Beyond the Scrabble Word List: Making more inclusive word games

Allison Parrish, IndieCade East February 2015

@aparrish

I'm the Digital Creative Writer-in-Residence at Fordham University and a research fellow at NYU's Interactive Telecommunication Program.

The best way to sum up my practice and research: I think about what happens to language when it comes into contact with systems that are usually considered "non-linguistic." Like poetic form, or the Internet, or games.

Games about spelling

"Word game" is an ambiguous term—after all, almost all games can involve words in some way or another. For the purposes of this talk, by "word games" I mean "games about spelling." Scrabble, Boggle, Quiddler: games where the individual units of play are letters, and the goal of the game is to form words.

In this talk, I'm going to focus on Scrabble, though I think the critique applies to most games about spelling. I'm going to show how Scrabble, through its structure, expresses a particular view about what kind of language use should be valued, and why those values are worth challenging. Then I'm going to show a few of my attempt to design word games that express different values.

This talk is more of a polemic/manifesto than anything else—I think I've identified a problem, and I think I have some good ideas about solutions, but I don't have any empirical evidence that my solutions do, indeed, solve the problem. It's all very preliminary. Be forewarned.

An anecdote

4

So to begin I want to tell a story about Scrabble. Once I was in Utah visiting my family over the holidays and we decided to play a board game. Someone suggested scrabble and I was like, okay. This is essentially how the game went:

5

First my mom played "north." A perfectly good play, worth 24 points.

Then my little sister played "fireman" for 26 points. And then it's my turn to play and I look at my rack and my eyes light up and I don't WANT to be an insufferable smart-ass, but what can I do? so I play...

QOPH. And "qi" and "pe" and "hm." It's a pretty good move, perfectly legal Scrabble, and it's worth 66 points... but is it worth the contention and strife caused by playing not one but FOUR weird words in one turn? This move had a deleterious effect on our fun. Everyone thought I was engaging in ostentatious brain-showboating. AND I KIND OF WAS. It wasn't fun for anyone and eventually we decided to play something else. So here's why I don't like scrabble:

8

Scrabble turns otherwise nice people into pedantic a**holes

A nicer way to phrase this would be: competitive Scrabble play requires a lot of arcane knowledge. You have to memorize a lot of words, both tiny and large. So when you're playing with people who haven't memorized all this hermetic vocabulary, it can lead to hurt feelings—and hurt feelings are no fun! It's worth mentioning that other games aren't like this—if I was better than someone in my family at soccer or street fighter, they probably wouldn't think I was being a smart-ass.

Why?

It's easy to chalk these problems up to the culture of Scrabble, or game balance issues, or simply the inability of certain individuals (ahem, me) to keep their smartassery in check for more than thirty seconds at a time. But it's worth digging a little deeper. I think the structure, mechanics and materials of Scrabble reflect larger cultural and societal problems. Allow me to explain!

Unigram frequency

- The "moving parts" of the game are individual letters

- Commonality and value of letters based on (some model of) letter frequency in English

- Letters drawn at random

Scrabble is based on what I'm going to call "unigram frequency"—by which I mean the "moving parts" of the game are individual letters, whose value and commonality are determined by the frequency of those letters in the English language. There's only one Z, but it's worth 10 points. There are a bunch of Es, but they're only worth one point. Letters are put into play by drawing them randomly, one at a time.

```
        Common letters with low value
          Rare letters with high value
    +     Low reward for longer words
    ─────────────────────────────────
          a**hole words like "qoph"
```

This means that the highest value words are words that are densely packed with rare letters.

Words densely packed with rare letters tend to be words that (a) are less commonly known and (b) have inconsistent or unusual spelling.

These are what I'm calling "smart-ass" words.

Good vocabulary and spelling skills are good things (according to Scrabble)

So Scrabble is, through its structure, expressing a value: knowing a lot of arcane words and being skilled at spelling are good things. These are the skills you need in order to be good at Scrabble.

"The dictionary was very much the complete curriculum"

Now, there is some strategy to Scrabble beyond being able to spell arcane words. But if you look online, the most commonly cited advice for improving your scrabble score is to memorize words. I got the quote in this slide from a study ("The Development of Expertise Within a Community of Practice of Scrabble Players")—basically, many expert players agreed that the gateway to higher level play is to memorize arcane words.

Okulicz, E.; Vialle, W.; and Verenikina, I., The Development of Expertise Within a Community of Practice of Scrabble Players, Learning and Socio-cultural Theory: Exploring Modern Vygotskian Perspectives International Workshop 2007, 1(1), 2007.
Available at:http://ro.uow.edu.au/llrg/vol1/iss1/4

Bad feelings

So at this point you're thinking, well okay, yeah. This is obvious: to be good at Scrabble, you have to be "good at" language. What's the big deal? Why does that fact lead to bad feelings?

I think that the issue here is that the kinds of skills involved in Scrabble aren't just valued *by Scrabble*. They're valued by society at large, in a way that can be oppressive. The bad feelings come from the fact that for someone to assert their expertise at a word game carries a different message (I think) than asserting their expertise at, say, soccer or poker.

To explain this, I want to unpack what it means to be "good at" language.

To have a "large vocabulary" and "spelling skills" implies that there is a canonical list of valid words with canonical spellings.

In other words: there is English that is "right" and English that is "wrong."

So first of all, the concepts of "good spelling" and "large vocabulary" presuppose two things: (1) that there is a canonical way of spelling English words; and (2) there is a canonical list of English words.

But this establishes a binary! If there's such a thing as "right" spelling for a word, there also has to be a "wrong" spelling. For there to be a list of words that are "valid," some other words have to be "invalid."

Now, this isn't necessarily problematic on its face! Most games set up binary distinctions between which moves are "valid" and "invalid." For example, some hands in poker are valid, others aren't.

Having your English considered to be the "right" English is a form of privilege.

The problem is this binary reflects and relies upon a harmful binary that exists in the real world. Ideas about what kinds of English are considered "correct" and "incorrect," "valid" and "invalid" are deeply intertwined with ideas about race, class, gender and ability.

What's considered "correct" English is usually the English spoken by neurotypical middle class white men.

What's worse: there's an idea that people who use language differently, who have "low" literacy, are "stupid" or "lazy."

More in-depth discussion of these ideas: http://kimkatrincrosby.squarespace.com/media/2012/4/15/the-africana-language-power-privilege-and-discourse.html

a better Scrabble player than you

I'm a better person than you.

So a good move in Scrabble is saying more than "I'm a better Scrabble player than you." It's saying "I'm a better person than you." I'm more *literate* than you. This is, I think, the source of the bad feelings you get when playing Scrabble.

And I think it's a phenomenon at least somewhat unique to word games: the skill that you use to play is very closely associated with a harmful form of privilege.

Games, by valorizing these skills, are implicitly supporting this harmful form of privilege.

Vocabulary size + "correct" spelling
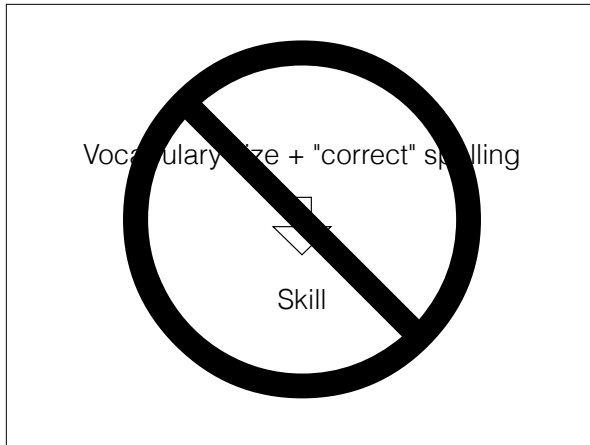
Skill

So the radical idea I'm proposing here is twofold. The first is that having an arcane vocabulary and being skilled at spelling are not necessarily *a priori* "virtuous" or "skillful"; and that game designers, by assuming that "good" players will have these skills, are unnecessarily limiting their audience.

To drive these points home, I want to go into a bit more detail about how both of these things—spelling skills and vocabulary size—aren't worth designing a game around.

Spelling is not real

The first point: spelling is not real. Or rather: spelling is the way that it is because someone decided it should be that way.

behaviour → behavior
saltpetre → saltpeter
analyse → analyze
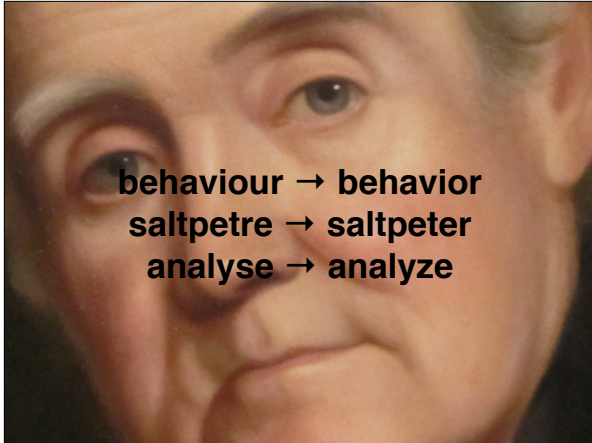
The idea of "standardized" spelling—by which I mean, the idea that a word might have only one valid spelling—is relatively contemporary, and really only dates back to the 18th century. Before that, every individual and community of practice had their own (sometimes very loose) conventions about how words should be spelled.

And even standardized spelling differs from one variety of English to the next! You can see here a very large and uncomfortable picture of Noah Webster, who was instrumental in introducing the distinction between American and British spellings like those seen here.

The important thing to know here is that Webster's reforms were essentially arbitrary. He decided that he thought some words should be spelled a certain way—for reasons he considered to be logical, aesthetic and politically expedient—and then other people decided to go along with it. ("Somebody thought of that, and someone believed it: the Rainbow Connexion.") There's nothing inherently "correct" about his spellings. (Or anyone else's for that matter.)

On spelling standardization: http://www.ruf.rice.edu/~kemmer/Histengl/spelling.html

ghoti

Another point to make about English: English spelling is *incredibly sophisticated* and requires a significant investment of resources to properly master. (The string "ghoti"—pronounced "fish"—is famously used to illustrate the inconsistencies in English spelling.) Tou*gh*, w*o*men, na*ti*on.

Learning how to spell is usually included as part of a formal education. Unfortunately, access to a quality formal education is available only to certain portions of the population.

"[He] might be dyslexic. He might have spent his entire childhood being shamed and belittled by his teachers and classmates because his brain works differently from theirs. He might come from an abusive or otherwise dysfunctional home, where focusing on his English studies comes a distant second to keeping the pieces of his body and spirit together. [...] He might be a refugee from a country whose native language has an entirely different writing system, or no writing system at all. He might be surviving any combination of these circumstances, and more."—*Literacy Privilege: How I Learned to Check Mine Instead of Making Fun of People's Grammar on the Internet*

This is a quote from "Literacy Privilege," an amazing essay that I'm incredibly indebted to for many of the ideas in this talk. Here, we have a list of things that might get in the way of someone being able to spell "properly."

To sum up: spelling is arbitrary, spelling is political, and access to spelling is obstructed along any number of axes.

source: https://paintingthegreyarea.wordpress.com/2012/11/26/literacy-privilege/

Dictionaries are imaginary too

Let's talk about word lists. The point I want to make here is this: you can't make a list of all "valid" words in the English language, and any attempt to do so is going to exclude the language of marginalized groups.

Many members of the audience here today are familiar with "SAT test" words—words that you learn simply for the purposes of proving your intelligence on standardized tests.
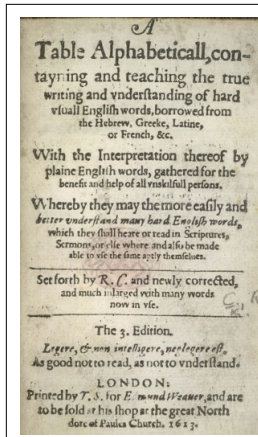
"... English words, gathered for the benefit & helpe of ladies, gentlewomen, or any other vnskilfull persons. Whereby they may the more easilie and better vnderstand many hard English wordes"—*A Table Alphabeticall*, published 1604

It turns out that learning words purely for the purpose of proving how smart you are has a long history! Here's the title page from "A Table Alphabeticall," one of the earliest English dictionaries. It claims to be "helpe"ful for "vnskillfull persons"—basically claiming that you should buy this dictionary so that you can sound smart. So it ever was.

(thanks to http://www.nytimes.com/2010/03/14/magazine/14FOB-onlanguage-t.html?_r=1&)

And so it continues today—"learning more words" is everywhere sold as snake oil to help you improve your lot in life. This is a screenshot of the copy from "Ultimate Vocabulary," available from http://www.ultimatevocabulary.com/ for the low price of $67. It promises "Instant Respect and Credibility"!
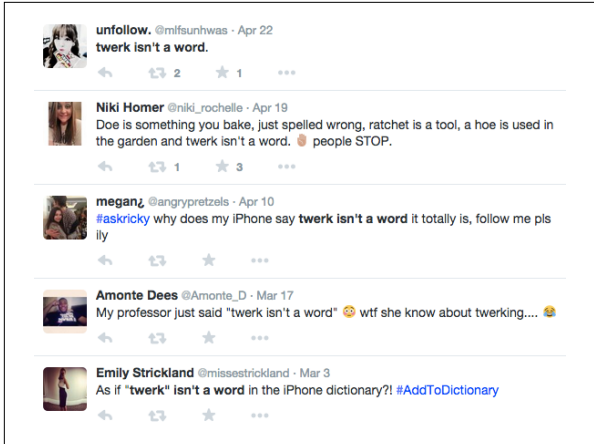
But what does "vocabulary" mean, really? Language changes *constantly*. New words are created all the time, and other words fall out of favor. Today, we can see these changes in real time: Here's a Twitter search for "twerk isn't a word" that demonstrates the process of a word coming into fashion. This demonstrates that you don't need to buy a book or download software to "build your vocabulary"—it's something we do collectively every day in the course of normal language use. We recognize new patterns of language use and incorporate them into our own talk.

Stay Cybersafe @SSCyberbullying · May 20
@GlobalCalgary @c_todd No.... this can't be possible!? **Selfie isn't a word**..

View conversation

reb @ribeeno · May 18
I can't even say **'selfie' isn't a word** because it's been added to the Oxford dictionary fuck knows why.

thatblueperson @thatbleuperson · May 12
WAIT. It says **selfie isn't a word**. Wasn't this crap added in the dictionary a bit ago?

Harris Hennig @harrishennig · May 12
"Selfie" isn't a word and it's swishy-sounding anyways

Here's the problem: Even when new words *are* added to dictionaries, to "official" word lists, the words that get added tend to be words from certain groups. Words from marginalized groups get ignored (or appropriated).

Here's a search for "selfie isn't a word." When the ever-ecumenical OED adds a word from a marginalized group, everyone gets up in arms about it, as you can see in this screenshot.

Which is to say: "official" word lists and dictionaries usually don't serve to catalog language, as much as they serve as a way to *exclude* certain kinds of language.

In sum: Having a "large vocabulary" usually doesn't mean "knowing a lot of words"—it means "knowing (and using) the *right* words."

So: both "spelling" is arbitrary, and "large vocabulary" is really just code for "privilege." Now what?

But! People *love* to play with words

Despite all this, people love to play with words.

Game designers who make word games have a leg up on everyone else, because nearly every human acquires a language fluently and has tremendous creative capacity in using that fluency. Language is an incredibly rich set of skills to draw from for games. So rich, in fact, that word games and games about spelling are incredibly popular! Everyone has expertise in language, and people love to play with things they have expertise in.

I Spy
Geography
Hangman
Ghost

Language is a common topic for folk games—here are just a few.

English Pig-Latin

Scrabble sucks → Abblescray uckssay

Luchazi (Bantu) word game

tuye kundzivo 'Let's go to the house' → yetu kundzivo

Zande (Sudanese) word game

baramu 'European' → rabamu

As further demonstration: There's a whole collection of language games like pig-latin, played cross-linguistically. Here are some examples from Luchazi and Zande that work very similarly to pig latin. There's something deep down in our brains that makes us want to do fun things with words.

Davis, Stuart (1993) "Language Games". The Encyclopedia of Language and Linguistics, Pergamon Press: Oxford and New York, pp. 1980-1985. http://www.indiana.edu/~lingdept/faculty/davis/LanguageGamesENC.pdf

And of course, the largest proof that word games are popular: people KEEP BUYING SCRABBLE—even if it makes them feel bad.

Toward word games for people, not a**holes

So what are some ways of making word games that make use of people's propensity for language play, but don't force them to be jerks?

Spelling ability and vocabulary size aren't related to intelligence

So the first thing for game designers to remember is this. It's safe to design games that use words without making this assumption.

...which is why people use assistive technology. Spell check and autocomplete aren't signs of a decadent society: they're helpful tools that smart people use to help make language (and in particular, English) easier to use. Given all the problems with English spelling and vocabulary that we've talked about today, this isn't surprising.

Word games that spell-check and autocomplete themselves

So here's a solution to these problems that I've been experimenting with. Make word games that spell-check and autocomplete themselves. Essentially: make games that AFFORD good spelling, instead of assuming it. Make games that make it easy for people to form long, familiar words.

Rewordable

Card game for 2+ players

Co-designed with Adam Simon and Tim Szetela

Based on the most common unigrams, bigrams and trigrams in English

Cards freely available at rewordable.com

The first experiment is... what happens if you just change the unit from unigrams to higher order n-grams? Rewordable is a game I helped design with my friends Adam and Tim a few years ago. It's a card game played with a deck of 160 cards, each of which has a unigram, bigram or trigram on it. ("Bigram" and "trigram" just mean groupings of two or three letters, respectively.) The bigrams and trigrams were selected because they're the most frequent of their kind in English words—sequences like "ing" or "er."

The idea is that players will be able to form longer, more satisfying words, because the sequences of letters on the cards themselves are longer. Having longer sequences of letters on the cards guide spelling instead of working against it, and make it more difficult for players to fall back on words densely packed with rare letters (smart-ass words). Here are some action shots. We're still working on getting the rules just right but I think by and large it accomplishes the goal of encouraging fluent word creation without the frustrations of scrabble. LOOK FOR A KICKSTARTER SOON.

Lexcavator is a video game I made. It's a cross between Boggle and Mr. Driller—you find and select words to clear them from the board, allowing your at-symbol to progress further into the game. To ensure that the words you find are interesting and fun, the board is generated with a statistical process called Markov chain generation. To quickly demonstrate—

```
R   R   A   C   F

E   A   T   H   U

C   A   T   R   O
```

Here's a simplified diagram of how Lexcavator board generation works. It starts with a few rows of random letters (weighted by English frequency), just to prime the process. Then for each cell in the next row, we randomly select a column of letters connected straight up or diagonally and populate that cell with a letter randomly, and then insert the letter most likely to occur in English words *following* the two letters above it. The net effect of this is to *guide* the player in making longer, more common English words, instead of working against them.

Corpus analysis

- ~15000 games downloaded from cross-tables.com, an archive of Scrabble game transcripts, containing a total corpus of ~670k words played

- ~620k words played in online sessions of Lexcavator, dumped from MongoDB

I collected a corpus of Scrabble games and a corpus of Lexcavator games and compared the two, graphing how often more "common" words were played in each game. I have no good reason to assume that the cross-tables data is representative of all Scrabble ever, but it seems like a reasonable place to start.
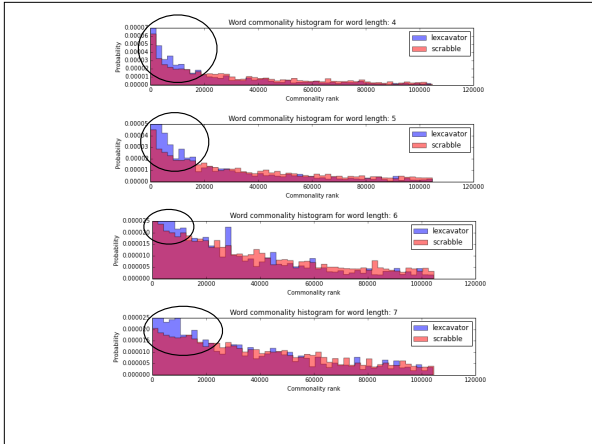
This is a graph of word frequency in the corpus, showing how often words with a particular commonality in English are found in each game. Each graph represents a different word length (4-letter words, 5-letter words, etc.) Frequency of common words is on the left of the graph, frequency of less common words on the right. The areas of solid blue are where Lexcavator's words are distributed, and the areas of solid red are where Scrabble's words are concentrated. You can see that Lexcavator's words are more bunched up toward the left of the graph, and Scrabble's are more evenly distributed across the entire range. So Lexcavator does indeed encourage people to form more common words. Success!

Thanks!

http://twitter.com/aparrish

http://www.decontextualize.com/

http://www.lexcavator.com/

http://www.rewordable.com/

A big thank you to the organizers of the conference for everything. It's been a privilege and a pleasure to participate.