

# The Fluid Reading Primer: Animated Decoding Support for Emergent Readers

Polle T. Zellweger and Jock D. Mackinlay  
Xerox Palo Alto Research Center, California and University of Aarhus, Denmark  
{zellweger, mackinlay}@daimi.au.dk

**Abstract:** We have developed a prototype application called the Fluid Reading Primer to help emergent readers with the process of decoding written words into their spoken forms. When a reader requests help for a particular word, the Fluid Primer uses interactive animation to break the word apart to visually indicate its constituent sounds. The Fluid Primer can also play back audio versions of each segmented constituent sound, as well as the blended complete word. We envision that our method could be incorporated into future e-books to provide scaffolding on demand for readers of arbitrary texts. Moreover, we believe that the animations shown by our tool provide a valuable visualization of the decoding steps that will help readers learn to perform the same steps more quickly and confidently on their own.

## Introduction

Although reading remains a crucial skill in the Information Age, learning to read is a difficult task. A carefully controlled study of adults in the United States by the National Center for Educational Statistics found that 22% could not answer simple questions about short newspaper articles (McGuinness 97, page 9).

Computers have been applied to a variety of aspects of teaching reading. Computer programs and games can help children develop reading readiness skills, such as replacing initial consonants to form new words (e.g., Reader Rabbit (The Learning Company 94)), read aloud to them while highlighting the words or phrases being spoken (e.g., Living Books (Brøderbund 92)), and provide pronunciations of unknown words on demand (also Living Books). Some programs reformat text to enhance readability (e.g., Proportional Reading, which displays one word at a time for a length of time that is proportional to the length of the word, with optional synchronized speech (Adams 98)). Other programs reformat text for comprehension (e.g., Live Ink, which analyzes sentence structure and displays semantic phrases (Walker 00)). Computers have been used to lengthen and enhance constituent sounds within words to help disfluent readers improve their audio processing for reading (Merzenich 96). Researchers have also applied speech recognition to listen to children read and to highlight incorrectly spoken words (Mostow 94). Eyetrackers have recently been used to judge from children's eye movements when to provide spoken assistance with a printed word (Sibert 00). A good overview of other electronic support for literacy can be found in (Topping 97).

However, most previous efforts have ignored a crucial and difficult piece of the reading process – namely, the actual steps involved in decoding a written word into its spoken form. Mastering the decoding process enables a reader to read any word, even unfamiliar or nonsense words.

We have developed a prototype application called the Fluid Reading Primer to help readers with the process of decoding written words into their spoken forms. When a reader requests help for a particular word, the Fluid Primer uses interactive animation to break the word apart to visually indicate its constituent sounds. The Fluid Primer can play back audio versions of each segmented constituent sound, as well as the blended complete word.

Because the Fluid Reading Primer can be used with arbitrary text, it should motivate early readers by allowing them to select material of interest to them. Furthermore, we believe that the animations involved in our tool provide a valuable visualization of the decoding steps that will help readers learn to perform the same steps more quickly and confidently on their own. Finally, since the Fluid Primer focuses on the decoding part of the reading task, it can be combined with other computer-based techniques for supporting readers.

Wise et al provided some decoding support in their computer-assisted remedial reading program (Wise 00). However, their support was given at a higher level (onset + rimes, such as *r/ing*, *st/ing*, rather than single sounds), and they used a much simpler visual presentation (white text on alternating green and blue backgrounds, with no animation). They found no benefits of their segmented form over whole-word support. We believe that our method's sound-level segmentation and careful animation will yield benefits, but verification awaits future empirical study.

The Fluid Reading Primer is part of a larger research project called Fluid Documents, which is exploring the use of interactive animation of typography to show additional information in computer-based documents. A user study indicated that animated annotations do not prevent readers from attending to the reading task (Zellweger 00). However, some subjects preferred static text to animated text. These results suggest the need for careful animation design. The Fluid Reading Primer extends this prior work by focusing on an educational task and incorporating audio into the interactive animation.

## Reading

Diane, Carmen, and Geoffrey McGuinness have recently developed a systematic and relatively simple way to conceptualize and teach the decoding process using conventional paper-based books and exercises (McGuinness 97, McGuinness 98). They argue convincingly that it is important to teach these steps explicitly because other current methods, such as phonics and whole language, fail many students. They have experienced broad success with their method with both reading remediation and new readers. This section summarizes the issues and their method.

### Reading English

We learn to hear and speak on our own, but we must be taught to read and write. Writing uses symbols to show sounds. We reverse this process to read – that is, readers decode symbols into sounds.

The English language, in particular, poses a challenging set of problems for readers. [Note: We focus primarily on children learning to read English as their first language. However, the basic concepts behind our tool are also applicable to learning to read other languages, to older individuals or to learning to read English as a second language.]

- Unfortunately, although English has 43 sounds, it has only 26 letters, so some sounds are represented by groups of letters.  
*Ex:* The letters “ch” show a sound /ch/ (as in “chair”) unrelated to the sounds shown by “c” or “h”
- To make matters worse, some letters and letter groups can represent more than one sound.  
*Ex:* The letter “o” shows the sound /o/ in “ox” and the sound /o-e/ in “old”  
*Ex:* The letters “ow” show the sound /o-e/ in “own” and the sound /ow/ in “town”
- Moreover, there can be multiple ways to show a single sound.  
*Ex:* The sound /a-e/ can be shown using 9 different combinations of letters: “a\_e”, “ai”, “ay”, “ea”, “ey”, “eigh”, “a”, “ei”, and “aigh” (the letter groups are shown in the order of their frequency of use in English words)

Recent research in reading suggests that a better way to teach reading is to start with the sounds, which students know already, and teach the corresponding letters or letter groups. This reverses the traditional phonics approach of letters to sounds.

### The decoding process

An emergent reader must learn three steps to decode words into sounds: visual segmentation, sound assignment, and sound blending.

*Step 1. Visual segmentation.* The reader must divide a word into groups of letters such that each group shows a single sound. A difficult word like “though” may have several possibilities that must be considered, including:

t h o u g h  
th ou g h  
th ough

*Step 2. Sound assignment.* The reader must assign a legal sound (again, typically choosing among several possibilities) to each letter group.

*Step 3. Sound blending.* The reader must remember each sound and sequence them in the correct left-to-right order to form the spoken word. No letter group’s sound may be omitted or reordered during this step, nor may any other sounds be inserted.

For a difficult word, the entire process may require several applications of these steps to try alternative segmentations and assignments. The process ends successfully when the resulting spoken form matches a known word that makes sense in the given context.

### Teaching the decoding process

The McGuinnesses teach the decoding process by starting with a Basic Code, in which each of the 43 sounds in English is associated with a specific letter or a pair of letters. After the emergent reader has learned the three decoding steps described above, the Advanced Code is taught as “spelling alternatives” for those sounds with more than one spelling, and “code overlaps” for those letter groups that show more than one sound. The code developed by the McGuinnesses uses every letter in the word in a left-to-right fashion to show the final spoken form, which simplifies the logic of the decoding process for the emergent reader. For example, the letter “k” in the word “know” is not thought of as silent. Rather, the letter group “kn” is another way to show the sound /n/. The so-called “silent e” requires only a bit more explanation: a few letter groups can be written in two separate parts. For example, the letter “a” and the letter “e” in the word “ate” work together as a single letter group “a\_e” to show the sound /a-e/.

Once the Advanced Code is mastered, the McGuinnesses teach two additional steps for decoding multi-syllable words: chunking and emphasis assignment. In the chunking step, which takes place before the steps described above, the reader must segment the word into chunks, or auditory syllables. Chunks are related to conventional syllables, but

they obey auditory rules rather than visual ones – for example, they do not break between two doubled letters; the doubled letters go to one side or the other of the break. In the emphasis assignment step, which takes place afterwards, the reader must choose which chunks should have strong or weak emphasis applied to them.

The McGuinnesses have also developed a home schooling curriculum that includes a few stories in which each word is printed using special typography to represent the results of the visual segmentation step, as shown below. Emergent readers who are trying to master the Advanced Code find these prepared stories much easier to read than the unmodified versions. This printed form, which they call the Phono-Graphix code, uses a minimal representation to avoid confusing early readers with new and unfamiliar symbols. Studies have shown that introducing special symbols can create a greater barrier to reading conventional text later (McGuinness 97, page 179).

Ex: McGuinness Phono-Graphix code: t e a c h            a t e

## Fluid Reading Primer and Animated Decoding Support

The Fluid Reading Primer provides animated online support for the steps in the decoding process. Our contribution is to use the computer a) to provide a powerful new dimension to the method described above and b) to broaden its applicability from a few specially prepared texts to potentially arbitrary texts.

The Fluid Reading Primer initially shows simple static text formatted in the usual way, which encourages readers to decode words on their own. When students need help decoding a word, they click on it with the mouse. The first click triggers an animation that demonstrates the visual segmentation step: the characters of the word move slightly together or apart to form the correct letter groups (Figure 1). The animation proceeds in a left-to-right order so as to emphasize the correct processing of groups into sounds, since sounds are inherently one-dimensional (time) and must be ordered properly, while text is inherently two-dimensional (x-y). This segmented form encourages the students to proceed with the sound assignment and sound blending steps on their own. If they need more help, they can click on the word again to fire a second multimedia animation that demonstrates the sound assignment step: again proceeding from left to right, letter groups are individually highlighted as the corresponding sound is played (Figure 2). Finally, a third click demonstrates the sound blending step: all of the letter groups are highlighted simultaneously while playing back an audio version of the complete word (Figure 3). Discontiguous letter groups use a special linked highlight (Figure 4).

- |  |   |
|--|---|
| <p>a) Then Snow White awoke and saw seven little dwarves.</p> <p>b) Then Snow White awoke and saw seven little dwarves.</p> <p>c) Then S now White awoke and saw seven little dwarves.</p> <p>d) Then S n ow White awoke and saw seven little dwarves.</p> <p>e) Then S n ow White awoke and saw seven little dwarves.</p> | <p>a) Then S n ow White awoke and saw seven little dwarves.</p> <p>b) Then S n ow White awoke and saw seven little dwarves.</p> <p>c) Then S n ow White awoke and saw seven little dwarves.</p> <p>d) Then S n ow White awoke and saw seven little dwarves.</p> |
|--|---|

**Figure 1:** Animation of the visual segmentation step for the word “Snow.” (a) mouse click, (b) word moves right, (c) (d) (e) letter groups move one after the other. As each letter group moves, its color changes from black to red.

**Figure 2:** Multimedia animation of the sound assignment step for the word “Snow.” (a) second mouse click on the word, (b) (c) (d) filled rectangles highlight each letter group separately, synchronized with audio of the corresponding sound. There is a slight pause between each letter group.

Then S n ow White awoke and saw seven little dwarves.

**Figure 3:** Multimedia animation of the sound blending step. The word is spoken while filled rectangles highlight the visual segmentation.

Then Snow Wh i t e awoke and saw seven little dwarves.

**Figure 4:** Highlighting a discontiguous letter group during the sound assignment step.

We envision that the Fluid Primer will ultimately have adjustable settings corresponding to a reader's current ability. An early reader may need slower animations than a more experienced one. A reader who is just beginning to read multi-syllable words would see a more complete animation, beginning with chunking, and proceeding through visual segmentation, sound assignment and sound blending for each chunk, followed by emphasis assignment and the spoken version of the entire word. A more experienced reader may generally require only the chunking and emphasis assignment steps, although full decoding assistance should remain available. The final minimal level of assistance would simply play back an entire spoken word on demand.

### Issues in effective animated user interface design

The user interface of the Fluid Reading Primer has been designed carefully to promote learning the decoding process. Although the basic concepts involved in animating the steps of the decoding process are relatively simple, subtle issues are involved in developing an effective animated design. This section discusses some of these issues.

The Fluid Reading Primer necessarily accentuates the decoding process in order to demonstrate and teach it. However, we use a careful and parsimonious design for displaying the text and the animations to promote easy transfer to printed books. Although it would be possible to create elaborate graphics or animations to entertain children, we have deliberately resisted this temptation. Recall that some experimental studies suggest that introducing special typographical symbols can either confuse readers or make them dependent on these representations (McGuinness 97 p 179). Elaborate graphics or animations are likely to distract readers in a similar way.

*Step 1: Visual segmentation.* We use a single distinguished color for the selected word throughout the three decoding steps (red is the default, but the color can be changed if desired). Because this color differs from the background color of the remaining text, the word remains a visual unit: its letter groups neither visually join adjacent words nor become separate "words" on their own. However, color does not distinguish between separate letter groups.

We use only whitespace to separate letter groups in the visual segmentation step. This choice differs from the McGuinness Phono-Graphix code, which also emboldens and underlines various characters, for two reasons. First, since the Fluid Primer uses a distinguished color and it only shows the segmented form of one word at a time, the spaces are more distinctive. Second, the animation in the visual segmentation and sound assignment steps further accentuate the distinction between letter groups. Following our design principle of parsimony, the closer we can remain to the printed form without compromising clarity, the better.

The discontinuous "vowel+e" letter groups pose special problems. They must indicate their two-part but related nature in the visual segmentation step. We use a light shaded background to join the two parts of these letter groups even in the visual segmentation step, where other letter groups have no background shading. This shaded background darkens in the sound assignment step (Figure 4).

The animation of the visual segmentation step provides a smooth transition between the conventional form of a word and its segmented form. The smoothness of the animation, which shows several intermediate positions of each letter, demonstrates that no letters appear or disappear in the transition to the segmented form; the letters simply move together or apart. Thus a reader is shown that *every* letter in a word becomes part of some letter group that shows a single sound in the following sound assignment step. This removes the concept of "silent" letters, which can lead to inaccurate reading strategies.

This animation is also intended to encourage readers to notice each letter group and hence to realize that it represents an actual sound that is part of this word, because studies show that disfluent readers may fail to notice letter groups (McGuinness 97). Each letter group changes color as a unit and then all of its letters move left at once. The distinguished color and the motion call attention to it.

Finally, the visual segmentation animation is designed to demonstrate the left-to-right ordering of the decoding process, because disfluent readers may also perceive letters out of order. After experimenting with a variety of movement strategies, we determined that a good way to emphasize the left-to-right ordering is to move each letter group exactly once, to the left, into its final position. We can see the importance of this as follows: if an earlier letter group moves *during* the positioning of later letter groups, it tends to distract attention from the desired focus on the later letter group. Similarly, if earlier letter groups move *after* the positioning of later letter groups, it breaks the left-to-right flow of the process. Thus the letters in the first letter group turn red and move smoothly into their final position, followed by the letters in the second letter group, and so on.

Ideally, the segmented word should still be centered between the surrounding words. This can be a bit tricky to accomplish while maintaining the emphasis on left-to-right ordering. We combine the following two strategies to avoid reader confusion: First, the initial layout algorithm places the words on the screen to allow room for each word to expand into its visually segmented form. Second, we borrow a trick from cartoon animation, which anticipates movement in one direction by a pre-movement in the opposite direction. Here we move the entire word a bit to the right and compress it slightly as feedback for the reader's first click (requesting the visual segmentation), to allow more space for moving left during the remainder of the visual segmentation step (Figure 1b).

*Step 2: Sound Assignment.* In the sound assignment step, we use filled rectangles behind the letters to create the impression that a separate visual layer is showing the sound assignments. These filled rectangles must neither be too dark (interferes with readability) nor too light (interferes with the perception of synchronized sound and letter group).

The animation of the sound assignment step again emphasizes left-to-right processing. Starting from the left, each corresponding sound is played back independently. The sound playback is synchronized with the appearance of a filled rectangle behind the corresponding letter group, thus demonstrating that this sound is a possible alternative for this letter group. Slight pauses between playing the sounds keep them audibly separate.

*Step 3: Sound Blending.* All rectangles are shown while the word is spoken aloud. A smooth animation is used to remove the filled rectangles and close the letter groups after it has been read.

## Fluid Primer Infrastructure

Several underlying dictionaries, tables, and files support the operation of the Fluid Primer:

- *A dictionary of words.* Each entry includes:
  - a. its segmentation into letter groups. Care must be taken here with the discontinuous e, which must appear in its own segment, but participates in a single sound in the sound assignment step
  - b. its assignment of which sound corresponds to each letter group
  - c. its segmentation into chunks
  - d. emphasis assignment for chunks
  - e. a pointer to an audio pronunciation of the entire word

Sample dictionary entries:

Word	Visual segmentation	Sound assignment	Chunking	Emphasis assignment	Audio file
snow	s n ow	s/ n/ o-e/			snow.au
white	wh i- t -e	w/ i-e/ t/			white.au
dwarves	d w ar v es	d/ w/ or/ v/ z/			dwarves.au
pillow	+ p i ll * ow	p/ i/ l/ * o-e/	pill * ow	+ pill * ow	pillow.au

This dictionary contains information that cannot be derived easily from standard dictionaries, so it must be created anew. Note that in general, it must also explicitly contain inflected forms of all words: top, tops, topped, etc. However, once created, a single dictionary can be consulted to help readers with arbitrary texts. If desired, regional pronunciations can be provided for selected words in a separate user dictionary. This feature is particularly useful for names.

Note that some words have different pronunciations for different word senses, such as “read” or “lead”. Simple versions of this situation, in which the different word senses form different parts of speech, could be distinguished by automatically parsing the containing sentence. If the different word senses form the same part of speech, such as “primer” (introductory book) and “primer” (undercoat of paint), the Fluid Primer could offer a choice to the teacher or reader.

- *Audio files pronouncing all 43 English sounds.* This small set of short audio files requires careful speech, preferably by a trained speaker, to isolate the sounds of individual English phonemes. High-quality recording and audio processing are likewise required to make the sounds as clear and distinct as possible. For example, aspirated sounds such as /p/ and /b/ may need to be accentuated, while their plosive portion may need to be softened.
- *A table mapping each sound to the letter groups that can show it.* (for interactive exercises; see below)
- *A table mapping each letter groups to the sounds that it can show.* (for interactive exercises; see below)

## Status and Future Work

Our prototype Fluid Reading Primer is implemented in Java JDK 1.1, which provides minimal support for animation and audio playback. Our animations are implemented in custom code using Java’s thread facilities. They are carefully tuned to synchronize with the playback speed of our audio files on a 200MHZ Pentium II. We plan to reimplement the prototype in JDK 1.2, which has better audio support. We expect to need better synchronization technology as we extend the prototype to support a larger range of texts on arbitrary computers.

We have not yet conducted an empirical evaluation of the Fluid Primer. However, we have shown it to several dozen parents and elementary school teachers in California and Denmark, and their reaction has been very favorable. We also got an encouraging reaction from two emergent readers who were being taught the McGuinnesses's home schooling curriculum. They quickly grasped that they should click on words to get help, and they did not find the animations confusing. We are currently exploring opportunities to evaluate the Fluid Primer with more early readers.

The Fluid Primer can be extended to allow emergent readers to perform the decoding steps on their own via interactive exercises. For example, in the visual segmentation step, readers could drag the mouse across several letters to collect them into a single letter group. Alternatively, they could choose from lists of legal English letter groups. In the sound assignment step, they might choose among visual representations of the legal sound alternatives for a letter group. Audio feedback for each sound alternative would be given on request.

## Discussion

The World-Wide Web is making it possible to access the entire literature of the world quickly and easily. However, the promise of this technology will be reduced if the retrieved literature cannot be read. Decoding words into their sounds is a key skill that will empower a child to use computer-based documents. In this paper, we have described how a computer-based document can also be used to teach an emergent reader this decoding process. Smooth animations demonstrate the three decoding steps: visual segmentation, sound assignment, and sound blending. The animations are carefully designed to illustrate the decoding process with a minimum of graphics and animation. Each step encourages the students to complete the decoding process on their own. Although a systematic evaluation of this technology awaits the development of a more robust prototype, we expect from our previous evaluation of Fluid Documents that these carefully designed animations are not likely to be disruptive to the emergent reader (Zellweger 00). The preliminary observations of our two sample emergent readers support this conjecture.

## REFERENCES

- 1.JF Adams. (1998). Proportional Reading. <http://www.proportionalreading.com/>
- Brøderbund. (1992). Living Books. The Learning Company <http://www.learningcompanyschool.com/>
- The Learning Company. (1994). Reader Rabbit. <http://www.readerrabbit.com/>
- C McGuinness and G McGuinness. (1998). *Reading Reflex: The Foolproof Phono-Graphix Method for Teaching Your Child to Read*. Simon & Schuster. See also <http://www.readingfoundation.com/phono.html>
- D McGuinness. (1997). *Why Our Children Can't Read and What We Can Do About It: A Scientific Revolution in Reading*. Simon & Schuster.
- MM Merzenich, WM Jenkins, P Johnston, C Schreiner, SL Miller, P Tallal. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science* **271**:77-81.
- 1.J Mostow, S Roth, AG Hauptmann, M Kane. (1994). A prototype reading coach that listens. *Proceedings of the AAAI'94 Conference on Artificial Intelligence*. See also <http://www.cs.cmu.edu/~listen/>
- J Sibert, M Gokturk, R Lavine. (2000). The Reading Assistant: eye gaze triggered auditory prompting for reading remediation. *Proceedings of the ACM UIST'00 Conference on User Interface Software and Technology*, 101-107.
- KJ Topping. (1997). Electronic literacy in school and home: a look into the future. *Reading Online*, May 1997.
- R Walker. (2000). Attribute-varied, multi-dimensional text presentation: Impact on reading comprehension and on-line reading eyestrain. Presentation at U of Minnesota Vision Science Colloquium, March 13, 2000. [http://vision.psych.umn.edu/www/Vision\\_Colloquium/Walker.html](http://vision.psych.umn.edu/www/Vision_Colloquium/Walker.html). See also <http://www.liveink.com/>
- B Wise, J Ring, R Olson. (2000). Individual differences in gains from computer-assisted remedial reading. *Journal of Experimental Child Psychology* **77**, 197-235.
- P Zellweger, S Regli, J Mackinlay, B Chang. (2000). The impact of Fluid Documents on reading and browsing: An observational study. *Proceedings of the ACM CHI'00 Conference on Computer-Human Interaction*, 249-256. See also <http://www.parc.xerox.com/istl/projects/fluid/>

## Acknowledgements

We thank Bay-Wei Chang for his instrumental role in the underlying Fluid Documents implementation.