Beyond Black ’n White: Addressing Cultural Bias in Word Gap Studies

Anita Pandey

“You are such a good drawer!” remarked 7-year-old Helena to my 11-year-old who was sketching a family of ducks by the wetland swamp near our home. While she clearly meant “artist,” her word “drawer” (like “painter,” where the suffix “-er” refers to the doer) got her point across. Two decades ago, psychologists Hart and Risley (1995) drew attention to what is today referred to as the (30-32-million) word gap, which has been cited time and time again as evidence of the need for more frequent interactions between parents and young children. Word gap studies also undergird much of the national discourse in favor of quality preschooling and early educational experiences—considered the foundation for lifelong success (Rich, 2013; Colker, 2014). True, children from lower income homes across the U.S. typically hear fewer Standard English (SE) words by the time they enter school, and this is believed to account for their relatively limited knowledge of the language of schooling or academic English. As is to be expected, when instruction and assessment are not culturally inclusive, dual language learners (DLLs) and speakers of stigmatized dialects of English, such as Black English Vernacular (BEV) tend to fall behind (Author, 2000). Yet, exactly how this finding is surprising is unclear. Indeed, it would appear axiomatic that children from low-income homes—few of whom employ SE in their everyday interactions—would hear, understand, and use fewer words from this dialect. This piece identifies cultural bias in the word gap debate, and offers recommendations.

The following are some arguments and questions aimed at pinpointing the bias in question:

1. **Children are Language Facilitators and Word Generators**

Parental interactions with children are not the primary catalyst for vocabulary and other forms of language, and certainly not in every single household. Children and young adults are fully capable of facilitating language and literacy. As demonstrated in *The Child Language Teacher* (2010), they are some of the best
facilitators for their peers-- and adults-- yet their language-instructional skills are rarely recognized and maximized. While research shows that interaction fosters language development, to assume that vocabulary growth is contingent on interactions stimulated primarily or solely by parents and/or adult caregivers is a misconception. How else, for instance, would we account for creoles—compact, child-created languages? Quality of engagement also matters. Language acquisition theory and practice suggest that interaction is an essential ingredient, alongside comprehensible input (CI—essentially understandable language), and timely and clear feedback (Author, 2012). However, it should be noted that the nature of feedback and the manner in which it is presented are culturally variable. To assume that only direct and positive feedback engender learning is questionable. In addition to being behavioristic in orientation, and suggesting that children merely mimic adults, it fails to account for language children produce that they did not get from adults.

Far more important than adults interacting frequently with children then is ensuring that children get to interact often. Gurgling and other forms of language used by infants and toddlers, however rudimentary, signal their readiness for interaction, and must be encouraged, since it is precisely by negotiating meaning that children grow their language and content—including their knowledge of specific sounds, word parts, words, phrases, clauses, and culturally variable discourse (Author, 2012). Any and all forms of interactions involving children should, therefore, be fostered and all language (including media-based) that children access must be counted in assessments of vocabulary and language growth—or lack thereof. These include child-child, child-technology (e.g., Wi-programs), child-teen; child-parent; and child-third-party (family and non-familial) exchanges. If not, resultant vocabulary assessments are incomplete and adult-centric, and children’s creative vocabularies, including add-on meanings they assign to conventional words, as well as words they continually manufacture (and that parents adore and recollect with nostalgia) are overlooked.

2. Déjà vu + Quantity isn’t Quality

Recall the Ebonics controversy, which was blown out of proportion? [The linguistic term “genetically related” in the initial Ebonics Resolution, for instance, was totally misconstrued] So too has the so-called “word gap” (Sparks, 2015).
Have you ever wondered how researchers arrived at a 30- to 32-million word difference? Were words used by the parent(s) alone counted? Were both parents present in every home when the data were collected? How were idioms and phrasal verbs tallied? These are just some of the questions that come to mind.

That dialectal variations are rarely considered in word gap studies is problematic. To contrast the lexicon of a primary speaker of SE with that of a speaker of another dialect is like comparing apples to oranges. The word gap argument is, therefore, moot and, in a word (LOL), biased. For one thing, quantity isn’t quality. Giving someone the eye, for example, communicates so much more in BEV than a single SE word could possibly convey. Indeed, the adage less is more is exemplified in the lexicons of several dialects, particularly pidgins and creoles. In Nigerian and many other varieties of English-based pidgin, for instance, the preposition for is multipurpose; this single word communicates practically all relational meanings that multiple SE prepositions are employed to convey—and context and intonation communicate varied and nuanced meanings.

Does a child really need 30-32 million words to excel and/or to sound erudite, or is word-mongering fashionable? If the latter were the case, why would the revamped SAT delete “obscure words” in favor of “relevant vocabulary” (NPR, June 3) and why would researchers (e.g., Marzano, 2000) observe that 50,000 to 60,000 words are more than adequate? If the average third grader is expected to know some 3,000 SE words, a 32-million-word difference between rich and poor children, while confounding makes little sense. Research suggests that students learn an average of 2,000 words per year. Even if we were to exaggerate and use 3,000 as the magic number, the most a student would learn in 12 to 15 years of schooling (including preschool—and assuming preschoolers mastered some 4000 words) would be 45,000. Many words were very likely counted multiple times in “word gap” studies. We would need to know whether the corpus was analyzed for lexical variety (i.e., whether a few words were repeatedly used in the recording period and counted many times or whether substantially more and a wider array of words were recorded in selected homes). So the question to ask is whether the word gap refers to 30-32 million different words (a difference of 2 million that’s not small by any standards) or whether this figure is merely a contentious extrapolation that must be taken with a grain of salt.
3. Not all Concepts are Captured in Words

Just ask yourself “Are words the sole and/or primary carriers of concepts or ideas?” Arguably, one can be expressive without using words--or multiple words--for that matter. Indeed, we can communicate key concepts through non-lexical means; language consists of more than words. So hearing, saying, and writing more words does not necessarily mean that a child’s concept knowledge is more expansive than that of a child that hears and/or uses fewer labels. We would need to agree on what it means to know a word versus, for instance, a concept. We would also need to research which words are labels for single concepts and which ones encode and represent more than one concept. If we were to research the structure of all 6,000 or so languages still in use in some form, we would find that each has some unique words and concepts, as well as some (like birth, death, love, and marriage) that are shared. We are also likely to find that while some words are labels for single concepts, others are broader in scope (i.e., homonyms and/or homophones). Pidgins/patois and creoles, for instance, are highly compact languages; fewer words and shorter sentences communicate multiple ideas that require a lot more words in SE.

4. Missing the Big Picture?: Words vs. Language

Aside from salutations (Hi), leave-taking devices (Bye, Adios), insults (idiot!), and threats and warnings (Stop! Enough!), we rarely communicate through single words. The order in which words, phrases, and clauses are assembled and, in speech, the intonation employed, collectively communicate meaning. A dependent clause, for instance, signals conditionality (e.g., Unless you . . .). It literally leaves you hanging, and is arguably more threatening than a single word (e.g., “unless”). Non-verbal language (e.g., eye contact, gestures, proxemics) and other forms of oral language that many stigmatized dialects typify are equally important. Indeed, non-lexical forms of language should count, too, yet how best can they be counted?

When we count words alone, we miss the big picture. Why else do some employ the French sentence Je no se quoi in lieu of a single word? The next time our word power fails us, how about impressing our listeners with (Yo) no se que? After all, the Spanish variant (Yo) no se que is similar to and, in fact, closer to the Latin source than the French equivalent, and Spanish is more widely spoken in the U.S.
Or are we convinced that French sounds fancier, much like many of us are fixated with the word gap and insist that more words equal greater knowledge?

We must remember that substantive language isn’t always lexicalized, that is, expressed through single words. When an Arabic speaker encounters a senior moment, rather than referencing a lexical loss, she typically apologizes for her momentary “language” amnesia. In Persian, Arabic, and Urdu, ghazals and shayar (i.e., debate-songs and couplets) actually garner greater praise than single words. Ironically, even SE speakers love using idioms. While not technically “words,” idioms are classified as words, because single-word equivalents are readily available. In a word (pun intended), the word is just one of the many language building blocks we employ to convey meaning.

5. **1 + 1 = More: Word on the Street**

Word has it that children from low-income households are word-strapped. Yet few, as you’ll agree, are at a loss for words. Most enjoy rapping and miming, and ‘talk’ up a storm, and are more vocal and confident when it comes to public speaking than children from professional and/or higher-income homes who, arguably don’t have to try as hard to be heard. The Czech adage *You live a new life for every new language you speak* illustrates the richness and dearth of meanings bi/multilinguals have access to—through language that lacks lexical equivalents in SE. To assume that children and adults from low-income families and communities are linguistically deficient because they utter and/or hear fewer words at home is stereotypic (Strauss, 2013) and unfounded. Let’s talk about words for concepts that are culturally unimportant in SE. Examples include words for distinct family members and relationships (e.g., paternal vs. maternal grandparents, delineated in Chinese and most African and Asian languages). Indeed, languages tend to lexicalize culturally significant concepts. When we consider the colorful and meaningful language that children who are not primary speakers of SE often employ, in addition to their seemingly limited SE vocabulary, then they actually know much more language and arguably more words than monolingual SE users. Many, for instance, know concepts for which English lacks words. Try translating the BEV word *trippin* into SE. You end up with multiple words, and a host of meanings, from annoying, to unconventional, to boastful.
6. Slow as a Turtle?: When Slow is Fast

Not all words convey the same meaning(s) in every language variety. I still remember the loud murmur that swept through my class when my student Samakai shared the meaning of his name. In Vai, a language spoken in Liberia, Samakai means *slow as a turtle*, yet in actuality, *slow* means *fast* in the West African cultural context. There, the turtle, while slow in speed is the fastest animal in intellect. Reversed meanings are not unusual. Another example of such day ’n night lexical differences is the simple word *owl*. While SE speakers rarely go around calling people *owl*(s), the owl is revered for its wisdom. To be associated with the owl is commendable in mainstream America. Not so in Southern Africa and India. In Bemba, Shona, Hindi, Gujarathi, Marathi, and even in Nepali and Urdu, the equivalent (*ulu*) is negative. In these cultures, the owl is considered incapable of distinguishing day from night, and as such, considered foolish. *Ulu* is hardly a compliment; it refers to an *idiot*.

Also, it should be noted that word recall speed isn’t necessarily indicative of lexical knowledge. While mainstream America is visibly individualistic and competition-cum-aggression-oriented, and we tend to equate lexical speed with efficiency and victory, *reaction time* is not a culturally inclusive measure of vocabulary knowledge or, for that matter, of language competence.

Like schools, primarily-SE-using homes tend to prize reading and writing. These are not innate skills but human creations and learned behaviors, unlike listening and speaking (primary literacy). As such, SE words easily confound not just children but many an adult, as well, since they are rarely written the way they sound. Mastering non-phonetic words therefore takes a considerable amount of time, effort, and strategy (see Author, 2012).

7. The Complexity of Simple

It should be noted that lexical complexity is relative and depends on one’s PL. So too are semantics, as illustrated above. To children whose PLs are Latinate, SE “simple words” (primarily Anglo-Saxon) are complex, and vice versa (Author, 2012). Therefore, exposing children to other dialects and languages early on and to synonyms and synonymous ways of communicating meaning (i.e., to more than words) introduces them to structural variety in language, and is highly advisable.
This way, children hear and steadily begin to use a mix of simple and complex words and other forms of language.

8. A Final Word about the Word Gap

Just because a child hears a word does not mean that i) she understands the meaning(s) it communicates, and ii) uses it correctly. So simply saying that middle class children hear and/or use more words doesn’t translate to superior vocabulary and, for that matter, to better concept knowledge and overall communicability. Moreover, equating word power with language competency (and by extension, content knowledge) is an unfounded and biased approach to instruction and assessment. After all, unlike in SE, words are not the primary or most important unit of language in every single variety. Consider for a moment the multiple words and expressions in SE that prioritize just a single unit of language, namely, the word. Anything that’s not well-worded is usually considered poorly written and disorganized in English—that is, in SE. From words like *foreword* to *afterword*, to *the last word*, and let’s not forget, *a word in edgewise*, SE is, *in a word*, wedded to the word. No wonder it boasts the largest vocabulary. Meanwhile, spelling bees thrive, signalling selected students’ academic prowess, and the SE dictionary continues to grow and stump even adults. Mandarin speakers might very well be at a loss for words if you ask them how many words a Chinese-speaking third-grader is expected to know, or to identify word boundaries in Chinese. Words are an arbitrary unit of language, which explains why word-for-word translations are rarely accurate. [Try translating Chinese characters. What do you end up with? Letters, multiple words, or word parts?]

Spanish, for instance, uses a single word *vamonos* where English uses two to three (*Let’s go* or *Let us go*), much like Bengali, Marathi, Hindi and even Urdu (i.e., “chalo,” pronounced in a variety of ways). Some might argue that *vamonos* constitutes two words, or that contractions aren’t single words. What good would counting do us? So ask yourself “Does the word gap really add up”?

I could go on and on, until I’ve used 32 million words, and we’d be drowning in words. Would my piece be more meaningful? “What’s the solution?” you ask. I propose *culturally-inclusive vocabulary instruction*. We have an obligation to correctly guage children’s *concept knowledge* in their *primary language* (PL) and in other varieties they know, so that we can effectively bridge observed *language*
*differences* (not *gaps* per se) in a culturally and linguistically inclusive manner. It’s our responsibility to correctly identify the PL and to use it alongside SE. Such an approach is the most effective. Just as children from low(er) income homes are expected to be familiar with SE terms, so too we must evaluate their educators’ and peers’ knowledge of their PL, and of culturally significant concepts and expressions in their languages. Did I hear you say “Word”!? 

**References**


