A Software Tool for Teaching Reading Based on Text-to-Speech Letter-to-Phoneme Rules

Marian J. Macchi

Dan Kahn

E-Speech Corporation
Princeton, NJ 08540

`mjm@espeech.com`

`dk@espeech.com`

Abstract

Native speakers of English who are good readers can "sound out" words or names from printed text, even if they have never seen them before, although they may not be conscious of the strategies they use. No tools are available today that can convey that knowledge to learners, showing them the rules that apply in English text. We have adapted the letter-to-phoneme component of a text-to-speech synthesizer to a web-based software system that can teach word decoding to non-native speakers of English, English-speaking children, and adult learners.

1 Introduction

Learning to read a language like English involves learning many different operations, including phonemic awareness, word recognition, fluency, verbal comprehension, and expression. The research in this project focuses on the pronunciation aspect of reading from the printed page: understanding how letters, or graphemes, in words are related to sounds, or phonemes.

Most people recognize that the relationship between English orthography and phonetic representation is complex and somewhat arbitrary. Although there is significant evidence that phonological information plays an important role in word reading (Kayner, Foorman, Perfetti, Pesetsky, and Seidenberg, 2001), the precise role of "phonics rules" that would allow a learner to "sound out" a printed word has been debated by educators as well as by cognitive psychologists, and many versions of phonics rules have been discussed by educators.

A classic paper by Clymer (1963) argued that most of the phonics generalizations taught in elementary school are not valid most of the time. Clymer found that for many of the rules, there were so many exceptions that the rule had little utility as a generalization for teaching learners to sound out a word of English. However, the Clymer results do not necessarily mean that phonic generalizations are not useful to readers. Since Clymer, there have been many papers that have suggested alternate formulations of the letter-to-phoneme rules for teaching reading. For example, a recent study by Johnston (2001) found one reason that Clymer considered phonics rules to be unreliable is because the rules he evaluated were too general. Today, there is no consensus on a set of rules, nor does there exist any complete, explicit rule system that "decodes" any word or proper name of English for learners.

E-Speech's letter-to-phoneme (LTP) software, developed over many years for text-to-speech and speech recognition applications, uses proprietary rules to produce pronunciations for any input text. We have adapted the LTP software into a prototype web-based, interactive online system that teaches word pronunciation by explicitly presenting rules for those words/names pronounced according to regular rules and by showing exceptions to the rules. The system allows students to view families of words that obey any given rule and to view words with the same letter patterns that obey different rules.

Our intent is to develop a system that can provide phonics training for beginning readers, either children or adults who are native speakers of En-

glish, as well as for nonnative speakers of English and language-disabled learners. We envision the system either as part of an interactive dictionary or general language-teaching package or stand-alone as an instructional tool for teaching word pronunciation.

A major challenge is to identify rules that are useful for learners and to present them effectively. We have begun to test our prototype system with nonnative speakers of English who were studying English as a second or foreign language. Our preliminary results indicate that the software was (1) useful in improving nonnative speakers' pronunciation of English words; (2) effective at teaching both "basic" pronunciation rules, such as those commonly taught in phonics programs, and some novel, proprietary pronunciation rules.

2   Software Design

The *Word Pronunciation* tool allows a student to enter *any* word or name – whether it is in any dictionary or not - and see our set of rules that account for its pronunciation.   The screen capture below shows the output for the word "photograph". The student can also hear the word pronounced, either normally or syllable-by-syllable.[1]
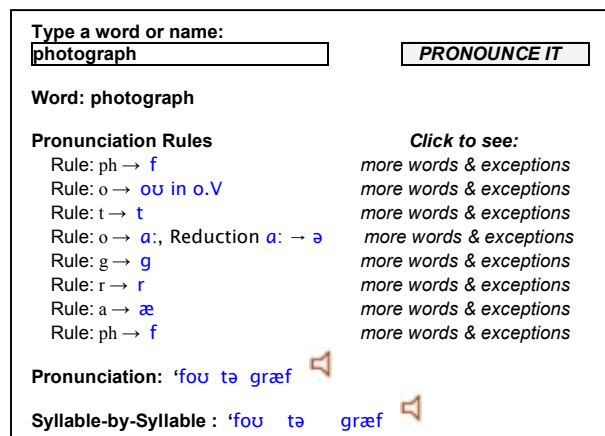
Figure 1 : Word Pronunciation tool display

In addition, a student can click on any rule and see other words obeying the same rule as well as words that are exceptions to the rule.[2]

The *Letter Pattern* tool allows a student to enter a letter or sequence of letters (ie, a letter pattern) and see the rules that apply to that pattern and exceptions to those rules. For example, a student confused by the fact that "how" and "snow" don't rhyme can enter the letter pattern "ow" and view the various generalizations (rules) that determine the pronunciation of this letter string in different contexts, as well as words that don't follow these generalizations (exceptions). The software underlying this tool allows the user to tailor the output to his needs of the moment. For example, one can choose 1-syllable versus multisyllabic words as targets for the rules, how many sample words to output by default, and how big a vocabulary from which to draw words. A simple example of the operation of this tool is illustrated in Figure 2**.**
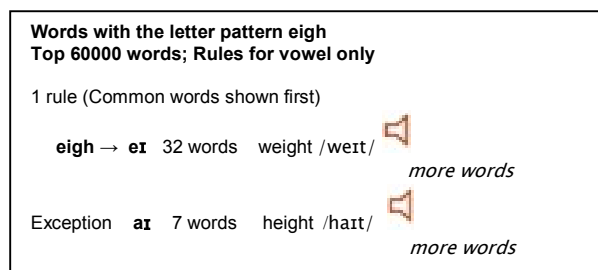
Figure 2: Letter Pattern tool display

While we developed the Letter Pattern tool for general use by learners, we used its underlying search engine in exercises designed to diagnose and teach pronunciation rules.

We implemented a framework for a self-paced set of exercises that allows the user to work alone to diagnose his pronunciation-rule weaknesses and learn the rules necessary to correct his errors. We used this framework to assess the effectiveness of our rules and system for learners.

In the typical exercise interaction, the user sees a sequence of words and must choose the correct pronunciation for each. He indicates his choice of

---

[1] The system uses the International Phonetic Alphabet (IPA) to represent phonetic transcriptions, because most of our target population, adult foreign-born learners of English as a second language, were familiar with this alphabet, since it is used in many English learners' dictionaries.

[2] In this prototype, we worked on presenting the segmental rules, that is, rules for pronouncing phonemes. Although our letter-to-phoneme software assigns lexical stress (to indicate which syllable bears primary stress in polysyllabic words), the stress assignment algorithm is quite complicated. Although the algorithm is accurate, it is too complicated for a human to apply for learning. We also ignored the rules for morphological decomposition, such as for analyzing "walking' into "walk" plus "ing" or "snowman" into "snow" plus "man".

pronunciation by clicking on one of several options, represented in the International Phonetic Alphabet (IPA) or by clicking on a speaker-symbol, so that he can hear the options spoken. An example of a test item would be the nonsense word "doke". If the user chooses an incorrect pronunciation for this word, he is told the correct pronunciation, as well as the relevant rule, which in this case is that an 'o' followed by a single consonant followed by a final 'e' is pronounced /oʊ/. The user can choose to see actual examples of the rule in action ("smoke", "home", etc.) and other rules involving 'o' (eg., the default pronunciation /ɑ:/, as in "hot").

In some cases, the exercises tested real English words, and in others, "nonsense" words (words that do not exist in English but are possible as words, because they have letter sequences occurring in English words). Only through knowing the general rules of English pronunciation can a student correctly predict the pronunciation of words he has never seen before. Figure 3 shows the exercise for knowledge of the letter "a" in the nonsense word "jate."
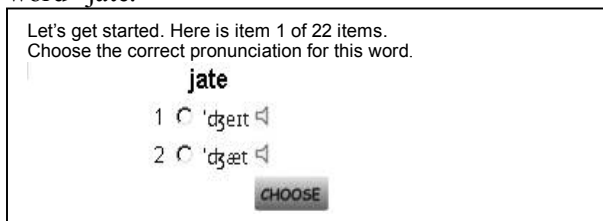

Figure 3. Exercise example

A wrong answer would cause the screen in Figure 4 to appear, in an attempt to teach the student the rule he apparently hadn't mastered.
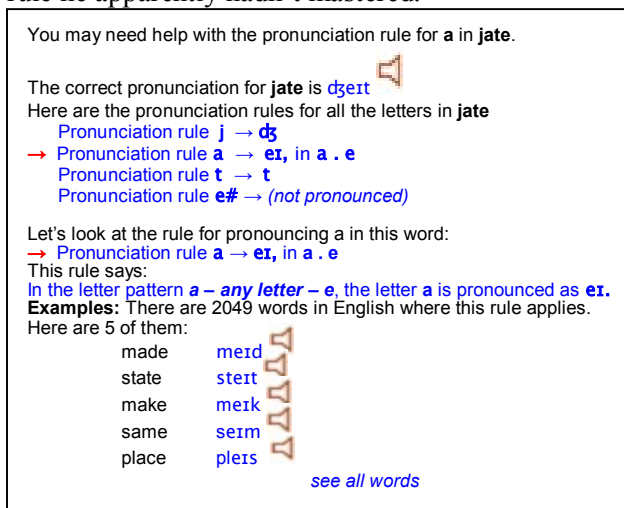

Figure 4. Exercise feedback lesson example

This "lesson" screen highlights the relevant pronunciation rule in the word. Because subjects told us that our pronunciation rule syntax, derived from our LTP rules, was "too mathematical" and hard to understand, the screen also displays an English language explanation for each rule (e.g., "In the letter pattern *a – any letter – e*, the letter **a** is pronounced as /eɪ/"). In our prototype, we developed a simple text-generation algorithm to translate from our "mathematical" rule syntax ("a → /eɪ/, in a.e") into normal English for the rules that we tested in our evaluation. Going forward, however, we will need to produce the explanations via a more sophisticated algorithm or simply hand-prepare explanations for the rules.

A subject can click on the "See all words" link to see more English words in which that rule applies. After the "lesson" the learner is given the opportunity to try again, in order to reinforce the correct pronunciation.

The design of the prototype incorporates several features that are important to its extension to a full learning system. First, the set of exercises is table-driven, so that is relatively easy to add a new set of exercises. This feature is important since a complete system will need a large number of exercises. Second, the system is designed so that the corpus of words that serve as examples of the rules can be changed easily. This feature is important since different user groups (e.g., adult nonnative speakers, children, speakers with reading disabilities) may require different kinds of words as examples.

3 Experimental Results

In addition to developing lexical resource tools, we conducted an experiment to determine (1) if our software could be useful in teaching nonnative speakers of English how to pronounce English words, and if so, (2) if both commonly-taught pronunciation rules and pronunciation rules that are idiosyncratic to the E-Speech letter-to-phoneme system can effectively be taught.

We considered testing the lexical resource tools directly by giving students lists of words and instructing them to use the tools to learn the pronunciation rules for the words. However, we felt that a more efficient way of testing our software would be to develop a set of exercises to diagnose and teach various pronunciation rules and then to test how effectively students learned from the exercises. We developed the design of the exercises

based on informal comments and results of pretests with more than 40 nonnative speakers of English.

3.1 Experimental Design

We sought to improve nonnative speakers' word pronunciation competence, aiming toward giving them the competence of native speakers of English. Therefore, we included both native and nonnative speakers as subjects. 10 nonnative speakers of English and 7 native speakers of English successfully completed the final set of exercises. Six non-native subjects were undergraduates or graduate students at Montclair State University who had been assigned to an English as a Second Language course based on their performance on an English language test administered by the university. The other four were nonnative speakers of English in Brazil, Bolivia, and Germany. Native languages of the subjects were German, Portuguese, Korean, Spanish, Polish, Bangla (Bangladesh), and Urhobo (Nigeria). The native English-speaking subjects were high school or college students who grew up in New Jersey.[3]

The subjects were assigned logins to the system and were instructed to complete a series of exercises, each of which would present different English pronunciation rules. A subject logged in to the system with a web-browser over the internet, saw a printed word and a set of possible pronunciations for the word (as described above). The student was instructed to listen to the set of choices and to choose the pronunciation that he thought was correct. Subjects were told that each exercise would consist of two parts. The first part of each exercise would identify the pronunciation rules with which a subject might need help and then teach the rule; the second part of the exercise would determine whether teaching the pronunciation rules was effective. In the teaching part of the exercise, each rule was presented several times, as it applied to different words. Subjects were allowed to repeat the first part of each exercise as many times as they wished, until they felt comfortable about proceeding to the test part of the exercise.

---

[3] Nonnative subjects were told that we had developed software for teaching word pronunciation and we needed nonnative speakers to try the software and see if it helped them to improve their word pronunciation. English-speaking subjects were told that we had developed software for teaching word pronunciation to nonnative subjects, and that we needed to compare the students' performance with native speakers' performance.

Our software logged the students' choice for each word in each part of each exercise and scored it as correct (1) or incorrect (0). We computed the percentage of correct choices, which we call the word pronunciation score. We also logged the number of times a student practiced with the diagnosis/lesson portion of each exercise, and the amount of time a student spent with each item.

The basic exercises were:

Basic: 1-syllable nonsense words representing "basic" rules, rules that are extremely common in English words. These are productive rules (English speakers apply them in nonsense words), and rules capturing these generalizations are commonly taught in phonics programs. Specifically the exercise teaches:

- **a** is pronounced /eɪ/ in the letter sequence **a - any letter - e** , as in **make**
- **a** is pronounced /æ/ by default, as in **cat** and analogous rules for the letters **e, i, o, u.**

The other exercises taught and tested rules from the LTP system, using English words rather than nonsense words as the material. These were:

LTP1: the basic rules for the letter **a**, plus the trisyllabic laxing rule (which we call the **3-syllable rule**), which causes underlying long vowels and diphthongs to shorten to a lax vowel in antepenultimate syllables:

- **a** is pronounced /eɪ/, in **a - any letter - e**, as in **make**
- **a** is pronounced /æ/ in **a – any letter - e** when **a** is 3 syllables from the end of a word (the "3-syllable rule"), as in **tragedy**
- **a** is pronounced /æ/,by default, as in **cat**

LTP2: rules for **a** before the letter **l**:
- **a** is pronounced /ɔ:/ in **all** at the end of the word, as in **ball**
- **a** is pronounced /ɔ:/, in **alt**, as in **salt**
- **a** is pronounced /eɪ/, in **a - any letter - e**, as in **sale** and **make**
- **a** is pronounced /æ/ by default, as in **pal** and **cat**

LTP3: rules for the letter **a** when it is followed by the letter **r**:
- **a** is pronounced /ɔ:/ in **war**
- **a** is pronounced /æ/ in **arr** followed by any vowel , as in **carry**
- **a** is pronounced /ɑ:/ in **ar** at the end of the word, as in **car,** and in **ar** followed by any consonant, as in **part**
- **a** is pronounced /eɪ/, in **a - any letter - e** , as in **care**

LTP4: rules for the letter **a** when it is preceded by the *phoneme* /**w**/:
- **a** is pronounced /ɔ:/ in **war**

- **a** is pronounced /ɑ:/ in /**w**/**a**, as in **watch** and **quality …** *except …*
- **a** is pronounced /æ/ in /**w**/**a** before the phonemes /**k/, /g/, /m/, /ŋ/,** as in **wag, swam, quack, swang**
- **a** is pronounced /eɪ/ in **a - any letter - e** , as in **wake**

Since each of these exercises included only several rules, the final exercise, ***LTP-all***, recapped the other exercises, in order to assess how well students could integrate all the rules.

   LTP-all: an integrated exercise*: **all** the rules for the letter **a** that were presented in the previous exercises, plus the rule

- **a** → /eɪ/ in **aste**, as in **paste**

   We chose this particular set of LTP rules because they would allow us to compare the "basic" rules, common to many phonics programs, and rules in our LTP system that are not taught in phonics programs.   All the non-basic rules in our experiment were pronunciation rules for the letter "a" and were chosen because they applied to many English words and represented a variety of formal types of rules.   For example, some had relatively simple contexts (e.g., "alt"), and some had complicated contexts. For most of the rules, the context was specified in terms of the surrounding *letters* (for example, the letter a when followed by any letter and the letter e).   For one rule, the context was specified in terms of the surrounding *phonemes*.   This latter type of rule is complicated because it requires that a learner first identify the phonemes for the letters surrounding the target letter "a".

   In the exercise on "basic rules", we used nonsense words to teach and test pronunciation rules. Our reasoning was that the strongest test of whether a student knows the rules is to test his pronunciation of nonsense words, since the only way he could possibly know how to pronounce a nonsense word is by applying the rules. Further, we felt there was strong evidence that the "basic rules" are productive in English.   That is, native speakers of English know these rules and apply them in novel words and nonsense words. For example, English speakers pronounce the "a" in nonsense words with "a" – consonant – "e" at the end of a word, (e.g., "pake", "glape", "nade") as /eɪ/. Consequently, we felt that teaching and testing with nonsense words would help give nonnative speakers the same competence as native speakers.

   However, for the other exercises we used real English words.   Our reasoning was that the nonbasic rules, although they apply to classes of Eng-

lish words, may not be productive in English. That is, native speakers of English might not apply the rules to nonsense words, even though the rule governs a class of existing English words. For example, in English, "oo" is most commonly /u:/ (as in "coo" and "cool"), but when the "oo" is followed by a final "k", the vowel is almost always pronounced /ʊ/ (e.g., "took", "book", "cook", "brook", "crook", "snook", though there are a few exceptions: "kook", "spook").   The question thus arises whether native speakers of English, who obviously know how to pronounce these words, have internalized the "ook rule" and apply it in novel words. Native speakers do not always pronounce novel "ook" words with /ʊ/; instead, they sometimes use /u:/ in nonsense words, like "mook", "dook", "vook" (see Treiman et. al., 2003). Of course, the fact that a rule is not productive does not mean that it is not useful for teaching students how to pronounce words; clearly it would be useful for students to know that "ook" is usually pronounced /ʊk/.   However, since we wanted to compare nonnatives' performance to natives' performance, and we were primarily concerned with teaching nonnative speakers how to pronounce real English words, we chose to teach and test real, as opposed to nonsense, words.[4]

   We did include one test of non-basic-rules that used nonsense words, anticipating that native speakers might perform differently from nonnatives, if the nonnatives, who had been explicitly taught pronunciation rules, applied them to nonsense words, even if natives did not apply them productively in nonsense words.

3.2 Experimental Results

   We present our results informally, without statistical analysis of significance, primarily because we have to date collected data a relatively small number of subjects. Consequently, we interpret our results as preliminary.

---

[4] We attempted to choose words that have relatively low frequency-of-occurrence, to minimize the chances that a nonnative speaker would simply know the word.
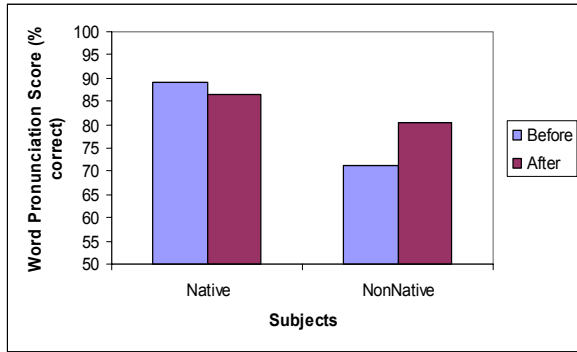
Figure 5. Overall Word Pronunciation Scores

Figure 5 shows word pronunciation scores averaged over all subjects and exercises, tabulated as "before" (word pronunciation scores before subjects were offered any lessons) and "after" (word pronunciation scores from the test parts of the exercises, after the lessons). As would be expected, native speakers had higher word pronunciation scores than nonnative speakers. Further, nonnative speakers had higher word pronunciation scores after completing the lessons than they did before the lessons, although, overall, they did not achieve native speakers' level of word pronunciation. Thus, our data suggests that, overall, nonnative speakers were able to learn aspects of word pronunciation from our system.
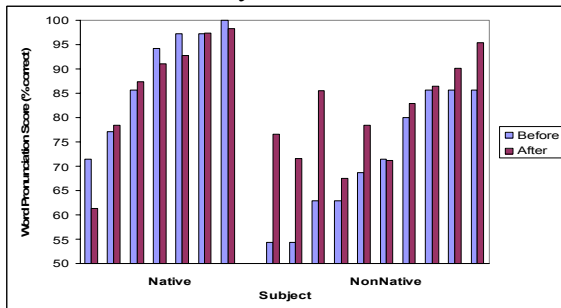

Figure 6. Word Pronunciation Scores by Subject

Figure 6 indicates that there was wide variability among the subjects. Some nonnative subjects' scores increased much more than others', and several subjects' scores did not increase or increased only slightly. Nonnative subjects with higher "before" scores, in general, did not increase as much as the nonnatives with low "before" scores, probably because their scores were high to start with.
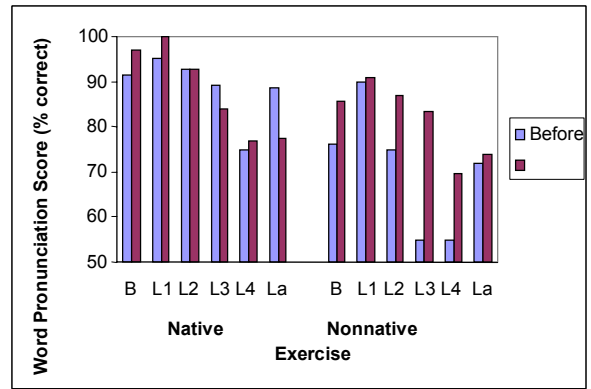

Figure 7. Word Pronunciation Scores by Exercise (B=*Basic*, L1=*LTP1*, L2=*LTP2*, L3=*LTP3*, L4=*LTP4*, La=*LTP-all*)

Figure 7 presents the same data, collapsed across subjects, for the different exercises, which represented different sets of pronunciation rules. We wanted to know whether some exercises proved more learnable than others. In general, as shown at the right side of the Figure 7, for nonnative speakers, for each exercise, the word pronunciation scores were higher after the lessons than before, although the effects were greater for some exercises than others. For native speakers, in contrast, there were no systematic differences in the before versus after scores. However, overall, scores were higher for some exercises than for other exercises even for native speakers. Examination of the native speakers' "incorrect" responses suggests that dialectal issues may have caused some native speakers to choose responses that we did not anticipate. For example, for the word "waffle", some native subjects chose the pronunciation /wɔːfəl/, although we had assumed that the pronunciation in these subjects' dialect was /wɑːfəl/.

Figures 8 and 9 suggest that some rules were useful to nonnative subjects. For example, the "basic" rules, in general, were effective; the nonsense words tested after the lessons elicited higher scores than those tested before any lessons. Of the other, letter-to-phoneme-based rules, the "war" rule and the 3-syllable rule seemed to be effective (the "before" bar for "war" is not displayed in Figure 9, because the before score was extremely low, only 20%).
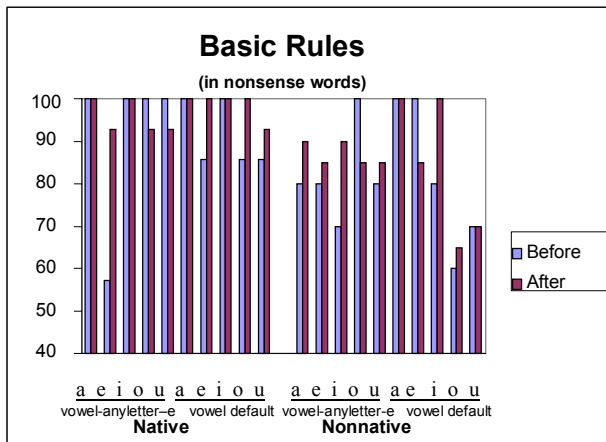
Basic Rules
(in nonsense words)

Figure 8. Word Pronunciation Scores for Basic Rules
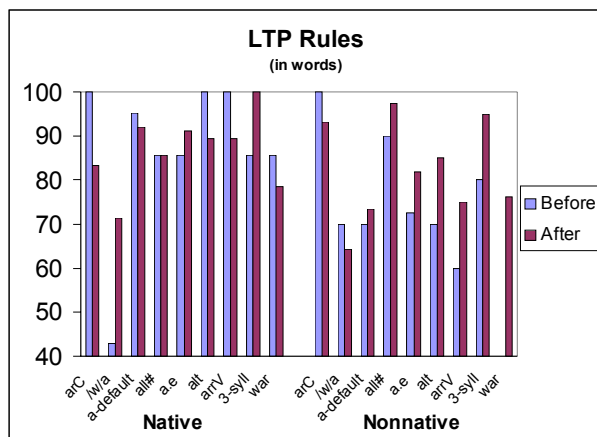
LTP Rules
(in words)

Figure 9. Word Pronunciation Scores for LTP Rules

A complicated rule, the /w/a rule (i.e., the rule that the letter "a" after the *phoneme* /w/ is pronounced /ɑ:/), appeared not to be useful to nonnative subjects. We found no evidence for the effect of teaching for another rule, the "aste" rule, because all nonnative subjects knew the pronunciation of the "aste" words before the lessons. However, there were differences between subjects. One source of this difference is probably due to differences in the nonnative subjects' pre-existing English knowledge; that is some subjects knew some word pronunciations in advance of the lessons. Consequently, for some rules we obtained data for only a few subjects.

As discussed above, since the items in the exercises testing our letter-to-phoneme rules were English words (even though there were not common words) how do we know that subjects' perform-

ance was due to our lessons; perhaps the subjects simply knew the words' pronunciation before participating in our experiments? How well do nonnative speakers apply the rules they learned to words that we can be sure they have never seen before? One of our exercises, **LTP-all** included a section that contained only nonsense words (e.g., "laterous", "plar", "swarg", "falt").
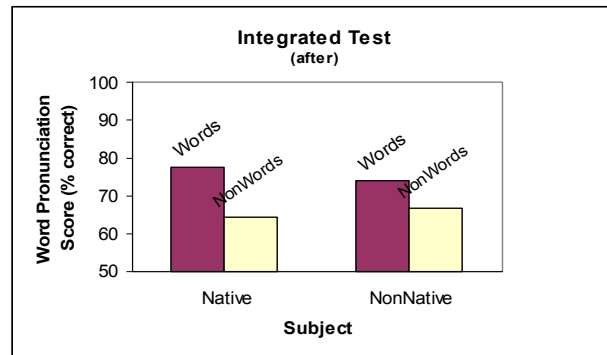
Integrated Test
(after)

Figure 10. Word Pronunciation Scores for nonwords versus words in the Integrated Rules Test **LTP-all** after lessons

Figure 10 presents the results of the nonsense word portion of **LTP-all**: pronunciation scores for real English words versus nonsense words for nonnative subjects and the analogous scores for native speakers, collapsed across subjects. Although there were between-subject differences, on average, both sets of subjects had lower scores for nonsense words than for words. If subjects based their scores for all test items – words as well as nonsense words – entirely on the word pronunciation rules that we included in our exercises, then we would expect their scores to be the same for words and nonsense words. Since words have an empirically correct pronunciation (they are given in a dictionary, for example), native speakers may be relying on a stored phonological representation for the word items. For nonsense words, however, the subjects must rely on rules or other principles. If subjects used rules or principles for pronouncing nonsense words different from the ones we expected, then the nonsense word scores would be lower than those for words. However, the pronunciation scores for the nonsense words for the nonnative subjects were *higher* than that for natives. This fact suggests that the nonnatives were, in fact, applying the pronunciation rules they had learned in our lessons to the nonwords.

In summary, our preliminary results indicate that: (1) our software was useful in teaching nonnative speakers of English how to pronounce English words; (2) both "basic" pronunciation rules and some novel, proprietary pronunciation rules were useful for teaching word pronunciation.

4 Future Research

The major directions for our future research are to select and reformulate LTP rules that are useful for teaching and then to obtain more user data for nonnative speakers and for native-English-speaking children learning to read.

First, we intend to produce a complete learner's rule system for English, based on the entire set of text-to-speech LTP rules. Since our text-to-speech system contains roughly 800 LTP rules, which we believe is too large a number for learners, a significant task is to reduce the number of rules. We intend to focus on rules that apply to large numbers of words and remove rules that apply to few words, relegating the words to which they apply to the exceptions dictionary. We also intend to attempt to collapse rules that apply in similar contexts.

Second, we need to determine whether learners can, in general, understand a pronunciation system that requires rule ordering. In our system, some rules are labeled as "default" rules, meaning that they apply if no other rules apply. Consequently, learners must know *all* the rules in order to know when to apply the default rule. If the number of rules is too large, learners may need a system in which each rule is independently unambiguous.

Third, we need to recruit more nonnative speakers of English who are good candidates for improving their word pronunciation skills. Some of our subjects in this experiment had relatively high word pronunciation scores before exposure to our lessons, so that observing any effect of our lessons was inherently limited. Therefore, we would like to recruit more subjects with lower word pronunciation ability, in order to get a better picture of the effectiveness of our system. Can we predict which students will have high pronunciation scores and which will have low scores based on a student's report of his or her experience in English
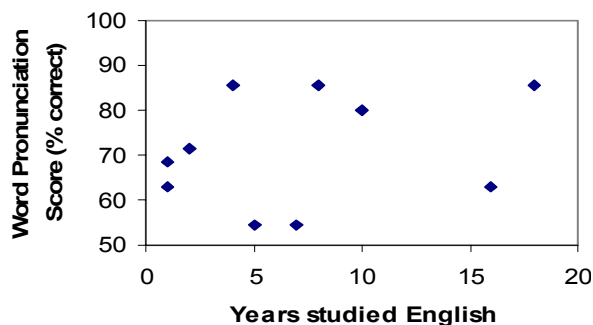

Figure 11. Relation between Word Pronunciation Score and Years Studied English

Figure 11 presents the average word pronunciation scores for each subject at the beginning of the exercises, before doing lessons. As shown, the reported length of time a student had studied English was not correlated with word pronunciation scores. Consequently, we cannot depend on a student's reported length of time studying English as predictive of his word pronunciation abilities. Instead, we will need to screen prospective subjects via pretesting.

Finally, although we have developed our system for nonnative speakers, we would like to test our system with native-English-speaking children learning to read. However, it is likely that the user interface and corpus of exemplar words will need to be different for the child population.

References

Clymer, T. (1963). The utility of phonic generalizations in the primary grades. *The Reading Teacher,* 50, 182-187.

Johnston, F. P. (2001). The utility of phonic generalizations: Let's take another look at Clymer's conclusions, *The Reading Teacher*, 55, 132-143

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., and Seidenberg, M. S. 2001. How psychological science informs the teaching of reading. *Psychological science in the Public Interest*, vol. 2. 2: 31-94.

Treiman, R, Kessler, B, and Bick, S. 2003. Influence of consonantal context on the pronunciation of vowels: a comparison of human readers and computational models. *Cognition* 88: 49-78.