

A Brief Analysis of The English Alphabet Code

“It isn’t easy being green,” Kermit the Frog used to say. Maybe that’s true – but neither is it easy being a child in England, America, Canada, Australia, or anywhere else in the English-speaking world where you have to learn to read (and spell) the English Alphabet Code.

In most countries around the world, nearly all children master the written code and learn to read and spell with relative ease. In these countries a “poor reader” is defined by reading speed, not by accuracy. In English-speaking countries, poor readers are slow as well, but their main problem is accuracy, not speed. This difference is a function of the difficulty of mastering the immensely complex written code for the sounds of the English language.

This will come as a shock to most people. But the facts speak for themselves.

The Nature of Codes

A code is an artificial set of signs or marks that represent (stand for) something ‘real’. That something may be *quantity*, represented by symbols called “numbers.” Or it could be *music*, represented by “musical notation”. This code consists of ‘notes’ made up of black and white circles and flags that sit on 5 lines/5 spaces (or 10 - depending on the musical instrument).

Codes for ‘quantity’ and for ‘music’ are “good codes” -- one symbol for each number, and only one way to represent musical pitch, duration, and rhythmic patterns. A good code is unambiguous. This is what is meant by a "transparent" code. Everyone can see how it works.

The logic of codes is very important. To learn a code you must know the difference between the code itself and what it stands for. **Quantities** exist in the real world. The written symbols for those quantities are ‘abstract’ – a set of arbitrary marks or signs designed to represent them. These marks are so arbitrary they must be agreed upon by *everyone*, otherwise mathematics could not exist. No one would ever dream of marking the quantity ‘*three of something*’ with more than one symbol, such as the symbol **3**, and the number symbol **6**, and the number symbol **21**, and expect the system to work!

Yet this is exactly what the English alphabet code does -- it marks the same sound in English with multiple symbols.

Facts that Cannot be Ignored. Here is another clue that the major difficulty in learning to read in English is due to our written code. In many countries, virtually every child learns to read *and spell* accurately in a matter of months. They do this at about age 6 or 7

- depending on when instruction begins. Comparisons between English and Austrian children showed that after one year of reading instruction, Austrian children were reading twice as fast, with 4 times the accuracy, as English children who had been learning to read for 4 years!

There is no “dyslexia” in countries with “good codes,” certainly not as defined by the “inability to decode the written symbols for a language.” Mastery of a writing system during the first year of school for all children is not unique to Austria, or exotic lands in remote places we have never heard of. Good codes can be found in Italy, Spain, Greece, Germany, Austria, Finland, Sweden, Norway, Korea, and many other countries, all of which have nearly a one-to-one correspondence between a sound in the language, and a symbol that represents it.

Think how bad the English people would be at mathematics, if the written code had multiple alternative symbols for the same number, and each symbol could represent more than one quantity.

The Structure of the English Alphabet Code

An alphabet code marks the smallest speech sounds people can hear. These are the individual consonants and vowels, known collectively as “phonemes.” Alphabetic writing systems were adopted in all cases where other (larger) units of speech would not work. For example, the Chinese have only 1200 syllables in their language, and a syllabary writing system works just fine. By contrast, English has over 55,000 legal syllables, and has to be written in an alphabet by default.

The Anglo-Saxons designed the first written code for the English language. It was a nearly perfect code, with one letter for each sound in the language. At the same time, they continued using written and spoken Latin in all academic, legal, and religious settings. Later, England was invaded by the Danes and the Norman French. New words with different spellings based on the same Roman script, added more layers to the code. And so it came to pass that the English alphabet code wandered from its pristine roots to a state of near chaos.

Here are the fundamental issues that any reading and spelling programme must address and deal with. The solution is NOT to hand out a list of ‘sight words,’ and instruct children to “guess” at meaning by looking at the pictures, then ‘hope’ children will teach themselves to read – standard fare in classrooms around the English speaking world. We have been doing this for over 100 years, and it doesn’t work.

The solution is to find the simplest, most logical way to teach our formidable alphabet code. Here is the problem in a nutshell:

1. Not enough letters.

The English writing system is based on the Roman alphabet. There are only 26 letters in

this alphabet for the 40+ sounds in the English language, and 3 of these letters are wasted (redundant). These are the letters c x q.

2. The creation of digraphs: 2 letters = one sound.

To solve the missing letter problem, new symbols should have been created, but they were not. Instead, Anglo-Saxon scholars combined letters in pairs to represent the left-over sounds -- like sh for the sound /sh/ in 'ship.'

This worked pretty well for consonants, but not for vowels. There are 6 vowels in Latin and 5 vowel letters, but there are 19 vowels in modern English, and still the 5 vowel letters.

As a result of all the above, the ultimate problem is this:

3. Multiple spellings for the same sound.

The sound /ee/ has ten spellings. The sound /ae/ has nine, and so forth. The ten spellings for the sound /ee/ are:

be, beat, beet, baby, key, deceive, believe, radio, marine, theme

4. Multiple ways to decode the same spellings: Read these words out loud and listen carefully to the vowel sounds:

sound, soup, touch----- hat, table, any

Note that our spelling system is context dependent - a spelling is entirely determined by the word it sits in. You can't teach a spelling in isolation.

The Solution.

The solution is to design a reading and spelling programme (the two must never be separated) based on the structure of the English Alphabet Code.

1. **Limits of the Code.** Map the limits of the entire spelling code. Find out how many spellings there are for the 40+ sounds in English. Fortunately, this has already been done by linguists, who estimate the total number of spellings to be between 350-400.

2. **Probability Structure.** Find the probability structure for these spellings to determine which spellings need to be taught. A probability structure is the calculation of the number of spellings *used the most* to those *used the least*. This calculation must be based on frequency in print (how often these spellings appear in print). It is really only essential to teach the main spellings, those used most in written text.

For example, it makes no sense to teach a child all the rare spellings like the **pt** spelling for the sound /t/ (**pterodactyl**), *as if this was equally likely to any other spelling*. Anyone can easily learn this rare spelling if and when it is encountered in print.

3. What Needs to be Taught? Decide how many alternative spellings need to be taught on the basis of the above analysis. For younger children, this is approximately 176 spellings

4. What Sequence is Necessary? Break down the information from 2 and 3 above to simple steps, always moving from the most basic (common words/common spellings) to the most complex.

5. Design Lessons to Maximize Learning Speed. Provide a set of lessons with multiple tasks that reinforce all possible sensory and motor systems: *listening* (phoneme analysis), *looking* (discriminate letter shapes/learn spelling patterns, visual tracking), *writing* (kinesthetic movement), and *speaking* (speech-motor system, auditory feedback) to anchor the spelling code in memory as quickly as possible.

Diane McGuinness.