Learning About...

...or “Everything You Wanted To Know About WordCore 144 But Were Afraid To Ask!”

©Russell Thomas Cross June 2007
**Table of Contents**

**Why is it called “WordCore?”**  
1. We store words in our head as the “building blocks” for talking.  
2. Words can be modified by using mental rules.  
3. Although there are hundreds and thousands of words in the English language, a very small number make up the actual number of words used.  
4. Words can be categorized into groups that have similar features.  
5. Word Prediction can reduce keystrokes.  

**Why are keys arranged the way they are?**  
1. Words are clustered by Part of Speech  
2. Many sentences start with a Pronoun and a Verb, or Verb and a Pronoun: *WordCore* is organized to make this easier.  
3. In *WordCore 144*, the top half of the core keys contain the highest frequency words and phrases  

**How does the color-coding work?**  
1. Color coding is common in AAC systems  
2. *WordCore* uses elements of the Fitzgerald key and Goosens’, Crain and Elder.  

**References**  

**Appendix 1: Word Counts**  

**Endnotes**
**Why is it called “WordCore?”**

*WordCore* is based on many features of spoken language. In effect, *WordCore* tries as much as possible to take how language happens in your head and do the same on a device in the world. It’s always a good idea to make devices work in the same way that people think because in the long term, this makes them easier to use. Let’s take a look at some of those “features of spoken language” that are so important.

1. **We store words in our head as the “building blocks” for talking.**

When we talk to each other, we typically talk in sentences. These sentences are made up of strings of words put together in some sort of order. Estimates for the number of words a person knows vary from 10,000 to 250,000! Even children are thought to progress from a 500-word vocabulary at age two through to a potential 14,000 by age six. On average, children appear to learn around 11 words per day from the age of 18 months to six years.

*WordCore uses words based on normal word usage.*

2. **Words can be modified by using mental rules.**

Many words are made up of a base form onto which endings can be added. For example, the word “want” can be changed to “wanting,” “wants,” and “wanted” by adding the endings “-ing,” “-s,” and “-ed.” Similarly, the words “cat” and “dog” can become “cats” and “dogs” by adding on the “-s” ending. This ability to take a WORD and add an ENDING is called morphology and it is a critical component of not just the English language but all languages.

*WordCore uses normal English morphology rules.*

---


\(^f\) Deutscher, G. (2005); Pinker, S. (1999);
3. Although there are hundreds and thousands of words in the English language, a very small number make up the actual number of words used.

Large-scale studies of English have shown that a relatively small number of words are used for most of our conversation\(^8\). This is true of normally developing children\(^b\), and individuals using AAC systems\(^j\).

WordCore uses a core\(^3\) vocabulary based on this wide range of vocabulary studies.

4. Words can be categorized into groups that have similar features.

The words “want,” “walk,” and “think” have something in common. So do “cat,” “leg,” and “tent.” In the same way, “I,” “you,” and “he” have more in common that “in,” “on,” and “under.”

The words in each one of these groups will behave in similar ways. The first group we call Verbs, the second Nouns, the third Pronouns and the fourth Prepositions. As a whole, these groups are referred to as Parts of Speech.

From a very early age, children are able to learn that words can be categorized as parts of speech\(^j\), and even use these categories to attach meaning to words\(^k\). Adults frequently add new words to their mental dictionary and can typically use all the correct forms of the word once they know its grammatical category\(^l\). The rules that describe how parts of speech can vary are well documented\(^m\) and using these rules in any language system is recommended\(^l\).

WordCore uses Parts of Speech as a method of representing and categorizing vocabulary.

5. Word Prediction can reduce keystrokes.

Using a single key to produce a single word is fast and efficient. Spelling letter by letter is very flexible (you can write any word you want) but slower than using single-hit keys

\(^b\) Raban, B. (1987)
or sequences of keys\textsuperscript{a}. For individuals with motor issues, reducing keystrokes is valuable, even if it doesn’t increase the overall rate of communication\textsuperscript{b}.

\textit{WordCore uses word prediction to save keystrokes}\textsuperscript{5}.


**Why are keys arranged the way they are?**

1. **Words are clustered by Part of Speech**
   As already mentioned, words in our head appear to be stored categorically, not alphabetically. *WordCore* reflects this on its overlay.

2. **Many sentences start with a Pronoun and a Verb, or Verb and a Pronoun: *WordCore* is organized to make this easier.**
   The commonest type of sentence that people use is called an *SVO sentence*: Subject + Verb + Object. A frequent Subject is a Pronoun (*I, you, he, she* etc.). The overlay reflects the Subject-Verb construction by having the Pronouns at the top left and the Verbs just to the right. Hence, you go from left to right for a Subject-Verb phrase, or right to left for a Verb-Subject phrase – which is a common form in *questions*.

3. **In *WordCore 144*, the top half of the core keys contain the highest frequency words and phrases**
   If you intend to use *WordCore* in **Key Mode**, the keys available in that mode have been chosen to try to ensure that the highest frequency words are available without the need to use the Scroll Key. For example, in *WordCore 144* the conjunctions *and, but, if, as, because, or and than* are all available in the extreme right-hand column of the core. If you switch to Key Mode, you still have access to *and, but* and *if* because these are the three most frequently used conjunctions. The same happens with all the other grammatical classes.

---

How does the color-coding work?

1. Color coding is common in AAC systems
Color coding is common in both low- and high-tech systems but there is no standard set of colors. The closest to a recognized set is based on an approach introduced by Elizabeth Fitzgerald in the 1920’s for teaching language to individuals with deafness. The original “key” was based on a set of six symbols, not colors, and each stood for a general grammatical class. The color notion was added and this has been used in programs such as Gateway, WordPower, the Ingfield Express, Clicker, and others. Here’s how the key is usually coded:

- People – Yellow
- Nouns - Orange
- Verbs - Green
- Social - Pink
- Descriptions - Blue
- Miscellaneous - White

Contrast this with Goosens’, Crain, and Elder (1992):

- Nouns – Yellow
- Prepositions – Green
- Verbs – Pink
- Descriptions – Blue
- Miscellaneous – Orange

2. WordCore uses elements of the Fitzgerald key and Goosens’, Crain and Elder.

<table>
<thead>
<tr>
<th>POS</th>
<th>COLOR</th>
<th>VARIATION 1</th>
<th>VARIATION 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pronouns</td>
<td>YELLOW</td>
<td>Lighter yellow: Object and Possessive forms</td>
<td>Darker yellow: Base pronoun forms</td>
</tr>
<tr>
<td>Verbs</td>
<td>GREEN</td>
<td>Light green: Helping verbs</td>
<td>Darker green: Common verbs</td>
</tr>
<tr>
<td>Adjectives</td>
<td>LIGHT BLUE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverbs</td>
<td>TURQUOISE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conjunctions</td>
<td>BROWN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepositions</td>
<td>PINK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determiners</td>
<td>ORANGE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questions</td>
<td>TAUPE</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

† Fitzgerald, E. (1954)
† Inman, N. (2001)
§ Blenkhorn, P. and Hawes, P. (2002)
¶ Blenkhorn, P. and Hawes, P. (2002)
The Adjectives and Adverbs have been coded as shades of the same color because both are types of “description” word, with adjectives describing nouns and adverbs describing verbs. The shading difference between the verb types is also based on this notion of trying to mark similarities between generic groups that also have some common features."
References


Appendix 1: Word Counts

<table>
<thead>
<tr>
<th>WordCore 45</th>
<th>Words</th>
<th>Phrases</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td>72</td>
<td>211</td>
<td>283</td>
</tr>
<tr>
<td>Activity</td>
<td>125</td>
<td>91</td>
<td>216</td>
</tr>
<tr>
<td>Pages</td>
<td>345</td>
<td>23</td>
<td>358</td>
</tr>
</tbody>
</table>

Total WordCore 45 vocabulary: 857 words and phrases

<table>
<thead>
<tr>
<th>WordCore 84</th>
<th>Words</th>
<th>Phrases</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td>74</td>
<td>398</td>
<td>472</td>
</tr>
<tr>
<td>Activity</td>
<td>144</td>
<td>70</td>
<td>214</td>
</tr>
<tr>
<td>Pages</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Total WordCore 84 vocabulary: 687 words and phrases

<table>
<thead>
<tr>
<th>WordCore 144</th>
<th>Words</th>
<th>Phrases</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td>135</td>
<td>560</td>
<td>715</td>
</tr>
<tr>
<td>Activity</td>
<td>197</td>
<td>122</td>
<td>319</td>
</tr>
<tr>
<td>Pages</td>
<td>76</td>
<td>33</td>
<td>109</td>
</tr>
</tbody>
</table>

Total WordCore 144 vocabulary: 1143 words and phrases

Note: These figures reflect the pre-stored vocabulary accessed via the Core, Activity Rows and Pages, but the Word Prediction system is fully functional for all other vocabulary items. WordCore provides a core of high frequency words supplemented by the word predictor for lower frequency items. Hence the reason that WordCore is noun poor with the prediction system doing all the hard work.
Endnotes

1 There are some differences between word frequencies in spoken language versus written language. The first person singular subjective personal pronoun (I) and second person personal pronoun (you) are used much more in spoken language than written (see Leech et al., 2001, for specific numbers). However, WordCore addresses these differences by (a) using complete grammatical category sets and (b) favoring spoken frequencies over written if there is a significant difference. An example of “complete grammatical category” is that ALL the Personal Pronouns are included in the core vocabulary, so the issue of spoken/written is no longer important.

2 The notion that kids lean “X words per day” isn’t as clear cut as it sounds. It’s not that on Monday, Frank knows 253 words then on Tuesday he has 264. Paul Bloom (2000) does a good critique of this notion and suggests that there is neither a straight line linear relation nor a “spurt,” but something more like an exponential rate of learning, with kids starting slow and getting progressively faster. As they get older, they become better at identifying words and adding them to a structured categorization system – once you “know” that a word ending in “-ing” is typically a verb, then if you here spoooting, you add spoot to your vocabulary, which in turn can be used to code spooots, spoooty, spoootily and so on.

3 The word core is used in a very specific way for a very specific meaning. It refers to a word that is used frequently across situations and across groups. The reason want is a core word is because it is used often at work, in school, at the dentists, at the movies, going fishing, and so on ad nauseum. The opposite of core is fringe, which refers to a word used infrequently and in limited situations. The word wrench is a fringe word because it is not used often at all and if it is, it is typically in a very prescribed situation. Some folk in AAC have tried to redefine the word core to mean something like “the important words in situation X.” This allows you to say hamburger is a core word “when you’re at a fast food restaurant.” This is misleading and dangerous redefinition of core and serves no useful purpose. In fact, it logically leads to you being able to define ANY word as core, and thus is not a defining characteristic at all!

4 Of course, it is possible for a word to belong to more than one grammatical category. The word light is notoriously fickle in that is can be used as a noun (“Turn on the light”); a verb (“Come on baby light my fire”); and as an adjective (“I’d prefer a light beer”). Within the adjective category, it can also have different meanings, such as light=not heavy; light=not dark; light=low-fat; light=not serious; and so on. In WordCore, a word is assigned to a grammatical class based on the frequency of its use as a grammatical entity. Thus talk is stored as a verb, not a noun (“I have to give a talk tomorrow”) or an adjective (“Did you see that talk show?”).

5 For individuals using switches to access a device, keystroke savings are always a good thing. Even if your rate of speech generation using word prediction is no faster than straight spelling, using fewer keystrokes leads to less fatigue and that is itself is desirable.

6 This is a feature of the PRC ECO communication aid whereby the screen can be split in half horizontally to show a keyboard at the bottom and a computer at the top. This is a different approach from using a “typical” on-screen keyboard, which can in itself interfere with seeing and accessing the rest of the screen. A more general issue is that ALL display-based devices suffer from the problem of having a screen that has to operate both as input AND output. This is what I refer to as “The Masking Problem” because at any given time, the input method will be masking the output, or the output will be masking the input. It is an inherent problem with such devices, but oddly one that is rarely mentioned. Devices with a separate input and output methods don’t have this masking problem.

7 In fact, the original Fitzgerald Key didn’t mention colors; it was concerned with grammatical categories and word order. The use of the “modified” Fitzgerald key seems to have started with Bliss – at which point color codes were added.
If you’re a guy, then of course, *taupe* does not exist in your mental color wheel. Blokes apparently can only perceive 7 colors, along with *black*, *brown*, and *white*. Men can describe anything by using the 7 colors plus the prefixes *it’s a sort of…, light and dark.*

This is a unique feature of *WordCore*. Other systems use basic colors but only *WordCore* using shade variation to capture linguistic commonalities.

Alert word counters will have noticed that WordCore 84 has fewer pre-stored words than WordCore 45. Doesn’t this mean that WordCore 45 is “better?” Truth is, simply counting the number of words in a system isn’t a measure of how “good” it is. The Oxford English Dictionary has every word in the English language stored, but “Green Eggs and Ham” is a much “better” read. Here’s an example of where WordCore 84 is actually more efficient than the 45:

To say the word “family” using WordCore 45 requires three keystrokes; SPELL, “F,” then selecting “family” from the word row on the page. Next time around, the word “family” still requires three keystrokes.

To say the word “family” with WordCore 84 requires four keystrokes; “f,” “a,” “m,” then “family” from the word prediction row. However, next time, you only need two keystrokes because the system learns from your use of the word predictor.

So, the 84 is actually more efficient in that it requires fewer keystrokes to generate the equivalent language to the 45-location version. You will also notice that having words like “this” and “think” as single hits on the 84 overlay also means that keystrokes are saved. So, overall, the 84 version performs better as measured by keystroke efficiency rather than pre-stored word count alone.

To a certain extent, I would argue that the problem of how to handle fringe is much more difficult than the problem of handling core. In fact, core vocabulary is easy to identify and so the trick is to get as much of it as possible in as small a space as possible with as few keystrokes as possible. But fringe is much trickier for four big reasons: (a) It is potentially huge in size, (b) it changes every day, (c) it contains items that often cross category boundaries, and (d) it is incredibly difficult to predict. The huge size is apparent if you simply open up a dictionary. The vast majority of words turn out to be nouns from *aardvark* to *zyrian* and entering all of them into a system is challenging, especially if you’re using a picture-based approach. Secondly, words are not only added to the English language on a daily basis, but current words can also shift their meaning. A trip to [www.urbandictionary.com](http://www.urbandictionary.com) is stunningly instructive on this issue! Thirdly, words like *fork, coat,* and *pear* can all have more than one meaning, so how do you categorize them? Finally, how do you predict which fringe words are going to be needed in any given situation? At best you play the probability game and say, “Given that I am in Burger King®, the probability of my using the word *burger* is likely to be higher than that of *necromancer.*” But is the statistical probability truly significant? I don’t know and I don’t know of any studies that help us here, except for perhaps the Balandin and Iacono 1999 study that effectively supports the notion that *any* fringe vocabulary items can pop up in even a situation as apparently prescribed as “during the lunch break.”