Decoding Multi-Clause Initialisms in Online Discourse

Jalyn Hancock

Dr. Patricia Cukor-Avila (Mentor and Co-editor)

Abstract

This study examines the extent to which native English speakers can understand lengthy initialisms with little to no context. In 2020, popular tweets on Twitter revealed that, when given a string of as many as 21 letters such as "ilybidkhttybikydlmb"[1] (each letter being the first letter of a word in the sentence) with little context, some people are able to decipher the meaning, while others cannot. The research questions for this study are: (1) To what extent are native English speakers able to understand the meanings of multi-clause initialisms without prior knowledge of their meanings? (2) How do age, gender, and amount and focus of Internet usage affect the understanding of these initialisms? Previous studies have investigated the significance of initialisms and their online use (McCulloch, 2019) as well as their importance to the physical world (Pop and Sim, 2009). However, there is a lack of substantial research concerning the understanding of multi-clause initialisms that are outside the sphere of pre-established abbreviations (such as *smh* and *ttyl*). A questionnaire was distributed and results from 26 participants were analyzed to determine if there is a viable link between aspects of one's lifestyle or cultural background and the ability to understand multi-clause initialisms. It was hypothesized that age and Internet usage will affect the understanding of the initialisms. The results suggest that age and gender do not affect the overall ability to decode initialisms but could affect the

ability to decode specific types of initialism. High usage of the Internet for social communication increased the likelihood of a participant to correctly decode the multi-clause initialisms.

Introduction

Initialisms are a type of abbreviation that uses and pronounces the first letter of each word in a phrase (for example, some pre-established/foundational abbreviations on the Internet are *ttyl*, *tbh*, and *brb*). The use of these shortened means of expression has evolved dramatically since their inception, adding increasingly complex meanings to their surroundings, and prompting the need for more study on their expansion. Since the development of the Internet, initialisms have shifted from primarily professional or technical language settings to more common, everyday language contexts. In past studies, researchers noted attitudes towards abbreviations in formal and informal environments (Squires, 2010) as well as how the abbreviations' meanings depend on their context (Pop & Sim, 2009), some of which will be discussed in the following section.

As with many aspects of communication, the rules of word-shortenings change when applied to an online platform. Instead of being used solely for military-related terms (like AWOL) or reducing lengthy names (like NAFTA for North American Free Trade Agreement), they are also now being used to save space for complete thoughts. Internet speech, sociolinguistics, syntax, semantics, and cognitive linguistics all contribute to lengthy, multi-clause initialisms that can be somewhat understandable with little to no context, as they are the building blocks to create such initialisms. This study will examine the extent to which native English speakers can understand lengthy initialisms with little to no context and if factors, such as Internet use and age, affect this ability.

Literature Review

Although initialisms are used in English to shorten text, the Internet and the surrounding culture have incorporated initialisms into everyday speech on the platform, to the displeasure of many English speakers who do not use initialisms in this manner. As Internet usage has increased, its language has become more popular, spreading beyond the screen. Squires (2010) states that adults have expressed concern about the effects the language might have on Millennials in their teenage years. She referenced multiple newspapers from the late 1990s and early 2000s, and found that the general attitude towards the abbreviations, whether used in the appropriate setting or not, expressed negative sentiments since they deviated from what is considered "Standard English" (Squires, 2010, p. 460). She concluded that the enregisterment of Internet language depends on the existing ideologies within that specific technological and cultural sphere.

Nwala and Tamunobelema (2019, p. 12), however, found that students tend to have a static use of acronyms and word-shortenings that are mostly utilized informally online, not showing any variance despite a change in the use of the text (i.e., an academic essay vs. a text to a friend) or mood of the conversation. While the study does not present an explanation for the students' language usage on Facebook, the researchers claimed that social media shapes individuals and society in a way that goes against pedagogy, as the language system that they use is "open and unconventional" (Nwala & Tamunobelema, 2019, p. 12). They stated that acronyms and neologisms that are mainly used on social media are generally not agreed-upon and that they are "linguistically unpredictable" in that they can have different meanings and be interpreted differently depending on the individual (Nwala & Tamunobelema, 2019, p. 10). An earlier study (Merchant, 2001) conversely argued that using the new language is a marketable

skill that could be helpful in the new age of technology and presented from a positive perspective.

Abrahão (2014) collected data from a more diverse age group that exhibits a sense of self-awareness amongst the younger participants, namely those with whom the new language is most commonly associated, even if they were not likely to use the specific examples given in the study. A large percentage of the respondents acknowledged that certain sentences containing abbreviations unique to the Internet are viewed as adolescent-like or representative of "modern" society and do not necessarily have "intelligent" characteristics. For example, when asked what they thought about the sentence "tbh, idkwat 2 do w dem……..", which is translated as "to be honest, I don't know what to do with them," most of the participants responded that they viewed it as "*teenager*" and "*lazy*" (Abrahão, 2014, p. 116). Although the descriptive options they were given to express their perception of the users are not extensive nor necessarily explanatory (some options were "adult," "teenager," "modern," and "lazy") and there was no verification of their understanding of the sentences, the results of the questionnaire supported the consensus that abbreviations are not accepted in all mediums, especially when misused.

It is true that initialisms belong to particular environments. When used out of context, they can have meanings other than the one intended. Pop and Sim (2009) explored different abbreviations in business English and how they vary among contrasting fields and departments (such as ELA having the meanings "Employment Law Alliance," "Enterprise License Agreement," "Excess Loss Account," and even "English Language Arts"). They claim that acronyms are not predictable since any part of the phrase can determine them, meaning that they can exclude conjunctions and focus on content words, such as CARE from Cooperative for Assistance and Relief Everywhere, or take letters that are not at the beginning of a word and

include them in the abbreviation, such as the "N" in PINC from Property Income Certificate (Pop & Sim, 2009, pp. 557-558). Pop and Sim (2009, p. 559) inspected the formation of numerous types of abbreviations, and they concluded that the chosen abbreviation is based on the purpose of the shortening, either to make it easier to pronounce when expressed as a word (e.g., CAPEX from *Capital Expenditure*) or simply to shorten it (e.g., CEO from *Chief Executive Officer*).

Beyond the surface level of initialisms are the cognitive elements behind the shortening of words and phrases. Ungerer (1991) explored the loss of meaning that occurs when terms become acronyms (e.g., WASP and SPOT) as opposed to other types of word formation, particularly in trade names (e.g., Quic and Tide). He found phonological and visual motivations that evoke certain emotions, and the choices made concerning the type of abbreviation used depend mainly on the end goal, whether it be for semantic or graphic support.

When understanding such abbreviations, people tend to undergo top-down and bottom-up processing. Top-down processing occurs when one's perception goes from broad to specific, while bottom-up is from specific to broad. Researchers have debated which is the most important type of processing for reading skills (Andrews, 2009). Bottom-up is argued to be critical for the foundation of an advanced working memory and reading comprehension built by word identification and the ability to decode (Stanovich, 2000). In contrast, those who support top-down processing claim that decoding makes comprehension too difficult by requiring the processing of small, abstract details of language rather than understanding the overall meaning (Goodman, 1986). For example, Tat and Azuma's (2016) study on top-down processing with and without context when reading text messages of various styles (such as deletion and substitution) found that sentences with omitted vowels cause the least amount of disruption in recognition due

to the minimal amount of information missing required to understand the words; also, false recognition of abbreviations was usually higher for abbreviation types that looked more similar.

Although there has not been extensive research into the current language of the Internet, especially due to the rate at which it continues to grow and change, studies on its foundation can provide an understanding of how initialisms are used in a manner unlike that of non-Internet English. McCulloch (2019), for example, covers numerous aspects of Internet language, such as the typography and community surrounding it. Through her exploration of initialisms, emojis, memes, and the effects of punctuation in general online discourse, she attempts to break down and explain the culture that they shape. While allowing for more concise language use, the ambiguity of initialisms can create confusion when not used in the proper context, and even more so on the Internet. However, initialisms may potentially be understood with minimal context, as this study will explore, as well as the extent to which native English speakers can make sense of them and whether background characteristics aid in their understanding.

Methodology

The study was conducted using an online survey with a descriptive mixed-methods design, similar to the technique used by Nwala and Tamunobelema (2019). Participants were gathered via Instagram, Twitter, and Facebook posts promoting the survey and word of mouth. The participants were native English speakers older than 18 years of age and of any socioeconomic status, race, sexuality, gender, and nationality. The survey asked participants about their Internet use and presented four strings of initialisms for them to decode, some with pre-established abbreviations (e.g., *lol*, *smh*, *lmao*). Each string was shown with and without context relevant to a particular background, as shown in Table 1. The strings were found on

social media sites, though mainly Twitter, and participants were asked if they had prior knowledge of each string's meaning.

Table 1. "Initialisms"

Initialism	Meaning	Context
ilysmbidkhttybikydlmb	"I love you so much but I don't know how to tell you because I know you don't love me back"	If you can read this: "ilysmbidkhttybikydlmb," you're broken
gyaitgdhbibymfa	"Get your ass in the goddamn house before I beat your motherfucking ass"	If you're black, you should know what this stands for: "GYAITGDHBIBYMFA"
llabocynkwygg	"Life's like a box of chocolate, you never know what you're gonna get"	If you love movies, you should know what this means: "llabocynkwygg"
ijwytbwmftroml	"I just want you to be with me for the rest of my life"	Only loyal people can read this: "ijwytbwmftroml"

The participants were first required to state what they believed to be the string's meaning and were then given a list of possible answers to select from that were most like their initial response. The answer choices were either alternatives provided by Twitter users who responded to the original tweet or choices that were created for the survey. They were then asked to explain how they came to the answer they submitted (e.g., focusing on a known abbreviation and

working around it to make the rest of the letters make sense) through multiple choice answers and whether the context, when available, was relevant or not as well as if it helped them understand the meaning. The responses were analyzed to determine how much of the string they understood and if there is a connection between the demographic information they provided and the selection of correct answers.

Results

In total, twenty-six viable survey responses were collected. Eight men and eighteen women completed the survey. The respondents' ages are categorized by birth year and placed into two groups: between 1955-1980 (8 respondents) and 1981-2003 (18 respondents). The races of the respondents are as follows: twelve identified as Black, five as White, four as Hispanic or Latino, three as mixed, one as Middle Eastern, and one as East Asian.

The categories for the answer choices of the initialisms' perceived meanings were created based on the responses. Each question shared the same three options: Y (they gave the correct answer), 0 (they did not give an answer), and Other (they gave an answer that was not related to the correct meaning). They may also have a category labeled "Sim." for "similar," meaning that their answer was very close to the actual meaning but was not exact (such as putting "like" instead of "love" for "ilysmbidkhttybikydlmb"). If a category has "part" next to it, the answer was partially complete. Only respondents that indicated that they had NOT known the meaning of the initialism prior to taking the survey were included in the findings.

Age

The percentage of correct responses by age group was not consistent across questions, as seen below in Figures 1-4.

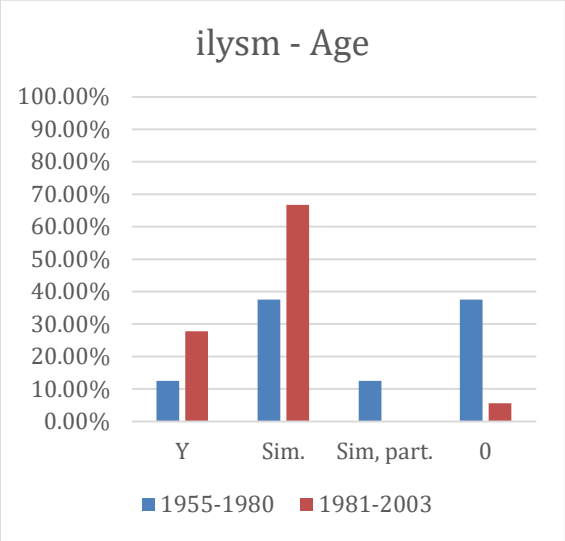

Figure 1. “ilysm - Age”

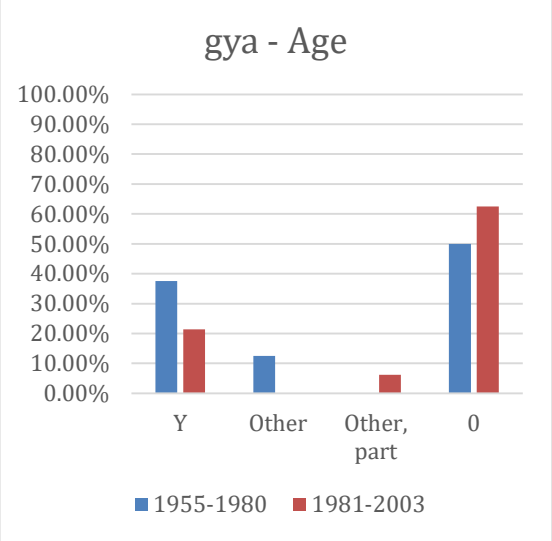

Figure 2. “gya - Age”

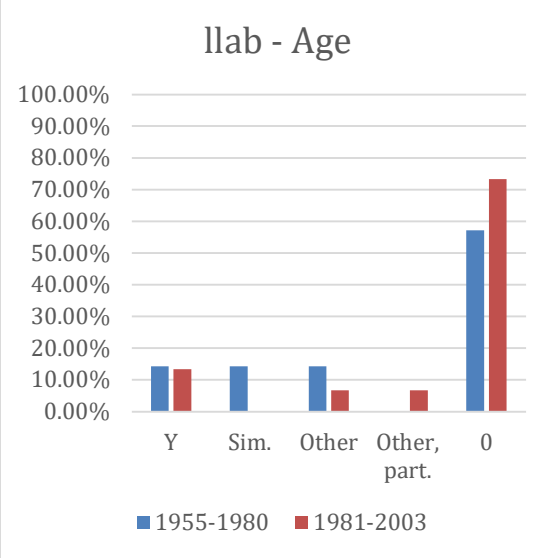

Figure 3. “llab - Age”

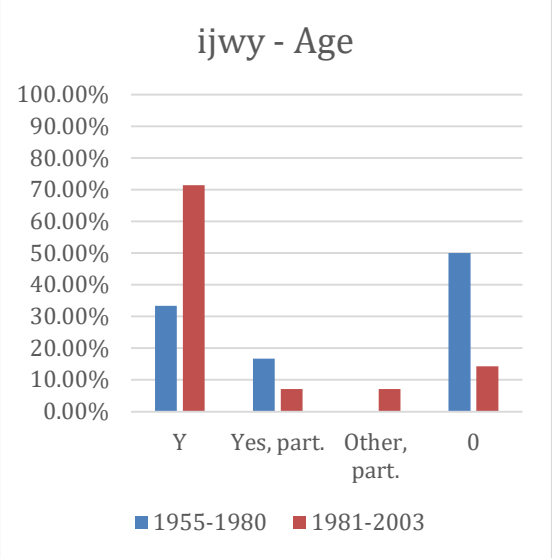

Figure 4. “ijwy - Age”

Younger respondents were more likely to get “ilysmbidkhttybikydlmb” and “ijwytbwmftroml” correct, the two initialisms that the older participants were more likely not to answer. However, older respondents answered “gyaitgdhbibymfa” correctly more often, which

was one of the two initialisms that the younger respondents were more likely to not answer. They both had around the same rate of correct responses for "llabocynkwygg," but the younger respondents were more likely not to answer it.

Gender

For two of the four initialisms presented, there were no notable differences in the success rate of a respondent according to gender: "gyaithbibymfa" and "llabocyknwygg." However, with the other two ("ilysmbidkhttybikydlmb" and "ijwytbwmftroml"), men were more likely not to answer, and women were more likely to get it correct, as seen in Figures 5-8.

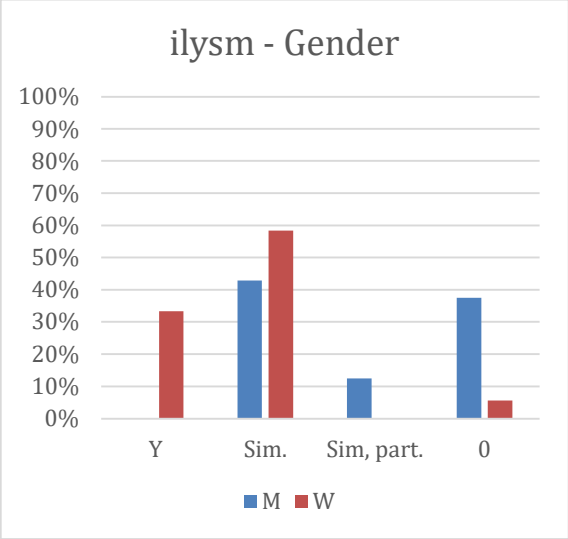

Figure 5. "ilysm - Gender"

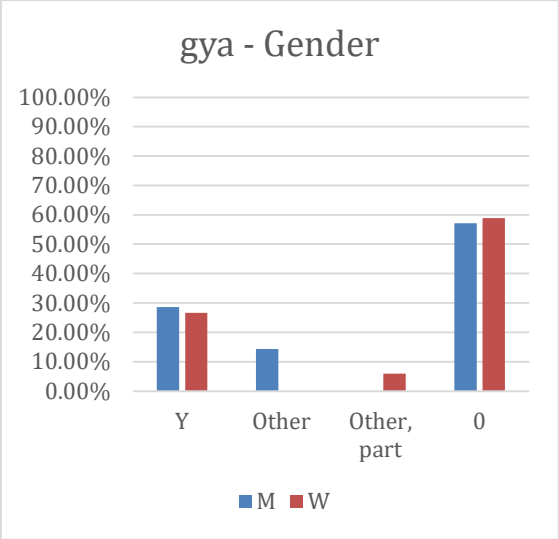

Figure 6. "gya - Gender"

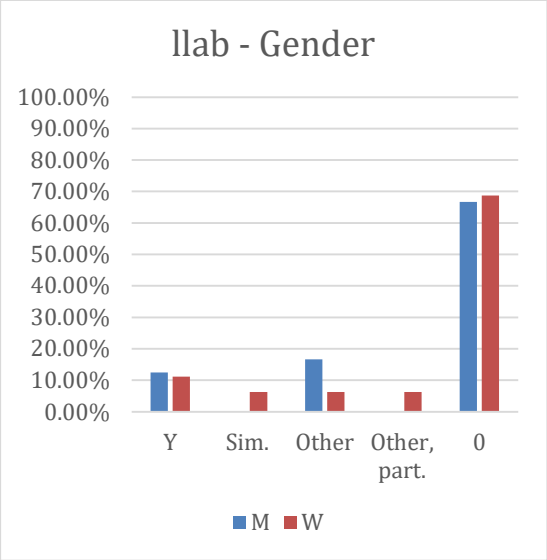

Figure 7. “llab - Gender”

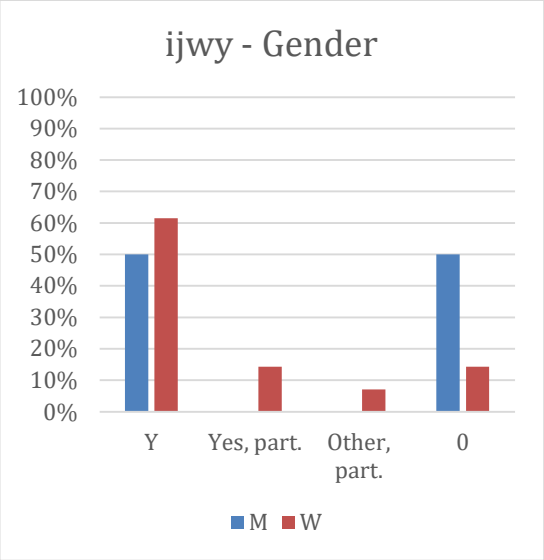

Figure 8. “ijwy – Gen

Amount of Internet Use

The participants were asked to state the amount of time they spent on the Internet every day for social communication. They were then put into three categories based on their Internet usage: “Low” for less than an hour a day, “Medium” for between one and five hours, or “High” for at least six hours a day. Figures 9-12 show, per initialism, the percentage of answer choices provided by participants in each category.

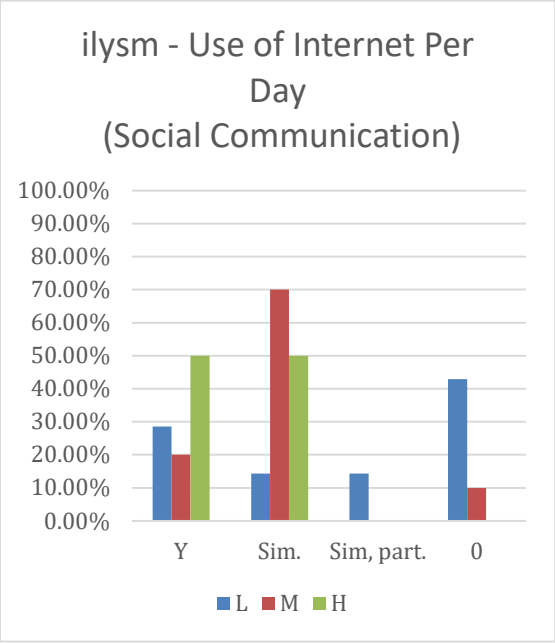

Figure 9. “Ilsym - Use of Internet Per Day”

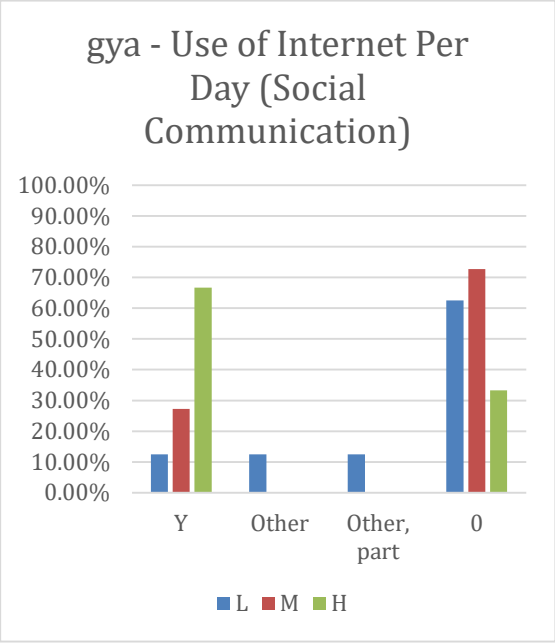

Figure 10. “gya - Use of Internet Per Day”

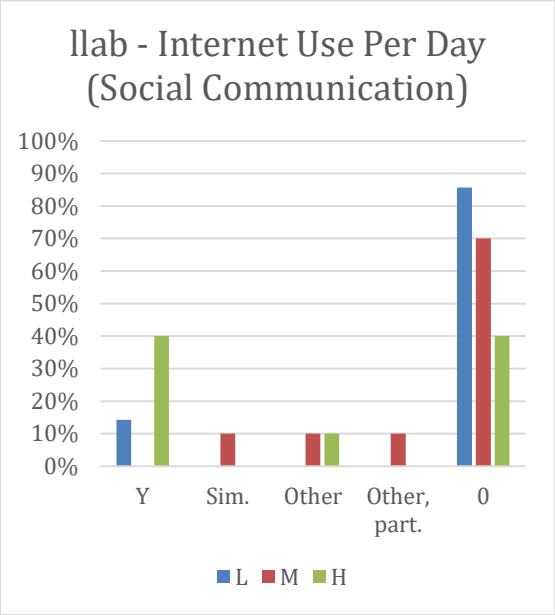

Figure 11. “llab - Use of Internet Per Day”

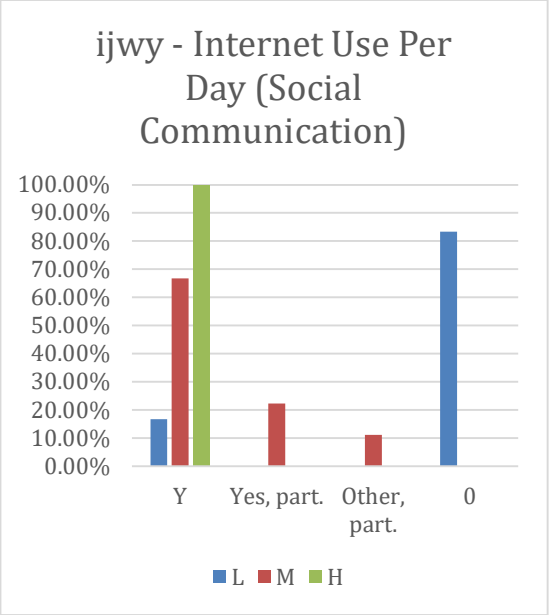

Figure 12. “ijwy - Use of Internet Per Day”

For each initialism, those who used the Internet for social communication for more than six hours a day were more likely to decode the initialism correctly and least likely to state that they weren’t able to decode it. The participants who had the least amount of Internet usage were not necessarily least likely to be able to understand the initialisms, as seen in Figure 9, but they were more likely to not provide an answer, except for the initialism “gyaitgdhbibymfa” in Figure 10.

Discussion

The most notable results seem to be the similarities between “llabocynkwygg” and “gyaitgdhbibymfa,” and “ilysmbidkhttybikydlmb” and “ijwytbwmftroml.” In both the gender and age categories, “llabocynkwygg” and “gyaitgdhbibymfa” have the most participants who could not create a meaning for them. The lack of pre-established abbreviations (i.e., “lol,” “brb,” etc.) could be a possible factor. Although “mf” could be a standalone abbreviation, its position later on in the initialism may have affected their ability to recognize it. Additionally, even though

“ilysmbidkhttybikydlmb” and “ijwytbwmftroml” had more correct or similar responses, “ijwytbwmftroml” does not contain any pre-established abbreviations.

Women were also more likely to decode the more emotional initialisms (“ilysmbidkhttybikydlmb” and “ijwytbwmftroml”) correctly or provide a similar answer, though the disparity in the amount of self-identifying men and self-identifying women makes it difficult to proclaim a strong relationship between the two.

The older participants’ ability to decode “llabocynkwygg” and “gyaitgdhbibymfa” could be attributed to the content of the initialisms. Although “llabocynkwygg,” representative of a quote from the famous 1994 movie “Forrest Gump,” is a known quote amongst many adults, the age at which the movie was seen, or when the quote became known to the participant, may have impacted their recollection of it. The results for “gyaitgdhbibymfa” could be seen as a reflection of the change in stereotypical Black parenting styles over the past several decades. When given the context of “If you are Black, you should know what this means,” paired with the initialism, some participants stated that, while it is known that not all individuals of a particular race speak the same way, they were able to use stereotypes to help guide them in their understanding of the initialism.

Regarding the other initialisms, however, most participants were of the opinion that having context or background information did not help. Instead, they usually chose to go letter-by-letter until the string made sense, or they focused on identifying the abbreviations that they knew and then made the letters around them make sense until they completed the string.

Although no concrete conclusions can be drawn from this study due to the limited amount of data, the results suggest that native English speakers are more likely to be able to decode initialisms that contain content they are already familiar with.

Implications

The increased use of multi-clause initialisms could show how chunking is used in the processing and storing of information. With the results of the survey suggesting that there are differences based on age, gender, and Internet use, there could be components of each category that affect how the brain is able to understand a string of what many individuals may initially perceive to be a random set of letters.

This study investigated several possible influences on the ability to decode initialisms. Future studies could focus on the breakdown of particular initialisms (such as those using similar abbreviations or containing similar cultural references) and how native English speakers decode them. The ability of non-native English speakers could also be tested in a similar manner and compared to the competencies of native speakers. The use of eye-tracking could be helpful with regards to detecting the order in which an initialism is decoded (starting at the beginning or with pre-established abbreviations) and how long it takes an individual to decode it.

References

Abrahão, V.D. (2014). U ok w dis?: Analysing language attitudes towards internet english. *Entrepalavras*, *4*(2), 105-128.

Andrews, S., & Bond, R. (2009). Lexical expertise and reading skill: Bottom-up and top-down processing of lexical ambiguity. *Reading and Writing, 22*(6), 687-711. doi:http://dx.doi.org.libproxy.library.unt.edu/10.1007/s11145-008-9137-7

Goodman, K. S. (1986). *What's whole in whole language? A Parent/Teacher guide to children's learning* Heinemann Educational Books, Inc., 70 Court St., Portsmouth, NH 03801 ($7.50). Retrieved from https://libproxy.library.unt.edu/login?url=https://www-proquest-

com.libproxy.library.unt.edu/books/whats-whole-language-parent-teacher-guide/docview/63091606/se-2?accountid=7113

McCulloch, G. (2019). *Because Internet: Understanding the New Rules of Language*. Riverhead Books.

Merchant, G. (2001). Teenagers in cyberspace: an investigation of language use and language change in internet chatrooms. *Journal of Research in Reading*, *24*(3), 293. https://doi-org.libproxy.library.unt.edu/10.1111/1467-9817.00150

Michael, A. N., & Tamunobelema, I. (2019). The social media and language use: The case of facebook. *Advances in Language and Literary Studies, 10*(4), 9-13. doi:http://dx.doi.org.libproxy.library.unt.edu/10.7575/aiac.alls.v.10n.4p.9

Pop, Mirabela & Ariana Sim. (2009). THE USE OF ACRONYMS AND INITIALISMS IN BUSINESS ENGLISH. Annals of Faculty of Economics. 1. 557-562.

Squires, L. (2010). Enregistering internet language. *Language in Society, 39*(4), 457-492. doi:10.1017/s0047404510000412

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers* Guilford Publications, Inc., Dept. 3R, 72 Spring St., New York, NY 10012 (Catalog no. 0565: $38). Retrieved from https://libproxy.library.unt.edu/login?url=https://www-proquest-com.libproxy.library.unt.edu/books/progress-understanding-reading-scientific/docview/62158931/se-2?accountid=7113

Tat, M., & Azuma, T. (2016). Remembering “txt” but not “text”: The effect of context and lexicality on memory for text message abbreviations. *Applied Psycholinguistics, 37*, 435-459.

Ungerer, F. (1991). ACRONYMS, TRADE NAMES AND MOTIVATION. [Abstract] *AAA: Arbeiten Aus Anglistik Und Amerikanistik, 16*(2), 131-158. Retrieved November 15, 2020, from http://www.jstor.org/stable/43023574