twisters	6.59	1.94	.90	.10
Control	6.34	1.92	.91	.11

Results and Discussion

As predicted, for response times the TT effect was observed [$t(41) = 2.07$]. Students made faster decisions on control lists than on TT lists. Students took longer to judge lists when the sequential arrangement of word lists was manipulated. These results indicate that the TT effect is sensitive to lexical-level activation. For proportions correct, the effect for list type was not significant [$t(41) = .48, p = .82$]. This finding is important because it indicates that the faster response times for control lists were not achieved at the expense of lower accuracy.

A list-based, rather than subject-based, independent-samples *t* test was also conducted on the response times. Mean response times across all subjects were computed for both the set of 15 negative TT lists and the set of 15 negative control lists. For each of the 30 lists, response times that were more than 2 *SD* from the mean were removed; outliers accounted for 7.3% of all responses. A marginally significant TT effect was observed [$t(28) = 2.02, p = .054$], providing support for the robustness of this effect across both items and subjects.

These results suggest that we were able to conduct a more sensitive examination than McCutchen et al. (1991) of the function of phonology during word identification. Whether the lists used by McCutchen et al. contained poor TTs, were too short, or both is not of importance here. What is important is the evidence that the TT effect occurs "upstream," suggesting that phonological activation during lexical access is high enough to cause interference.

Although these changes in methodology permitted some clarification regarding the role of phonology in lexical access, they also called for a reexamination of the previous assumption that phonological information is also activated during the memory and comprehension phases of reading to the point where interference may occur. In each of the previous studies using TTs, participants read sentences. It is possible that the slower reading times for TT sentences were due to articulatory interference occurring during lexical access of each word and not during working memory's attempt to maintain the phonological representations until comprehension was accomplished. A better test of whether the TT effect is sensitive to working-memory interference would be to require participants to keep words active in working memory. Also, because the type of TTs used in the present study is arguably different from the TTs used in the second experiment of McCutchen et al. (1991), as indicated by the different results, it is possible that the type of interference observed in Experiment 1 may occur only during individual word identification and not in tasks requiring that several

words be kept active in working memory. This issue was addressed in Experiment 2.

EXPERIMENT 2

Experiment 2 was designed to see whether the type of articulatory interference observed in Experiment 1 would also occur in a memory span task. The finding that adult readers experience difficulty in short-term serial recall tasks when forced to remember lists of words that are phonemically similar, such as *tap pan fat fan pat*, as compared with words that are not similar in sound, such as *met nip cog gap bus*, has been consistently demonstrated in memory span experiments over the past two decades. This finding has been termed the “phonemic similarity effect” (Conrad & Hull, 1964). However, most of these studies have used words that either rhyme or share a vowel sound (as in the previous example) to create phonemic interference that is detrimental to a person’s ability to recall the words in their serial order.

There have been no studies investigating a phonemic similarity effect using TTs like the ones used in Experiment 1. If a phonemic similarity effect is found in a memory span task, it could be argued that the type of TTs used in the present study cause interference that is detected during individual word recognition and persists during the memory phase. However, if no phonemic similarity effect is found, it could be argued that the articulatory interference caused by TTs is limited to individual word identification, which would challenge the conclusion of McCutchen et al. (1991) that the TT effect occurs “downstream.”

Method

Participants. Eighteen students enrolled in an undergraduate educational psychology course at a large state university volunteered for course credit, none of whom had participated in Experiment 1. Students participated for about 10 min.

Materials. One hundred fifty words were selected from the original 600 words. Two sets (TT and control) of 30 five-word lists were then constructed using the same 150 words. The following are examples of the lists used.

1. Control: *shores freshman glossy shearing frowned.*
2. Tongue twister: *flagpole fractional flopping flagrant flick.*

The two types of word lists were parallel in the sense that each student was presented with the same 150 words twice. Again, the only difference between TTs and controls was the sequential arrangement of the words.

Design and Procedure. The design consisted of one within-subjects factor: list type (TT vs. control). Students participated as a group in a classroom setting. All stimulus materials were presented on an overhead projector. Words were displayed in lowercase, in a single row, from left to right, with one space between each word. TT and control lists were presented alternately. The students’ task was to read each list and to recall as many words as possible by writing them on numbered response sheets. The experimenter turned on the overhead light, and a list appeared for 5 sec. Then the experimenter turned off the light and the students were given 20 sec to write down the words. At the end of the 20 sec, the experimenter would say, “Stop,” indicating that the next list would appear in 5 sec and that students should quickly finish writing and direct their attention to the overhead screen. This was repeated until all 60 lists had been presented with 1-min breaks given about every 12 trials. All students

were given four practice trials with corrective feedback and indicated that they understood the task before they began the exercise.

Scoring. Responses were scored by a person who was blind to the purpose of the study. For each list, a student received a score ranging from 0 to 5 for the number of words correctly written. Only errors in plurality of a word were permitted. The scorer summed the number of words correctly recalled for both tongue-twister and control lists and then divided each by the number of lists (30).

Results and Discussion

A dependent-samples *t* test was conducted on the mean number of words recalled. Table 2 presents the means and standard deviations for the average number of words correctly recalled by list type. The phonemic similarity effect was observed [$t(17) = 2.67$, $SEM = .04$]. Students recalled more words from control lists than from tongue-twister lists. Thus, it appears that the phonological activation that interfered with attempts to recognize individual words with similar initial articulatory units in Experiment 1 also interfered with attempts to maintain those words in working memory.

GENERAL DISCUSSION

The purpose of this study was to demonstrate that the TT effect rests mainly on articulatory, rather than working-memory, processes. This is consistent with the strategy recommended by Ziegler and Jacobs (1995)—“to determine the tasks and boundary conditions, in which phonological codes provide sources of constraint for . . . word recognition” (p. 569). In Experiment 1, we had subjects make lexical decisions on lists of letter strings whose initial consonants were either phonemically similar (TTs) or not phonemically similar (controls). This task did not involve working-memory processes because subjects were simply required to check through the letter strings in the list, looking for words that were misspelled. The results suggest that the TT effect is sensitive to this type of single-word processing, without any memory or sentence-processing demands. In Experiment 2, we had subjects attempt to remember sets of five words that were either TTs or controls. The results suggest that the TT effect is also sensitive to multiple-word processing when subjects are required to remember the temporal order of several words.

Taken together, the results from both experiments suggest that the phonological interference observed in the memory span task in Experiment 2 may have simply been due to maintaining the phonological activation that was observed in Experiment 1 during lexical access. In other words, there very well could be only one mechanism for the TT effect—the articulatory interference ob-

Table 2
Means and Standard Deviations for Average Number of Words (Out of Five) Correctly Recalled in Experiment 2 ($n = 18$)

List Type	Words Recalled	
	<i>M</i>	<i>SD</i>
Tongue twisters	3.37	.82
Control	3.48	.85

served during lexical access, which also appears during working memory. This has previously been suggested by Zhang and Perfetti (1993) as a more parsimonious explanation for the TT effect, assuming that words are implicitly articulated in working memory.

In their recent book, Gathercole and Baddeley (1993) suggested that phonology "plays only a very minor role in reading at the single-word level" (p. 199). Our findings suggest that phonological processes play a greater role in lexical access than was originally thought. This view is supported by the recent work of Ziegler and Jacobs (1995) and Johnston et al. (1995). Ziegler and Jacobs had subjects indicate whether a target letter was present in a letter string. Subjects performed worse at this seemingly graphemic task when the letter string was a pseudohomophone (e.g., *TAIP*). Ziegler and Jacobs concluded that phonological information that is activated from printed words is an early and major constraint in visual word recognition. Johnston et al. used lexical decision and sentence evaluation tasks to evaluate the role of phonology during lexical access and sentence processing. They concluded that phonology functions to access lexical meanings in addition to processing sentences. Johnston et al. also found that prelexical use of phonology depended on the type of instruction learners received.

Perfetti, Zhang, and Berent (1992) have asserted that the issue of whether phonological information is used pre- versus postlexically is no longer useful in a theory of word identification. For example, although phonological coding has been observed in prelexical processes using the English writing system (e.g., Perfetti, Bell, & Delaney, 1988), prelexical phonology has not been found using the Chinese writing system (Perfetti & Zhang, 1991). Postlexical phonological processing, on the other hand, has been shown to be the same for English and Chinese (Zhang & Perfetti, 1993). The role of phonology in a word identification theory should simply be a characteristic of human reading processes rather than being writing system specific (Zhang & Perfetti, 1993). For this reason, we choose the term "at-lexical" to refer to phonological coding that occurs not before lexical access, but rather during the process of accessing the lexical meaning of a word, to apply generally to word identification regardless of the writing system.

As previously mentioned, McCutchen et al. (1991) proposed a model of reading in which phonetically indexed representations are first stored in working memory to help distinguish words from one another and aid in keeping the words in their proper temporal order in the sentence so that they can then be used in the higher level process of comprehension. This *reference-and-order-securing* hypothesis can be slightly altered to account for the findings presented here. The TT effect is due to pho-

nemic interference that occurs during articulatory, rather than working-memory, processes. Phonology, then, is used during lexical access to *reference* individual words and then is also used in working memory to *secure* the order of several words.

REFERENCES

- AYRES, T. J. (1984). Silent reading time for tongue-twister paragraphs. *American Journal of Psychology*, **97**, 605-609.
- BADDELEY, A. D., & LEWIS, V. J. (1981). Inner active processes in reading: The inner ear, inner voice, and inner eye. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 107-129). Hillsdale, NJ: Erlbaum.
- BERENT, I., & PERFETTI, C. A. (1995). A rose is a reez: The two-cycles model of phonology assembly in reading English. *Psychological Review*, **102**, 146-184.
- CARROLL, J. B., DAVIES, P., & RICHMAN, B. (1971). *Word frequency book*. New York: American Heritage.
- CONRAD, R., & HULL, A. J. (1964). Information, acoustic confusion, and memory span. *British Journal of Psychology*, **55**, 429-432.
- COSTIN, D. (1988). MacLab: A Macintosh system for psychology labs. *Behavior Research Methods, Instruments, & Computers*, **20**, 197-200.
- CROWDER, R. G., & WAGNER, R. K. (1992). *The psychology of reading: An introduction*. New York: Oxford University Press.
- GATHERCOLE, S. E., & BADDELEY, A. D. (1993). *Working memory and language*. Hove, U.K.: Erlbaum.
- HABER, L. R., & HABER, R. N. (1982). Does silent reading involve articulation? Evidence from tongue-twisters. *American Journal of Psychology*, **95**, 409-419.
- HANSON, V. L., GOODELL, E. W., & PERFETTI, C. A. (1991). Tongue-twister effects in the silent reading of hearing and deaf college students. *Journal of Memory & Language*, **30**, 319-330.
- JOHNSTON, R. S., THOMPSON, G. B., FLETCHER-FLINN, C. M., & HOLIGAN, C. (1995). The functions of phonology in the acquisition of reading: Lexical and sentence processing. *Memory & Cognition*, **23**, 749-766.
- MCCUTCHEEN, 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**, 87-103.
- MCCUTCHEEN, D., & PERFETTI, C. A. (1982). The visual tongue-twister effect: Phonemic activation in silent reading. *Journal of Verbal Learning & Verbal Behavior*, **21**, 672-687.
- PERFETTI, C. A., BELL, L., & DELANEY, S. (1988). Automatic phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory & Language*, **27**, 59-70.
- PERFETTI, C. A., & ZHANG, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 633-643.
- PERFETTI, C. A., ZHANG, S., & BERENT, I. (1992). Reading in English and Chinese: Evidence for a "universal" phonological principle. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 227-248). Amsterdam: North-Holland.
- ZHANG, S., & PERFETTI, C. A. (1993). The tongue-twister effect in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1082-1093.
- ZIEGLER, J. C., & JACOBS, A. M. (1995). Phonological information provides early sources of constraint in the processing of letter strings. *Journal of Memory & Language*, **34**, 567-593.