



The **SCIENCE** of **HIGH-FREQUENCY** Words



LINDSEY CONNER, Ed.D.
Literacy Specialist

LAPREA EDUCATION
WHITE PAPER SERIES

EXECUTIVE SUMMARY

High-frequency words are among the first words children encounter in beginning reading and some of the most frequently occurring words in written English. Because these words appear so often in text, helping students recognize them automatically is essential for developing fluent reading (Blevins, 2024). Yet despite decades of research on how the brain learns written words, high-frequency word instruction remains one of the most misunderstood components of early literacy instruction.

For many years, well-intentioned educators believed that these words needed to be memorized as whole units because English spelling was thought to be largely irregular and unpredictable. As a result, instruction often emphasized flashcards, visual memorization, repeated copying, and weekly word lists. Unfortunately, these instructional routines do not reflect how skilled readers permanently store words in memory. Research consistently demonstrates that skilled readers do not recognize words by their overall visual shape or through rote memorization (Ehri, 2004; Dehaene, 2009). Instead, **automatic word recognition develops through orthographic mapping**, the process of connecting a word's pronunciation, spelling, and meaning in long-term memory (Ehri, 2014; Kilpatrick, 2015). This process depends on students attending to the individual letters and sounds within words, even when a spelling contains an unexpected pattern.

Fortunately, the English writing system is far more regular than many educators realize. Approximately half of English words are completely predictable based on common sound-spelling correspondences, while the vast majority of remaining words are regular except for a single unexpected spelling (Hanna et al., 1966). In fact, less than five percent of English words are considered truly irregular (Joshi et al., 2008). Rather than teaching students to memorize entire words, effective instruction helps them analyze which parts of a word are predictable, identify the portion that is unexpected, and connect those spellings to the word's pronunciation and meaning (Ehri, 2014; Blevins, 2024).

At Laprea Education, we believe high-frequency word instruction should reinforce systematic phonics instruction. Rather than treating high-frequency words as exceptions to the alphabetic principle, we teach them as opportunities to strengthen it. Through explicit analysis of sounds, spellings, and meanings combined with cumulative reading, spelling, and writing practice, students develop the orthographic representations necessary for immediate, automatic word recognition (Ehri, 2014; Kilpatrick, 2015). Ultimately, we want all words to become sight words, recognized immediately and automatically from memory. High-frequency word instruction is one opportunity to build the durable word knowledge and word-learning processes that support fluent reading for a lifetime.

“

Ultimately, **we want all words to become sight words**, recognized **immediately** and **automatically** from memory.



A Matter of Terminology

One reason high-frequency word instruction has generated so much confusion is that several important terms are often used interchangeably, despite referring to different concepts. As a result, educators may unintentionally treat all commonly occurring words as words that must be memorized, when research suggests otherwise. Before discussing effective instruction, it is important to clarify these distinctions.

High-Frequency Words

High-frequency words are simply the words that occur most often in written language (Blevins, 2024). Lists such as Dolch and Fry were developed to identify words students encounter frequently in text because recognizing these words automatically contributes to fluent reading.

Importantly, high-frequency describes how often a word appears, not how it should be taught. Many high-frequency words are completely decodable using previously taught phonics patterns (e.g., can, in, had, get, not). Others contain one or more unexpected spellings but remain largely decodable (Ehri, 2014; Moats, 2020).

Sight Words

Perhaps no literacy term is more misunderstood than *sight word*. In everyday conversation, educators often use the term to describe common words that students are expected to memorize. However, this is not how reading researchers define the term.

According to Ehri (2005, 2014), a sight word is **any** word that can be recognized instantly and automatically from memory. A sight word is not defined by its frequency, spelling, or appearance on a word list. Instead, it is defined by the reader. Once a word has been securely stored in long-term memory through orthographic mapping, it becomes a sight word.

This means that **acetylcholine** may be a sight word for a neuroscientist, while **the** may not yet be a sight word for a beginning reader. Importantly, **sight-word reading is the goal of reading instruction**. Every word, whether high-frequency, multisyllabic, or domain-specific, has the potential to become a sight word as readers repeatedly connect its spelling, pronunciation, and meaning through successful reading experiences (Ehri, 2014; Kilpatrick, 2015).

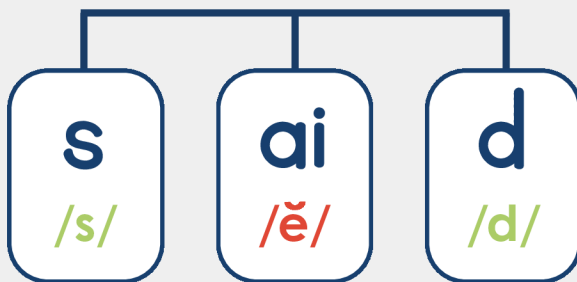
Irregular Words

Many high-frequency words contain spellings that do not follow the most common sound-spelling correspondences. These are often called irregular words. However, research suggests that very few English words are completely irregular. Approximately 50 percent of English words are fully predictable using common sound-spelling correspondences, and another 36 percent are predictable except for one phoneme (Hanna et al., 1966). In fact, less than five percent of English words are considered truly irregular (Joshi et al., 2008).

Recognizing that most so-called irregular words are only **partially** irregular has important instructional implications. Rather than asking students to memorize the entire word, teachers can help students identify the decodable components while drawing attention to the unexpected grapheme(s) (Ehri, 2014; Blevins, 2024).

It is also important to recognize that irregularity depends not only on the word itself, but also on the knowledge of the reader. Some words are **temporarily irregular** because they contain spelling patterns students have not yet been taught. As students expand their knowledge of phoneme-grapheme correspondences, these words become increasingly regular. Other words contain spellings that remain uncommon even for proficient readers and may require explicit attention whenever they are introduced. For instructional purposes, these words can be considered **permanently irregular**. Considering regularity on a continuum rather than as a binary status can support educators as they teach high-frequency words (Moats, 2020).

said



Said is considered a **high-frequency word** with a **permanently irregular** phoneme-grapheme correspondence.

It has three phonemes:

- two have **regular grapheme correspondences** (**s** and **d**)
- one has an **irregular grapheme correspondence** (**ai**)

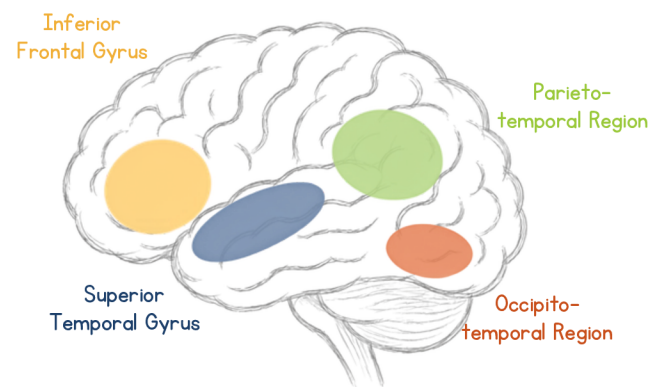
How the Brain Learns Written Words

For many years, researchers believed skilled readers recognized words by their overall visual appearance, similar to how we might recognize a familiar face or business logo. However, research in cognitive science and neuroscience has overwhelmingly rejected this idea, revealing a much different process for how the brain learns and recognizes written words. Researchers assert that **skilled readers do not identify words by their shape or by memorizing them as visual wholes**. Instead, readers recognize words because those words have become permanently stored in memory through a process known as **orthographic mapping** (Ehri, 2014; Kilpatrick, 2015).

Orthographic mapping is the mental process that bonds a word's pronunciation, spelling, and meaning together in long-term memory, allowing the word to be recognized instantly and effortlessly the next time it is encountered (Ehri, 2014). This process depends on readers being able to perceive the individual phonemes within spoken words and connect those sounds to the graphemes that represent them, while also retrieving the meaning of the word. When these connections are secure, the written word can be recognized quickly and automatically.

The Reading Brain

The human brain is not biologically designed for reading in the same way it is designed for spoken language. Reading is a relatively recent human invention, so there is no single area of the brain that evolved specifically for reading. Instead, learning to read requires the brain to build a new network by connecting systems originally developed for other purposes, including spoken language, visual processing, and meaning (Dehaene, 2009).



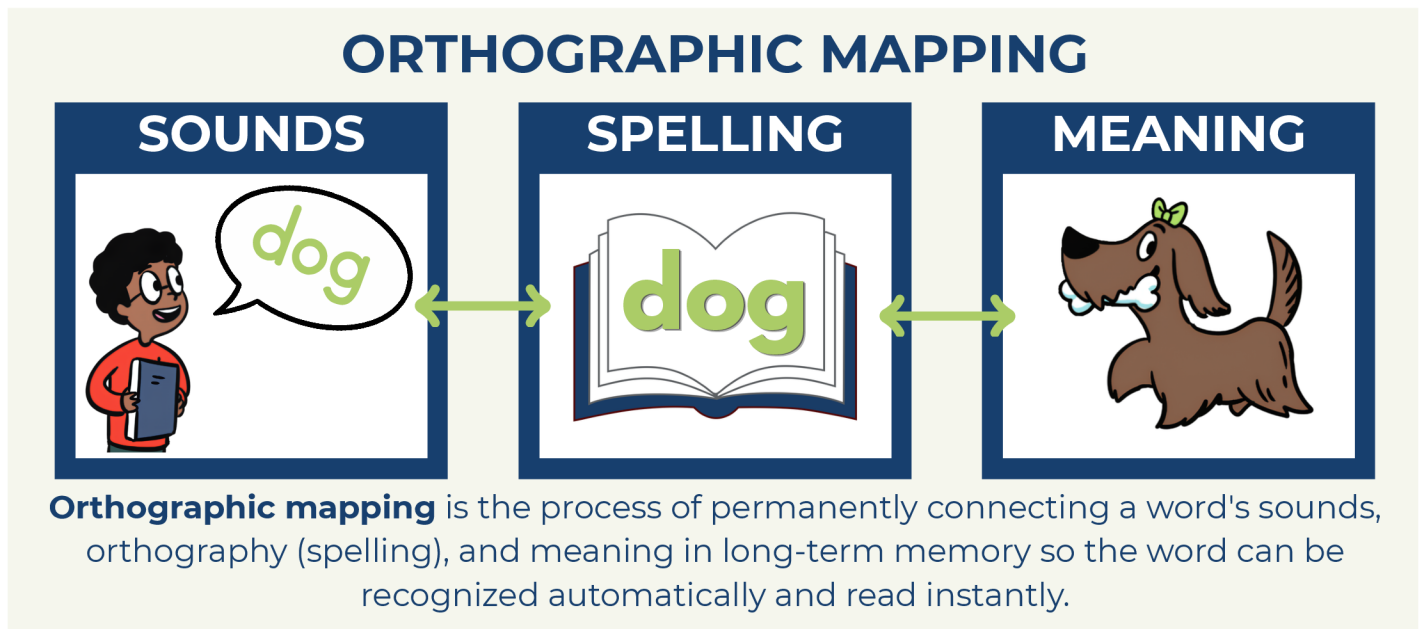
Several regions work together within this reading network. Written words are processed visually in the **occipito-temporal region**, which becomes increasingly specialized for the rapid recognition of letters and letter patterns. The **parieto-temporal region** helps connect those letters and patterns with the speech sounds they represent, while the **inferior frontal gyrus** supports the analysis and articulation of those sounds. The **superior temporal gyrus** supports spoken language processing, helping connect the word's pronunciation with its meaning. With skilled reading, these regions work together efficiently, allowing printed words to be connected with their pronunciations and meanings (Dehaene, 2009; Shaywitz & Shaywitz, 2008).

As students learn to read, these regions become increasingly efficient at connecting letters with speech sounds and language. Through repeated, meaningful practice, these connections become stronger, faster, and more automatic. Over time, this allows skilled readers to recognize familiar written words almost instantly, without needing to consciously sound out each word (Dehaene, 2009; Ehri, 2014).

Orthographic Mapping

Importantly, although the goal is for words to become automatically recognized through orthographic mapping, orthographic mapping itself is a cognitive process that occurs within the brain, **not** a teaching strategy or classroom routine. Educators can, however, facilitate orthographic mapping by providing opportunities for students to analyze the sounds in words, connect those sounds to specific spellings, understand word meanings, and repeatedly retrieve and apply those connections during reading and writing (Ehri, 2014; Miles & Ehri, 2019).

This helps explain why simply showing students a word repeatedly or asking them to copy it multiple times is often insufficient for lasting word learning. **Repeated exposure alone does not ensure that the connections needed for permanent word recognition are established.** Instead, instruction should direct students' attention to the internal structure of a word: the sounds within, how those sounds are represented in print, which spellings are regular or irregular, and the word's meaning. Attending to these features strengthens the orthographic representations that support accurate, automatic word recognition (Ehri, 2014; Dehaene, 2009).



Once a word has been successfully mapped, readers no longer need to consciously decode it letter by letter because the word has become a **sight word**. This does not mean the word was memorized visually, but rather its orthography, phonology, and meaning have been securely connected in long-term memory.

In fact, every successfully decoded word, from high-frequency words to multisyllabic words to content-specific terminology, has the potential to eventually become a sight word.

Why Memorization Falls Short

Understanding orthographic mapping also helps explain why whole-word visual memorization is an inefficient approach to word learning. Human memory is not particularly well suited to retaining arbitrary visual patterns as precise, unconnected forms. In an experimental study of visual memory, Cornelissen and Greenlee (2000) asked participants to remember randomly generated visual block patterns and found that memory for the patterns declined rapidly as the interval between presentations increased, with an estimated half-life of approximately three seconds.

Although the study did not examine word reading specifically, its findings illustrate the limitations of relying on visual memory alone to retain arbitrary patterns.

Written words, however, are certainly not arbitrary visual patterns to skilled readers. Instead, the interconnected representations established through orthographic mapping provide a far more efficient system for storing and automatically recognizing the thousands of written words readers encounter over time (Ehri, 2014).

Perhaps even more importantly, when students are asked to rely on visual memorization to learn words, their attention is diverted from the phoneme-grapheme correspondences that support orthographic mapping. Yoncheva et al. (2015) found that directing learners' attention to grapheme-phoneme relationships during initial learning produced patterns of neural activation more characteristic of those associated with skilled reading than instruction emphasizing whole-word learning. **In other words, how educators ask students to attend to words makes a difference.**

Additionally, words learned primarily through visual memorization may also remain fragile. Without adequate knowledge of phoneme-grapheme correspondences, students may develop less secure representations of written words and have greater difficulty transferring their knowledge to unfamiliar words and new contexts (Ehri, 2005; Miles & Ehri, 2019). **As text demands increase, reliance on visual memorization becomes increasingly insufficient**, often revealing weakness in word reading skills that explicit instruction in phonemic awareness and phonics could have addressed much earlier (Ehri, 2005; Kilpatrick, 2015).



The goal of high-frequency word instruction, therefore, is to help students build increasingly robust orthographic representations, allowing thousands of words to become securely stored in long-term memory and instantly retrievable across a lifetime of reading.

Principles of Effective High-Frequency Word Instruction

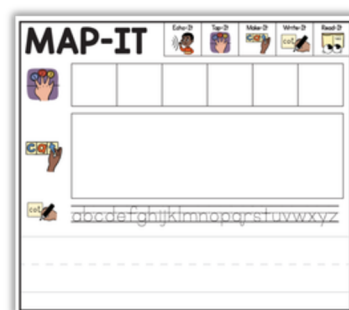
Understanding orthographic mapping fundamentally changes how we think about teaching high-frequency words. If words become permanently stored by connecting their sounds, spellings, and meanings, then instruction should be designed to strengthen those connections rather than encourage visual memorization.

In other words, effective high-frequency word instruction is about creating the conditions under which orthographic mapping can occur.

Decades of reading research point to several characteristics of effective high-frequency word instruction. While instructional routines may vary across contexts, these general principles remain remarkably consistent across the literature.

1 Draw Attention to Sounds Before Spellings

Orthographic mapping depends on accurate connections between phonemes and graphemes. Before students analyze the letters in a word, they should first be able to pronounce the word and identify the individual sounds within it (Ehri, 2014). Once the sounds have been identified, students can connect each phoneme to its corresponding spelling, making note of any portion that is unexpected.



This process directs students' attention to the structure of the word rather than its overall appearance. The goal is not for students to remember what the word looks like, but to understand how its sounds are represented in print.

2 Teach the Unexpected Parts Explicitly

Educators may be surprised to learn that few English words are completely irregular. More often, only one portion of a word differs from the most common spelling pattern (Hanna et al., 1966; Moats, 2020). Effective instruction helps students identify which parts of the word are predictable and which require special attention, sometimes called the “heart part”.

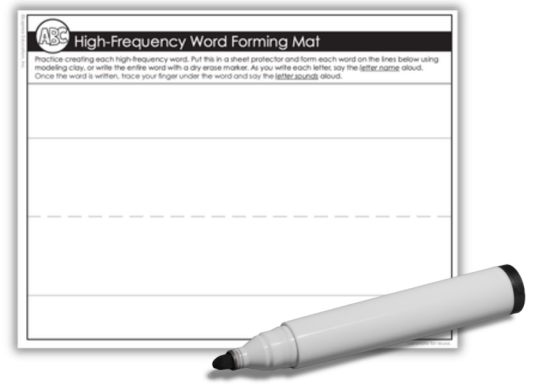
For example, in the word **from**, students can recognize that the initial /f/ and /r/ consonant sounds and the final /m/ sound are represented by expected spellings. The vowel **o**, however, represents an unexpected pronunciation and deserves focused attention. By isolating the irregular portion rather than treating the entire word as an exception, teachers can reinforce students’ phonics knowledge while building orthographic understanding.



3 Pair Decoding and Encoding

Reading a word is only **one** component of learning it. Students develop stronger orthographic representations when instruction intentionally integrates both decoding and encoding. These reciprocal processes provide students with repeated opportunities to connect a word's sounds and spellings, strengthening the phoneme-grapheme connections needed for orthographic mapping (Ehri, 2014; Graham & Hebert, 2011).

Writing the words adds another layer of reinforcement. Research with preliterate children suggests that handwriting experience can support the development of neural systems involved in letter recognition and processing (James & Engelhardt, 2012). When incorporated into word-level instruction, handwriting can therefore complement the decoding and encoding practice used to strengthen orthographic representations.



4 Provide Cumulative Review

Orthographic representations strengthen through successful encounters and retrieval over time. More broadly, research on learning and memory demonstrates that retrieval and distributed practice support stronger long-term retention than repeated exposure or massed practice (Roediger & Karpicke, 2006; Rosenshine, 2012). Applied to high-frequency word instruction, this suggests that a word encountered once during a weekly lesson is unlikely to receive sufficient practice for automatic recognition.

In practice, high-frequency word instruction should follow a planned scope and sequence that is **cumulative** rather than one-and-done. As new words are introduced, previously taught words should continue to be intentionally pulled forward for brief review and application. Students might read and spell them during word work, retrieve them during a quick warm-up or dictation routine, encounter them in decodable or connected text, and use them in their own writing.



Every Word Can Become a Sight Word!

At its core, the science of high-frequency word instruction is about understanding how the brain learns written language. Once educators understand the process of orthographic mapping, many long-held assumptions about "sight words" begin to change. **Words become instantly recognizable because they have been orthographically mapped, not because they have been memorized.**

This distinction fundamentally changes the purpose of high-frequency word instruction. The goal is not simply to help students remember the words on a weekly list. Rather, it is to help students build the permanent connections among sounds, spellings, and meanings that allow words to become part of their growing sight vocabulary. Every successful encounter with a word strengthens these connections, making future recognition faster, more accurate, and increasingly automatic.

Fortunately, this process aligns naturally with explicit, systematic phonics instruction. Rather than treating high-frequency words as exceptions to the alphabetic principle, educators can help students apply their phonics knowledge to identify the predictable portions of words while explicitly teaching the few spellings that are unexpected. When students read, analyze, spell, write, and repeatedly retrieve these words over time, instruction creates the conditions under which orthographic mapping is most likely to occur.

Ultimately, every word students encounter is capable of becoming a sight word! Our responsibility as educators is to provide instruction that helps those words become permanently mapped rather than temporarily memorized. This subtle shift in thinking transforms high-frequency word instruction from an exercise in rote memorization into an opportunity to strengthen the neural networks that support skilled reading.



References

- Blevins, W. (2024). *Differentiating phonics instruction for maximum impact: How to scaffold whole-group instruction so all students can access grade-level content*. Corwin.
- Cornelissen, F. W., & Greenlee, M. W. (2000). Visual memory for random block patterns defined by luminance and color contrast. *Vision Research*, 40(3), 287–299. [https://doi.org/10.1016/S0042-6989\(99\)00137-6](https://doi.org/10.1016/S0042-6989(99)00137-6)
- Dehaene, S. (2009). *Reading in the brain: The new science of how we read*. Viking.
- Ehri, L. C. (2004). Teaching phonemic awareness and phonics: An explanation of the National Reading Panel meta-analyses. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 153–186). Paul H. Brookes Publishing Co.
- Ehri, L. C. (2005). Development of sight word reading: Phases and findings. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 135–154). Blackwell Publishing. <https://doi.org/10.1002/9780470757642.ch8>
- Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, 18(1), 5–21. <https://doi.org/10.1080/10888438.2013.819356>
- Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, 81(4), 710–744. <https://doi.org/10.17763/haer.81.4.t2k0m13756113566>
- Hanna, P. R., Hanna, J. S., Hodges, R. E., & Rudorf, E. H. (1966). *Phoneme-grapheme correspondences as cues to spelling improvement*. U.S. Department of Health, Education, and Welfare, Office of Education.
- James, K. H., & Engelhardt, L. (2012). The effects of handwriting experience on functional brain development in pre-literate children. *Trends in Neuroscience and Education*, 1(1), 32–42. <https://doi.org/10.1016/j.tine.2012.08.001>
- Joshi, R. M., Treiman, R., Carreker, S., & Moats, L. C. (2008). How words cast their spell: Spelling instruction focused on language, not memory, improves reading and writing. *American Educator*, 32(4), 6–16, 42–43.
- Kilpatrick, D. A. (2015). *Essentials of assessing, preventing, and overcoming reading difficulties*. John Wiley & Sons.
- Miles, K. P., & Ehri, L. C. (2019). Orthographic mapping facilitates sight word memory and vocabulary learning. In D. A. Kilpatrick, R. M. Joshi, & R. K. Wagner (Eds.), *Reading development and difficulties: Bridging the gap between research and practice* (pp. 63–82). Springer. https://doi.org/10.1007/978-3-030-26550-2_4
- Moats, L. C. (2020). *Speech to print: Language essentials for teachers* (3rd ed.). Paul H. Brookes Publishing Co.
- Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- Rosenshine, B. (2012). Principles of instruction: Research-based strategies that all teachers should know. *American Educator*, 36(1), 12–19, 39.
- Shaywitz, S. E., & Shaywitz, B. A. (2008). Paying attention to reading: The neurobiology of reading and dyslexia. *Development and Psychopathology*, 20(4), 1329–1349.
- Yoncheva, Y. N., Wise, J., & McCandliss, B. D. (2015). Hemispheric specialization for visual words is shaped by attention to sublexical units during initial learning. *Brain and Language*, 145–146, 23–33. <https://doi.org/10.1016/j.bandl.2015.04.001>